Microbiology (1998), 144, 1387–1396 Printed in Great Britain

Implications of the *xylQ* gene of TOL plasmid pWW102 for the evolution of aromatic catabolic pathways

Sirinun Aemprapa† and Peter A. Williams

Author for correspondence: Peter A. Williams. Tel: +44 1248 382362. Fax: +44 1248 370731.
e-mail: P.A.Williams@bangor.ac.uk

School of Biological Sciences, University of Wales, Bangor, Gwynedd LL57 2UW, UK

Pseudomonas putida **strain O_2C_2 is able to grow on toluene, *m*-xylene and *p*-xylene through benzoate and the corresponding methylbenzoates (toluates). The catabolic genes are encoded on a large TOL plasmid, pWW102, of >220 kb. The complete catabolic genes were cloned on four large overlapping restriction fragments covering a total of 28 kb of the plasmid, which was carefully mapped by restriction enzyme analysis. The presence of the *xyl* genes on the cloned DNA was confirmed by assay of representative enzymes of both operons. Virtually all the genes were located on the cloned DNA by hybridization of Southern blots with gene-specific probes from related pathways of other catabolic plasmids. Within the limitations of available restriction sites, the analysis showed that the genes are in two blocks. The major block carries the *meta* pathway operon *xylXYZLTEGFJQKIH* with the two regulatory genes *xylSR* immediately downstream. The upper pathway operon *xylUWCMAB(N)* is about 2–3 kb downstream of the regulatory genes and transcribed in the same direction as the *meta* pathway operon. Within each operon the gene order appears to be identical to that found in other TOL plasmids, but the relative location of the operons most closely resembles that found on plasmid pWW53, although there is no evidence of any *xyl* duplications on pWW102. The nucleotide sequence of the *xylQ* gene for the acetaldehyde dehydrogenase (acylating; ADA), together with the 3′-end of the upstream *xylJ* (for 2-oxopent-4-enoate hydratase) and the 5′-end of the downstream *xylK* (for 4-hydroxy-2-oxovalerate aldolase), was determined. The *xylQ* gene was ligated into expression vector pTrc99a and high levels of XylQ protein were detected by enzyme assay and by SDS-PAGE. All three genes *xylJQK* showed a high degree of homology with genes encoding isofunctional proteins from other *Pseudomonas meta* pathways, the highest being with the naphthalene catabolic genes *nahLOM* from the plasmid of *Pseudomonas* sp. NCIB 9816. The implications of the sequence homologies to the evolution of these pathways are discussed.**

Keywords: TOL plasmid, *Pseudomonas putida*, toluene/xylene catabolism, acetaldehyde dehydrogenase (acylating), *xylQ*

INTRODUCTION

The pathway (Fig. 1) for catabolism of toluene and some substituted toluenes, such as *m*- and *p*-xylenes, in *Pseudomonas* appears to be almost always plasmid-determined (Williams & Worsey, 1976) and shares many features with the catabolic pathways of other aromatic compounds (Assinder & Williams, 1990; Harayama & Rekik, 1993). The genes encoding the catabolic pathway on the archetype of these TOL plasmids, pWW0, have

†**Present address:** Lampang Agricultural Research and Training Centre, Rajamangala Institute of Technology, PO Box 89 Maung, Lampang 52000, Thailand.

Abbreviation: ADA, acetaldehyde dehydrogenase (acylating).

The GenBank accession number for the sequences reported in this paper is AF043925.

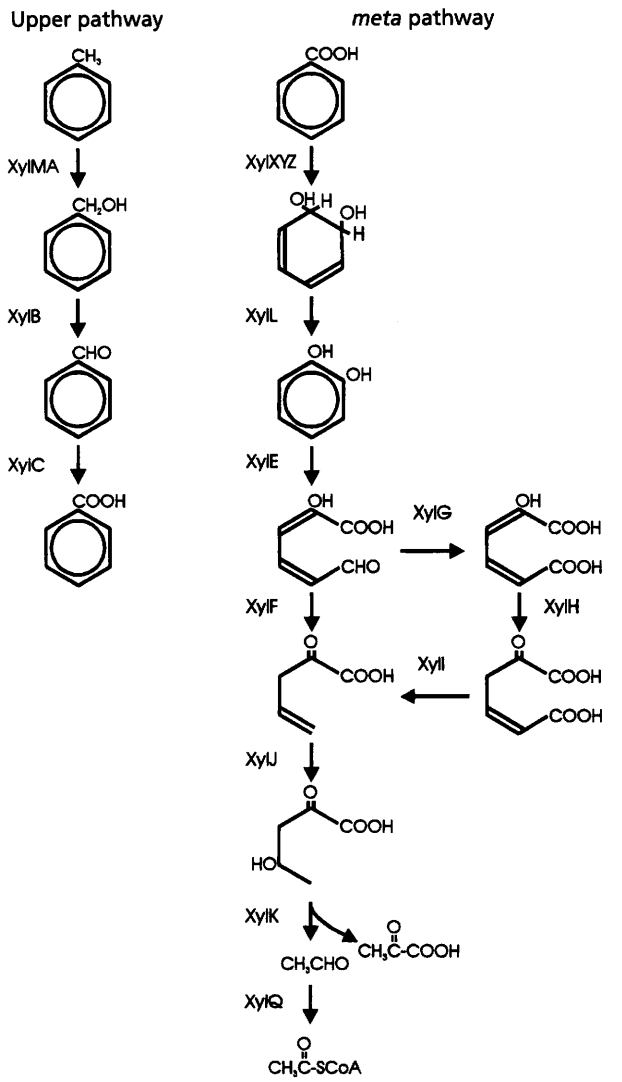

Fig. 1. *xyl* catabolic pathway encoded by TOL plasmids. The enzymes catalysing each of the transformations are denoted by their gene-associated nomenclature. *xyl* genes not encoding known metabolic reactions are *xylU,W,N,T,R,S*. The function of XylT is proposed as being involved in reactivating oxidized catechol 2,3-dioxygenase (XylE) (Polissi & Harayama, 1993). XylR and XylS are positive regulatory proteins. The functions of XylU, XylW and XylN are not known.

been studied in great detail (see Assinder & Williams, 1990) and a number of other TOL plasmids have been investigated to a lesser extent (Pickup & Williams, 1982; Keil & Williams, 1985; Keil *et al.*, 1985; Osborne *et al.*, 1988). They all share a common organization into two operons: an 'upper pathway' operon, *xylUWCMABN* (Harayama *et al.*, 1989; Williams *et al.*, 1997), encoding the sequential oxidation of the methyl substituent to a carboxylic acid (e.g. toluene to benzoate), and a '*meta* pathway' operon, *xylXYZLTEGFQKIJH* (Harayama & Rekik, 1990), encoding the conversion of the carboxylic acids to central metabolites. Whereas the gene order within each of the operons appears to be conserved, the relative orientations of the two operons and the distances between them differ from plasmid to plasmid, and most carry duplications of structural and/or regulatory genes (Chatfield & Williams, 1986; Osborne *et al.*, 1988; O'Donnell & Williams, 1991).

Pseudomonas putida O_2C_2 is a bacterium isolated from soil from the Netherlands which is able to grow on toluene and methyl-substituted toluenes and which contains a large TOL plasmid, pWW102. This paper reports the cloning of the *xyl* catabolic genes from pWW102. The nucleotide sequence of *xylQ*, for acetaldehyde dehydrogenase (acylating; ADA), is presented and the implications of comparison with genes for other ADA proteins to the evolution of aromatic pathways is considered.

METHODS

Maintenance of bacterial strains and plasmids. The strains and plasmids used in this work are listed in Table 1. Strains were routinely maintained on nutrient agar or on sensitivity test agar (Lab M) supplemented with appropriate antibiotics: kanamycin at 15 μg ml^{-1} and ampicillin at 100 μg ml^{-1}.

Enzyme assays. Cell-free extracts, other than for ADA, were prepared according to Cane & Williams (1982). The assays for catechol 2,3-dioxygenase (XylE) and for 2-hydroxymuconic-semialdehyde dehydrogenase (XylG) were carried out as described by Sala-Trepat & Evans (1971) and for 2-hydroxymuconic-semialdehyde hydrolase (XylF) according to Duggleby & Williams (1986). For the ADA (XylQ) assay, *Escherichia coli* containing recombinant plasmids bearing *xylQ* was inoculated at 37 °C for 5 h in LB, when IPTG was added for induction of the $lacI^q$-regulated *tac* promoter of pTrc99a vector, and the cultures were incubated overnight at 30 °C in the presence of IPTG. Extracts were prepared and assayed for ADA activity by the method of Shingler *et al.* (1992).

Protein electrophoresis. Cell-free extracts containing ADA were prepared as described above. SDS-PAGE was carried out by standard procedures in a PROTEANII cell (Bio-Rad).

DNA extraction and manipulation. Shotgun cloning was performed using DNA of plasmid pWW102 which was prepared by the sucrose gradient method of Wheatcroft & Williams (1981). The vector chosen for cloning was pUC18. Restriction endonuclease digestions and DNA ligations with T4 ligase were performed by standard procedures (Sambrook *et al.*, 1989) and recombinant transformants were selected on media containing antibiotics appropriate to the vector. Plasmid DNA was prepared for rapid screening by the method of Holmes & Quigley (1981) and purified by CsCl centrifugation (Guerry *et al.*, 1973) or using Qiagen kits. Colonies carrying recombinant plasmids expressing catechol 2,3-dioxygenase activity were detected as a result of their yellow colour after spraying with 10 mM catechol (Worsey *et al.*, 1978), whilst other recombinant plasmids were detected by loss of β-galactosidase activity or by colony hybridizations with probes constructed from homologous genes from other bacteria and pathways.

DNA–DNA hybridization. Restriction fragments were separated by agarose gel electrophoresis and transferred by Southern blotting (Southern, 1975) to Biodyne filter membranes.

Table 1. Strains and plasmids used in this study

Strain/plasmid	Relevant details	Source
Strains		
P. putida O_2C_2	Wild-type Mtol$^+$ Mxy$^+$	This study
E. coli NM522		Vieira & Messing (1982)
E. coli JM109		Yanisch-Perron *et al.* (1985)
Plasmids		
pTrc99a	Expression vector	Pharmacia
pUC18	Cloning vector	Vieira & Messing (1982)
pWW102	Wild-type plasmid from strain O_2C_2	This study
Recombinant plasmids		
pWW102-3	17 kb *Hin*dIII fragment of pWW102 in pUC18	This study
pWW102-6	11·6 kb *Bam*HI fragment of pWW102 in pUC18	This study
pWW102-7	8·6 kb *Hin*dIII fragment of pWW102 in pUC18	This study
pWW102-8	8·6 kb *Hin*dIII fragment of pWW102 in pUC18 (reverse orientation)	This study
pWW102-9	8·2 kb *Eco*RI fragment of pWW102 in pUC18	This study
pWW102-10	8·2 kb *Eco*RI fragment of pWW102 in pUC18 (reverse orientation)	This study
pWW102-61	4·0 kb *Bam*HI–*Sac*I fragment of pWW102-6 in pUC18 (coordinates 6·2–10·2 kb)	This study
pWW102-62	3·1 kb *Sac*I–*Sma*I fragment of pWW102-6 in pUC18 (coordinates 10·3–13·4 kb)	This study
pWW102-62E	3·1 kb *Sac*I–*Sma*I fragment of pWW102-6 in pTrc99a	This study
pWW102-622	2·2 kb *Kpn*I–*Sma*I fragment of pWW102-62 in pUC18 (coordinates 10·8–13·0 kb)	This study
pWW102-623	2·1 kb *Eco*RI–*Sma*I fragment of pWW102-62 in pUC18 (coordinates 10·9–13·0 kb)	This study
pWW102-624	1·0 kb *Sal*I fragment of pWW102-62 in pUC18 (coordinates 11·1–12·1 kb)	This study
pWW102-63	3·6 kb *Sma*I fragment of pWW102-6 in pUC18 (coordinates 13·2–16·7 kb)	This study
pVI300	0·87 kb *Bal*31–*Pst*I fragment of pVI150 in pMMB66HE	Shingler *et al.* (1992)
pVI308	1·02 kb *Hpa*I–*Sac*I fragment of pVI150 in pMMB66HE	Shingler *et al.* (1992
pVI309	1·02 kb *Dde*I–*Eco*RI fragment of pVI150 in pMMB66HE	Shingler *et al.* (1992)
pVI310	1·13 kb *Sal*I–*Sal*I fragment of pVI150 in pMMB66HE	Shingler *et al.* (1992)
pVI312	1·14 kb *Sau*I–*Bst*EII fragment of pVI150 in pMMB66HE	Shingler *et al.* (1992)
pVI313	1·08 kb *Not*I–*Xho*I fragment of pVI150 in pMMB66HE	Shingler *et al.* (1992)
pVI314	0·82 kb *Sca*I–*Pvu*II fragment of pVI150 in pMMB66HE	Shingler *et al.* (1992)
pMT100	1·7 kb *Bam*HI fragment of pWW0 in pBR322	J.-L.Ramos, CSIC, Granada, Spain
pWW0-7000	3·9 kb *Kpn*I–*Bam*HI fragment of pWW0	Williams *et al.* (1997)
pVI1:00	2·4 kb *Eco*RV–*Eco*RV fragment in pBluescript	Bartilson & Shingler (1989)
pWW110-4010	2·2 kb *Xho*I–*Xho*I fragment of pWW110 in pUC18	Carrington *et al.* (1994)
pWW110-6003	2·7 kb *Eco*RI–*Eco*RI fragment of pWW110 in pUC18	Carrington *et al.* (1994)
pWW53-3301	17·5 kb *Hin*dIII fragment of pWW53 in pKT230	Keil *et al.* (1985)

Cloned fragments to be used as DNA probes were extracted according to the method of Girvitz *et al.* (1980) and labelling and detection were carried out by the ECL method (Amersham International).

Nucleotide sequence determinations. Sequences of DNA inserts in pUC18 were determined with a Pharmacia DNA Sequencing kit using the appropriate synthetic oligonucleotide primers. The complete sequences of both strands of the *xylQ* region were determined from the sequences of 15 individual fragments. Custom-synthesized oligonucleotide primers were used to determine the sequences across the junctions of subfragments from this region. Some sequencing reactions were performed by Alta Biosciences (University of Birmingham) using an ABI automatic sequencer (Applied Biosystems).

Analysis of DNA and protein sequences. Databases were searched for homologous gene sequences using the BLAST programs (Altschul *et al.*, 1990). Manipulation and analysis of the nucleic acid sequences were done with the DNASTAR (Madison, WI, USA) nucleic acid analysis software package.

RESULTS

Strain O_2C_2 and plasmid pWW102

Strain O_2C_2 was isolated from an oil-contaminated soil in the Netherlands by W. A. Duetz (Institut für Biotechnologie, ETH Hönggerberg, Zürich, Switzerland) by directly plating supernatants from washed soil onto

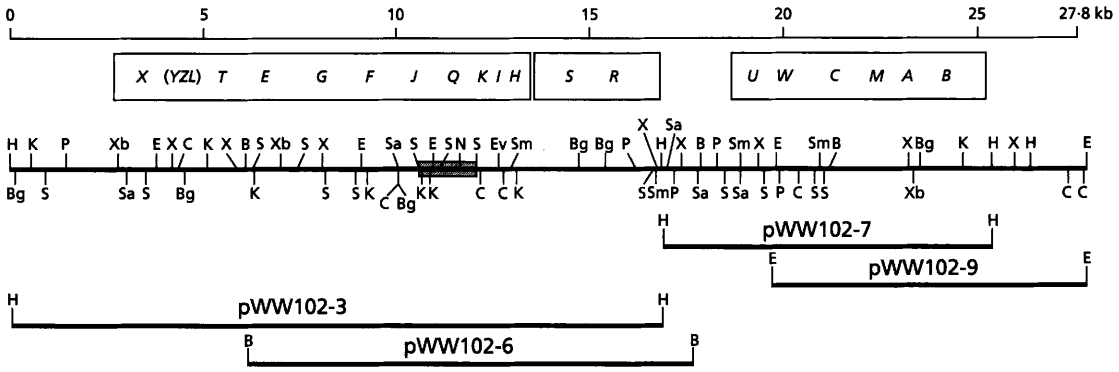

Fig. 2. Restriction map of the complete *xyl* catabolic genes of pWW102. The inserts of the major fragments cloned into recombinant plasmids are denoted below the restriction map. The approximate position of the *xyl* genes is shown above the map: the range of the possible positions of each gene as determined from Southern hybridizations is presented in Table 2. The region sequenced is marked by the grey box. B, *Bam*HI; Bg, *Bgl*II; C, *Cla*I; E, *Eco*RI; Ev, *Eco*RV; H, *Hin*dIII; K, *Kpn*I; N, *Nru*I; P, *Pst*I; S, *Sal*I; Sa, *Sac*I; Sm, *Sma*I; X, *Xho*I; Xb, *Xba*I.

minimal agar plates with toluene vapour as the sole carbon source. It has been classified by the National Culture of Industrial and Marine Bacteria as a *Pseudomonas putida* and it produces a blue-white fluorescence on King's medium B (King *et al.*, 1954). It grows on toluene, *m*-xylene, *p*-xylene, benzoate, *m*-toluate (3-methylbenzoate) and *p*-toluate (4-methylbenzoate): spraying colonies grown on any of these substrates with catechol produced the yellow colour characteristic of the presence of induced catechol 2,3-dioxygenase. It therefore has the phenotype characteristic of TOL-plasmid-containing strains (Worsey & Williams, 1975; Williams & Worsey, 1976). A large plasmid, pWW102, can be isolated from cultures and measurement and summation of the sizes of fragments of the plasmid digested with a range of restriction enzymes indicate it is between 220 and 270 kb. Growth through one cycle of benzoate in liquid culture for about 20 generations, which took 5–6 d, produced loss of the characteristic TOL phenotype (growth on toluene, xylenes, toluates) in 100% of the colonies. All retained the ability to grow on benzoate but no longer produced the yellow colour with catechol or contained any plasmid DNA. No reversion to growth on any of the TOL-specific substrates could be detected.

Cloning of the upper pathway operon

To identify restriction fragments carrying genes of the upper pathway operon, plasmid pWW0-7000, which carries *xylUW* and the 5′-end of *xylC* from TOL plasmid pWW0 (Table 1; Williams *et al.*, 1997), was used as a hybridization probe against Southern blots of digests of pWW102. Three fragments reacted positively: an 8·6 kb *Hin*dIII fragment and two *Eco*RI fragments of 8·2 and 8·7 kb. The appropriate areas were excised from agarose gels, the DNA was extracted and ligated into pUC18, and transformants were checked for hybridization against pWW0-7000 by dot-blotting. Two plasmids with an 8·6 kb *Hin*dIII insert in opposite orientations (pWW102-7 and pWW102-8) and two with an 8·2 kb *Eco*RI insert in opposite orientations (pWW102-9 and pWW102-10) were identified. When the inserts of pWW102-7 and pWW102-9 were mapped for several restriction sites (Fig. 2), comparison of the maps indicated an overlap of about 5·5 kb which was confirmed by cross-hybridization (data not shown).

Functional analysis of upper pathway genes

E. coli containing each of pWW102-7 and pWW102-9 produced pink colonies which turned blue on LB agar plates and produced insoluble blue particles in LB. This reaction is diagnostic for the conversion of tryptophan, present in the medium, to indigo and can be effected by the presence of cloned *xylMA*, encoding xylene monooxygenase (Mermod *et al.*, 1986; Keil *et al.*, 1987). Two enzymes of the upper pathway operon *xylUWCMABN* were identified by direct assay. Significant activity of benzyl alcohol dehydrogenase (XylB) was demonstrated in the strain containing pWW102-9 and of benzaldehyde dehydrogenase (XylC) in strains carrying either pWW102-7 or pWW102-9 (data not presented).

Location of genes of the upper pathway operon

Southern hybridization of digests of pWW102-7 and pWW102-9 located *xylUW*, *xylC*, *xylA* and *xylB* in the positions shown (Table 2). This location of *xylB* appears to conflict with the fact that XylB activity was expressed only from pWW102-9 and not from pWW102-7. It must be that the gene extends beyond the end of pWW102-7 (at coordinate 25·2 kb) but insufficiently far to hybridize at detectable levels beyond the *Hin*dIII site. No probes were used for *xylM* or for *xylN*.

Cloning of the *meta* pathway operon

Two plasmids carrying *meta* pathway genes were identified from a library of plasmid fragments inserted

Table 2. Summary of Southern blot hybridization of gene-specific probes against digests of plasmids containing fragments from pWW102 DNA

Probe	Gene(s) carried	Source of probe genes*	*xyl* homologue	Coordinates (kb)†
pWW110-6003	*bphX*	pWW110	*xylX*	2·7–3·9
pVI1300	*dmpQ*	pVI150	*xylT*	5·9–7·0
pWW110-4010	*bph(T)E*	pWW110	*xyl(T)E*	5·9–8·0
pVI:00	*dmpBC*	pVI150	*xylEG*	6·9–9·3
pVI1308	*dmpD*	pVI150	*xylF*	8·0–9·3
pVI1309	*dmpE*	pVI150	*xylJ*	10·2–10·9
pVI1310	*dmpF*	pVI150	*xylQ*	10·9–12·2
pVI1312	*dmpG*	pVI150	*xylK*	12·1–13·1
pVI1313	*dmpH*	pVI150	*xylI*	12·4–13·3
pVI1314	*dmpJ*	pVI150	*xylH*	13·3–14·8
PCR*xylH*‡	*xylH*	pWW102	*xylH*	
pMT100	*xylS*	pWW0	*xylS*	13·0–15·2
pWW53-3340§	*xylR*	pWW53	*xylR*	15·2–16·6
pWW0-7000	*xylUW(C)*	pWW0	*xylU(C)*	19·1–20·7
PCR*xylC*‡	*xylC*	pWW0	*xylC*	19·1–23·1
PCR*xylA*‡	*xylA*	pWW0	*xylA*	23·1–25·2
PCR*xylB*‡	*xylB*	pWW0	*xylB*	23·1–25·2

* Sources of probe DNA are the biphenyl catabolic plasmid pWW110 (Carrington *et al.*, 1994), TOL plasmids pWW0 (this study), pWW53 (Keil *et al.*, 1985) and pWW102 (this study), and phenol catabolic plasmid pVI150 (Shingler *et al.*, 1992); all pVI plasmids were prepared and supplied by V. Shingler, Umeå University, Sweden.

† The coordinates are presented as the minimum span of the gene(s) as deduced from the hybridization obtained and the restriction map presented in Fig. 2. Because of the location of available restriction sites, they represent a region potentially larger than the size of the gene itself but within which it is substantially located.

‡ Specific gene probes prepared by PCR using primers derived from the published gene sequences.

§ 1·8 kb *Bgl*II–*Hin*dIII fragment of pWW53-3301 extracted from gel and used directly to probe.

into pUC18 by using the ability of strains carrying a cloned catechol 2,3-dioxygenase to convert catechol to its yellow product, 2-hydroxymuconic semialdehyde. These contained large overlapping inserts, a 17 kb *Hin*dIII fragment (pWW102-3) and an 11·6 kb *Bam*HI fragment (pWW102-6) (Table 1). Both plasmids were completely mapped using 14 restriction enzymes (Fig. 2).

Functional analysis of *meta* pathway genes

Confirmation of the presence of *meta* pathway genes on pWW102-3 and pWW102-6 was obtained by assaying three representative enzymes, catechol 2,3-dioxygenase (XylE), 2-hydroxymuconic-semialdehyde hydrolase (XylF) and 2-hydroxymuconic-semialdehyde dehydrogenase (XylG) (data not shown).

Location of the genes of the *meta* pathway operon

The locations of *xyl* genes on the inserts were determined by hybridizations of Southern blots of restriction digests of plasmid pWW102-3 and three plasmids subcloned from pWW102-6 (pWW102-61, pWW102-62 and pWW102-63, carrying a 4·0 kb *Bam*HI–*Sac*I, a 3·1 kb *Sac*I–*Sma*I and a 3·6 kb *Sma*I fragment, respectively) (Table 1). Three sources of *meta* pathway genes, all from *Pseudomonas* catabolic plasmids, were used to prepare gene-specific probes: (1) cloned fragments of the *meta* pathway genes of the phenol catabolic plasmid pVI150 carrying parts or complete genes from *dmpB* (for catechol 2,3-dioxygenase) downstream (Shingler *et al.*, 1992); (2) for the genes for conversion of benzoate to catechol, fragments carrying *bph* genes from the biphenyl catabolic plasmid pWW110 (Carrington *et al.*, 1994); (3) a 0·7 kb PCR-generated fragment of *xylH* genes from the archetypal TOL plasmid pWW0. The hybridization results showed that the gene order of the *meta* pathway of pWW102 is *xylXYZLTEGFQKIJH* (Fig. 2), identical to the lower pathway operon on TOL plasmid pWW0 (Harayama & Rekik, 1990, 1993).

Location of regulator genes *xylS* and *xylR*

Using probes prepared from *xylS* and *xylR* from TOL plasmids pWW0 and pWW53 (Table 2), the homologues from pWW102 were located between coordinates 13·0 and 15·2 kb and between 15·2 and 16·6 kb, respectively (Fig. 2).

Linkage between two catabolic operons

Cross-hybridization of Southern blots of pWW102-6, containing a *Bam*HI insert, and pWW102-7, containing a *Hin*dIII insert, showed that they had in common a 1 kb *Hin*dIII–*Bam*HI fragment at their termini (data not shown). The two fragments are therefore adjacent and overlapping and, together with the other overlapping fragments, span a contiguous 28 kb of pWW102 DNA (Fig. 2).

Nucleotide sequence of the *xylQ* gene

From the hybridization results, *xylQ* (for ADA) was shown to be located between the *Kpn*I and *Sal*I sites at coordinates 10·0 and 12·2 kb (Fig. 2) and on the subclone pWW102-62 (Table 1). Three subclones of this region were made carrying as inserts a 2·2 kb *Kpn*I–*Sma*I fragment (pWW102-622), a 2·1 kb *Eco*RI–*Sma*I fragment (pWW102-623) and a 1 kb *Sal*I–*Sal*I fragment (pWW102-624) (Table 1). The nucleotide sequence on both strands of 1611 bp was determined from these plasmids. The ORF of *xylQ* (921 bp) encodes a protein with a predicted molecular mass of 33 kDa. Upstream of *xylQ* is the 3′-end of an ORF of 82 codons, giving 25 amino acids which showed 96·4% homology with the C-terminus of NahI from pWW60-22 (Platt *et al.*, 1995), and is therefore identified as *xylJ*, the gene for 2-oxopent-4-enoate hydratase. Downstream of *xylQ* is the 5′-end of an ORF of 140 codons giving a protein sequence of 134 residues which is 75·2% identical with the N-terminus of NahM, and is thereby identified as *xylH*, the gene for 4-hydroxy-2-oxovalerate aldolase.

Enzyme assay and electrophoretic analysis of expressed ADA proteins

The 3·1 kb *Sac*I–*Sma*I insert of pWW102-62 was excised and re-ligated into expression vector pTrc99a to form plasmid pWW102-62E. After induction with IPTG, both *E. coli* JM109(pWW102-62) and *E. coli* JM109-(pWW102-62E) expressed high ADA activities [400–500 mU (mg protein)$^{-1}$] compared with induced cells of *P. putida* O_2C_2, where the activity was about 12 mU (mg protein^{-1}). Extracts were also submitted to SDS-PAGE and showed the presence of a strongly expressed protein with a molecular mass of 35 kDa, in fair agreement with that predicted from the deduced amino acid sequence of the corresponding ADA protein (data not presented).

DISCUSSION

The large plasmid pWW102 of *P. putida* strain O_2C_2 appears to confer the phenotypic properties of other TOL plasmids. It too is counterselected by growth on benzoate minimal medium (Williams & Murray, 1974; Williams & Worsey, 1976), and the plasmid is totally lost after just one cycle of 20 generations. High frequencies of loss have been noted before but those plasmids which exhibited it appeared to segregate deletion mutants in which some of the TOL-encoded phenotypes were retained (Pickup & Williams, 1982; Keil & Williams, 1985; Osborne *et al.*, 1988; O'Donnell & Williams, 1991). pWW102 is the first example where the high frequency of loss is associated with complete loss of the plasmid.

The organization of the *xyl* operons on pWW102 is different from any described before. However, within each of the two catabolic operons and the two adjacent but divergently expressed regulatory genes, we can detect no difference from any of the previously reported plasmids. Given that (a) our analysis of gene location has a degree of imprecision due to the location of restriction sites available, (b) we did not use gene-specific probes for all 24 of the genes (probes for *xylY,Z,L,N* were absent) and (c) we did not determine the directions of transcription, there is nothing which would make us doubt that the internal operonic organization differed from that in pWW0. The organization of the operons is novel and, compared with the archetype pWW0, the two catabolic operons are reversed in order, *xylUWCMABN* being located downstream of *xylXYZLTEGFJQKIH*. pWW102 does bear considerable similarity to pWW53 in that all its *xyl* genes are in the same orientation and proximity as are the pWW53 *meta* pathway operon 1, the regulatory pair *xylSR* and the single upper pathway operon. We have acquired no evidence during this study of pWW102 to suggest that pWW102 has any of the *xyl* gene duplications found on pWW53 (Osborne *et al.*, 1988) or on pWW15 (O'Donnell & Williams, 1991).

Sequencing of *xylQ* confirmed that the gene encoding ADA is within the *meta* pathway operon of pWW102 in the place indicated by DNA–DNA hybridization. On either side of *xylQ* are partial reading frames confirming the gene order is *xylJQK*, just as found on the TOL plasmid pWW0 (Harayama & Rekik, 1990). This group of three genes is found in the same order in a number of pathways which involve extradiol cleavage of the aromatic ring. In the *xyl meta* pathway operon of TOL plasmids, the *dmp* operon of phenol catabolic plasmid pVI150 (Shingler *et al.*, 1992) and the second *nah* operon of naphthalene catabolic plasmid pWW60-22 (Platt *et al.*, 1995) it is located towards the end of a conserved gene cluster (homologous to *xylTLEGFJQKIH*) encoding the conversion of catechol to central metabolites: in each case, however, the first genes of the operon differ and encode the conversion of a chemically stable compound (benzoate, phenol and salicylate, respectively) to catechol. The same group of homologues of these three genes is also found in the same order (a) as the terminal three genes in the adjacent cumate (*cmtFHG*) (Eaton, 1996) and toluene (*todGIH*) (Lau *et al.*, 1994) operons of *P. putida* F1, (b) in several biphenyl (*bph*) 'upper' operons, where they are involved in the catabolism of the 2-hydroxypenta-2,4-dienoate formed as the product after cleavage of the first ring of the biphenyl (Kikuchi *et al.*, 1994; Hofer *et al.*, 1994; Kimura *et al.*, 1997; GenBank accession no. D85853) and (c) in the pathway for cumene (isopropylbenzene) catabolism (*cumFGH*) (Habe *et al.*, 1996). In all these latter examples the three genes are found with adjacent

(a)

```
                    11         21         31         41
TodI           ...MTRKVKA AIIGSGNIGT DLMIKILRHG QHIEMGAMVG IDPASDGLAR
BphG_KKS102    ...MTRKLKA AIIGSGNIGT DLMIKIISTA DSVEMGAMVG IDAASDGLAR
XylQ_pWW0      ...MNKKLKV AIIGPGNIGT DLMIKVMRNA QYLEMGAMVG IDPASDGLAR
XylQ_pDK1      ...MSKKLKV AIIGPGNIGT DLMIKVMRNA EHLEMGAMVG IDPASDGLAR
DmpF           ...MNQKLKV AIIGSGNIGT DLMIKVLRNA KYLEMGAMVG IDAASDGLAR
CmtH           MQTHDGKTAV AIIGSGNIGT DLMVKILRHG KHLRMGAFVG IDPESDGLKR
NahO           ...MSKKLKA AIIGPGNIGT DLVMKMLR.S EWIEPVWMVG IDPNSDGLKR
XylQ_pWW102    ...MSKKLKA AIIGPGNIGT DLVMKMLR.S EWIEPVWMVG IDPNSDGLKR
BphJ           ...MTKKIKC ALIGPGNIGT DLLAKLQR.S PVLEPIWMVG IDPESDGLKR
BphX2_KF707    ...MTKKIKC ALIGPGNIGT DLLAKLQR.S PVLEPIWMVG IDPESDGLKR
CumG           ....MNKIKC ALIGPGNIGT DLLYKLQR.S TVLEPVWMVG VDPTSEGLRR
                   * *  * ** ***** **  *          * *  *  * ** *
                    51         61         71         81         91
TodI           GERMGVAITH EGVEGLTRLP VFNEIDVVFD ATSAGAHVKN EALLRERKPG
BphG_KKS102    AQALGVATTH EGAEGLTRLP VFADIDIVFD ATSAGAHVKN DALLRHHKPA
XylQ_pWW0      AQRMGVATTH EGVEGFINLP EFADIDFVFD ATSASAHVQN DALLRRAKPG
XylQ_pDK1      AQRMGVATTH EGVEGLINLP EFADIDFVFD ATSASAHVRN DALLRRAKPG
DmpF           AQRMGVTTTY AGVEGLIKLP EFADIDFVFD ATSASAHVQN EALLRQAKPG
CmtH           AERMGVPTVS TGIEGLLSHP DFGSIGFVFD ATSAGAHARH EQLLRPH..G
NahO           ARDFGMKTTA EGVDGLLPHV LDDDIRIAFD ATSAYVHAEN SRKLNAL..G
XylQ_pWW102    AREFGMKTTA EGVDGLLPHV LDDDIRIAFD ATSAYVHAEN SRKLNAL..G
BphJ           AREMGIKTTA DGVDGLIPHM QADGVQIVFD ATSAYVHADN SRKVNAL..G
BphX2_KF707    AREMGIKTTA DGVDGLIPHM QADGVQIVFD ATSAYVHADN SRKVNAL..G
CumG           AAAMGLKTTS DGVDGLLPHI ATDGILIAFD ATSAYVHAEN SRKLNEL..G
                   *     *  *                 * **** *
                   101        111        121        131        141
TodI           LRMIDLTPAA IAPYCIPVVN GDDHLDAT.. NVNMVTCGGQ ATIPMVAAVS
BphG_KKS102    MRVIDLTPAA IGPYCIPVVN GEDHLAAL.. NVNMVTCGGQ ATIPMVAAVS
XylQ_pWW0      IRLIDLTPAA IGPYCVPVVN LEEHLAKL.. NVNMVTCGGQ ATIPMVAAVS
XylQ_pDK1      IRLIDLTPAA IGPYCVPVVN LEEHLAKL.. .VNMVTCGGQ ATIPMVAAVS
DmpF           IRLIDLTPAA IGPYCVPVVN LEEHLGKL.. NVNMVTCGGQ ATIPMVAAVS
CmtH           VRVIDLTPAA IGPFVVPAVN IEQHLDAP.. NVNMVTCGGQ ATIPMVAAVS
NahO           VLMVDLTPAA IGPYCVPPVN LKQHVGRLEM NVNMVTCGGQ ATIPMVAAVS
XylQ_pWW102    VLMVDLTPAA IGPYCVPPVN LKQHVGRLEM NVNMVTCGGQ ATIPMVAAVS
BphJ           ALMIDLTPAA IGPFCVPTVN LKEHVGKGEM NVNMVTCGGQ ATIPMVAAVS
BphX2_KF707    ALMIDLTPAA IGPFCVPTVN LKEHVGKGEM NVNMVTCGGQ ATIPMVAAVS
CumG           VMMIDLTPAA IGPFCVPPVN LLKHVGQGEM NVNMVTCGGQ ATIPMVAAVS
                   ****** * *   * **    *         ********* **********
                   151        161        171        181        191
TodI           RVAKVHYAEI IASISSKSAG PGTRANIDEF TETTSKAIEV VGGAAKGKAI
BphG_KKS102    RVAKVQYGEI VASISSKSAG PGTRANIDEF TETTSKAIEV VGGAAKGKAI
XylQ_pWW0      RVAKVHYAEI VASIASKSAG PGTRANIDEF TETTSKAIEV IGGAAKGKAI
XylQ_pDK1      RVAKVHYAEI VASIASKSAG PGTRANIDEF TETTSKAIEV IGGAAKGKAI
DmpF           RVAKVHYAEI VASISSKSAG PGTRANIDEF TETTSKAIEV IGGAAKGKAI
CmtH           RVVPVEYAEI VASISSRSAG PGTRANIDEF TETTSNAIVR VGGAQRGKAI
NahO           RVQPVAYAEI VATVSSRSVG PGTRKNIDEF TRTTAGAIEQ VGGAREGKAI
XylQ_pWW102    RVQPVAYAEI VATVSSRSVG PGTRKNIDEF TRTTAGAIEQ VGGAREGKAI
BphJ           RVQPVAYGEI VATVSSKSAG PGTRKNIDEF TRTTAGAVEK VGGAKKGKAI
BphX2_KF707    RVQPVAYGEI VATVSSKSAG PGTRKNIDEF TRTTAGAVEK VGGAKKGKAI
CumG           RVQPVTYGEI VATTASKSIG PGTRANIDEF TRTTGSAIEI VGGAKKGKAI
                   **  * * ** **   * * * **** ***** * **  *     ** * ****
                   201        211        221        231        241
TodI           IVLNPAEPPL MMRDAVYTLS .DFADIDQIE ESVQRMADAV QAYVPGYRLK
BphG_KKS102    IVLNPAEPPL IMRDTVYTLS .DFADTAKIE DSVQRMAAAV QAYVPGYRLK
XylQ_pWW0      IVMNPAEPPL IMRDTVFVLS .ETVDQAQVE ASVEEMTSAV QAYVPGYRLK
XylQ_pDK1      IIMNPAEPPL IMRDTVFVLS .EAVDQAQVE ASIEEMASAV QAYVPGYRLK
DmpF           IIMNPAEPPL IMRDTVYVLS .AAADQAAVA ASVAEMVQAV QAYVPGYRLK
CmtH           IILNPAEPPL IMRDTVYCLV PGHADRQAIV ESVERMASAV AAYVPGYRLK
NahO           IVINPAEPPL MMRDTIHCLT DSEPDQAAIT ASVHAMIAEV QKYVPGYRLK
XylQ_pWW102    IVINPAEPPL MMRDTIHCLT DSEPDQAAIT VSVHAMIAEV QKYVPGYRLK
BphJ           IILNPAEPPL IMRDTVHCLL ESEPDQAKIT ESIHAMIKEV QKYVPGYKLV
BphX2_KF707    IILNPAEPPL IMRDTVHCLL ESEPDQAKIT ESIHAMIKEV QKYVPGYKLV
CumG           IIINPAEPPM IMRDTVHCLT ETEPDRERIT ASVQKMVKEV QKYVPGYKLV
               *  ******   ***    *              *   *   *   ***** *
                   251        261        271        281        291
TodI           QRVQFDRIEA DCPIRIPGVG DRMNGLKTSI FLEVEGAAHY LPAYAGNLDI
BphG_KKS102    QRVQFDRIEA DRPIRIPGVG DRLTGLKTSV FLEVEGAAHY LPAYAGNLDI
XylQ_pWW0      QKVQFDVIPE SAPLHIPGLG .TFSGLKTSI YLEVEGAAHY LPAYASSLDI
XylQ_pDK1      QKVQFDMIPA SAPLHIPGLG .TFSGLKTSI YLEVEGAAHY LPAYAGNLDI
DmpF           QQVQFDVIPE SAPLNIPGLG .RFSGLKTSV FLEVEGAAHY LPAYAGNLDI
CmtH           QTVQFDEFKG RMPQETGSAQ ..APRLKVSV FLEVEGAGHY LPSYAGNLDI
NahO           NGPVFD.... .......... ....GNRVSI FMEVEGLGDY LPKYAGNLDI
XylQ_pWW102    NGPVFD.... .......... ....GNRVSI FMEVEGLGDY LPKYAGNLDI
BphJ           NGPVFD.... .......... ....GLRVSV YLEVEGLGDY LPKYAGNLDI
BphX2_KF707    NGPVFD.... .......... ....GLRVSV YLEVEGLGDY LPKYAGNLDI
CumG           NGPVFD.... .......... ....GKRISL FLEVEGLGDY LPKYAGNLDI
                                                 ****   * ** **  ***
                   301        311        321        331
TodI           MTSAALRTAE KLAERLLASL VA
BphG_KKS102    MTSAALRTAE KMAQRLLATL TA
XylQ_pWW0      MTSAALATAE RMAQSLLNA
XylQ_pDK1      MTSAALATAE RMAQSLLNA
DmpF           MTSAALATAE RMAQSMLNA
CmtH           MTSAALATAE RIAARQPVAA
NahO           MTAAALRTGE MFAEEIAAGT IQLPRRDIAL A
XylQ_pWW102    MTAAALRTGE MFAEEIAAGT IQLPRRDIAF A
BphJ           MTAAAARTAE MFAEEILAGQ LTLQPVHA
BphX2_KF707    MTAAAARTAE MFAEEILAGQ LTLQPVHA
CumG           MTAAAVRTAE MFAEEISSGK LKLGPTIETQ K

               ** **  * *
```

(b)

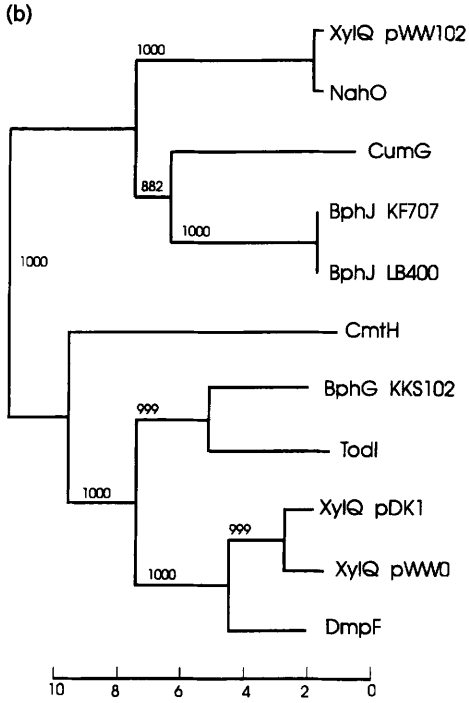

Fig. 3. Comparisons of amino acid sequences of ADA proteins from *Pseudomonas* catabolic pathways. The sequences were obtained from GenBank with accession numbers: XylQ (plasmid pWW102), this work; NahO, U13232; CumG, D63377; BphJ (strain KF707), D85853; BphJ (strain LB400), X76500; CmtH, U24215; BphG (strain KKS102), D16407; TodI, U09250; XylQ (plasmid pDK1), R. J. Benjamin, personal communication; XylQ (plasmid pWW0), M94186; DmpF, X60835. (a) Alignments produced by CLUSTAL W 1.7 (Thompson *et al.*, 1994), using the BLOSUM series of amino acid weight matrices with gap penalty 10 and gap extension penalty 0·1. Residues conserved in all sequences are marked with an asterisk. (b) Dendrogram drawn from the CLUSTAL alignment using the built in neighbour-joining function (Saitou & Nei, 1987). The numerical values are the bootstrap values with 1000 trials. The scale represents percentage divergence.

genes either unrelated or only distantly related to those which surround them in the *xyl*, *nah* and *dmp* operons.

A model for the evolution of the *Pseudomonas meta* pathway operons which has been proposed involves the recombination of pre-evolved genetic modules to give the gene arrangements seen in modern strains (Cane & Williams, 1986; Harayama *et al.*, 1987; Assinder & Williams, 1988; van der Meer *et al.*, 1992; Williams & Sayers, 1994). In the case of the *xyl meta* pathway operon, we can see evidence for three such 'modules'. The first and largest is the homologous sequence

equivalent to *xylT–xylH* found in the *meta* pathway operons for catabolism of toluene/xylene (Harayama & Rekik, 1993; Keil *et al.*, 1985), phenols (Shingler *et al.*, 1992), salicylate (and therefore naphthalene) (Assinder & Williams, 1988) and biphenyl (Carrington *et al.*, 1994). The second is located within these operons as the 'submodule' equivalent of *xylJQK* found also in other catabolic operons (see above). The third is upstream within the same operon, the *xylXYZL* genes for conversion of benzoate to catechol, homologues of which turn up as *benABCD* in the chromosomal β-ketoadipate pathway in both *Acinetobacter calcoaceticus* and *P. putida* (Neidle *et al.*, 1991, 1992).

If all *xyl meta* pathway operons, for example, share a common ancestor which arose by a modular assembly, then it would be expected that the genes and proteins found in all modern *xyl* pathways would be more closely related to each other than to those in any other pathway, such as a *nah* pathway, which would have arisen by an independent assembly and subsequently divergently radiated. However, alignment of the amino acid sequences of *Pseudomonas* ADA proteins (Fig. 3) shows this expectation not to be the case. Thus XylQ of pWW102 is more closely related to NahO of pWW60-22 than to the other XylQ proteins of TOL plasmids pWW0 and pDK1. Other comparisons of genes within the *Pseudomonas meta* pathway operons have shown a similar result.

In a detailed comparison of the nucleotide sequences of the *meta* pathways of pWW0, pVI150 and NAH7, Harayama & Rekik (1993) showed that the 5′-end of the 'catechol module' on pWW0 was more homologous to the isofunctional *nah* genes of NAH7, whereas its 3′-end was closer to the *dmp* genes of pVI150. They proposed that the *xyl* operon was the result of recombination between ancestral *nah* and *dmp* operons. Analysis of the nucleotide sequence of the *xylE* gene on TOL plasmid pDK1 showed that it had an apparent discontinuity in its homology and that its 5′-end was more homologous to *nahH* from NAH7 whereas the 3′-end was closer to *xylE* of TOL plasmid pWW0 (Benjamin *et al.*, 1991). Similarly the *xylS* regulatory gene on TOL plasmid pDK1 is also clearly a chimera of two different *xylS* homologues, one of each of which is found on another TOL plasmid, pWW53 (Assinder *et al.*, 1993). These results suggest that recombination between two sequences, not necessarily encoding the catabolism of the same original substrate, may give rise to hybrid or mosaic patterns within individual genes or within operons. Such an event has been recorded between the duplicate *meta* pathway operons of TOL plasmid pWW53 (Osborne *et al.*, 1988).

The ADA proteins fall into two clearly differentiated clusters (Fig. 3). This is very apparent in some areas where there are strong identities within the two subclusters but not between them. These are (a) residues 257–274, where all the proteins in the cluster containing XylQ from pWW102 contain the same deletion of 18 residues, and (b) three locations where the two clusters differ by identical small deletions: at positions 29, 98–99 and 129–130 in the alignment of Fig. 3(a). Exactly the same pattern of clustering is found when the protein sequences of the upstream genes (XylJ homologues) and of the downstream genes (XylK homologues) are compared (data not shown), which shows that in all these strains the *xylJQK* homologues have the same pattern of ancestry. However, there are no clear lineages within any of this cluster of three genes which directly correlate with their metabolic function (Fig. 3b). Thus the pWW102 ADA is closest to the homologous sequence on the naphthalene plasmid pWW60-22 but is in the opposite subcluster to the two TOL plasmid proteins from pWW0 and pDK1. In a parallel manner, two of the *bph* ADA proteins are close to the pWW0 XylQ, whereas another *bph* protein and TodI, the genes for which are very similar in arrangement to *bph* genes, are in the other subcluster and closely related to the pWW102 protein. There are two possible explanations: either the gene organizations we see conserved in these pathways have evolved separately on more than one occasion or, more likely in our opinion, recombination between homologous genes in these operons has resulted in mixing of the sequences between different metabolic pathways in a manner equivalent to gene conversion, thus obscuring their origin.

ACKNOWLEDGEMENTS

We should like to thank Wouter Duetz for strain O_2C_2, Vicky Shingler for sending all the plasmids derived from pVI150 and Linda Shaw for her continued help and advice during the course of this work and in the writing of this paper.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

Assinder, S. J. & Williams, P. A. (1988). Comparison of the *meta* pathway operons on NAH plasmid pWW60-22 and TOL plasmid pWW53-4 and its evolutionary significance. *J Gen Microbiol* **134**, 2769–2778.

Assinder, S. J. & Williams, P. A. (1990). The TOL plasmids: determinants of the catabolism of toluene and the xylenes. *Adv Microb Physiol* **31**, 1–69.

Assinder, S. J., De Marco, P., Osborne, D. J., Poh, C. L., Shaw, L. E., Winson, M. K. & Williams, P. A. (1993). A comparison of the multiple alleles of *xylS* carried by TOL plasmids pWW53 and pDK1 and its implications for their evolutionary relationship. *J Gen Microbiol* **139**, 557–568.

Bartilson, M. & Shingler, V. (1989). Nucleotide sequence and expression of the catechol 2,3-dioxygenase-encoding gene of phenol-catabolizing *Pseudomonas* CF600. *Gene* **85**, 233–238.

Benjamin, R. C., Voss, J. A. & Kunz, D. A. (1991). Nucleotide sequence of *xylE* from the TOL pDK1 plasmid and structural comparison with isofunctional catechol 2,3-dioxygenase genes from TOL pWW0 and NAH7. *J Bacteriol* **173**, 2724–2728.

Cane, P. A. & Williams, P. A. (1982). The plasmid-encoded metabolism of naphthalene and 2-methylnaphthalene in *Pseudomonas*: phenotypic changes correlated with structural modification of the plasmid pWW60-1. *J Gen Microbiol* **128**, 2281–2290.

Cane, P. A. & Williams, P. A. (1986). A restriction map of naphthalene catabolic plasmid pWW60-1 and the location of some of its catabolic genes. *J Gen Microbiol* **132**, 2919–2929.

Carrington, B., Lowe, A., Shaw, L. E. & Williams, P. A. (1994). The lower pathway operon for benzoate catabolism in biphenyl-utilizing *Pseudomonas* sp. strain IC and the nucleotide sequence of the *bphE* gene for catechol 2,3-dioxygenase. *Microbiology* **140**, 499–508.

Chatfield, L. K. & Williams, P. A. (1986). Naturally-occuring TOL plasmids isolated from soil carry either two homologous or two nonhomologous catechol 2,3-oxygenase genes. *J Bacteriol* **168**, 878–885.

Duggleby, C. J. & Williams, P. A. (1986). Purification and some properties of the 2-hydroxy-6-oxohepta-2,4-dienoate hydrolase (2-hydroxymuconic semialdehyde hydrolase) encoded by the TOL plasmid pWW0 from *Pseudomonas putida* mt-2. *J Gen Microbiol* **132**, 717–726.

Eaton, R. W. (1996). *p*-Cumate catabolic pathway in *Pseudomonas putida* F1. Cloning and characterization of DNA carrying the CMT operon. *J Bacteriol* **178**, 1351–1362.

Girvitz, S. C., Bacchetti, S., Rainbow, A. J. & Graham, F. W. (1980). A rapid and efficient procedure for the purification of DNA from agarose gels. *Anal Biochem* **106**, 492–496.

Guerry, P., Leblanc, D. J. & Falkow, S. (1973). General method for isolation of plasmid deoxyribonucleic acid. *J Bacteriol* **116**, 1064–1066.

Habe, H., Kimura, T., Nojiri, H., Yamana, H. & Omori, T. (1996). Cloning and nucleotide sequences of the genes involved in the *meta* cleavage pathway of cumene degradation in *Pseudomonas fluorescens* IP01. *J Ferment Bioeng* **81**, 187–196.

Harayama, S. & Rekik, M. (1990). The *meta*-cleavage operon of TOL degradative plasmid pWW0 comprises 13 genes. *Mol Gen Genet* **222**, 113–120.

Harayama, S. & Rekik, M. (1993). Comparison of the nucleotide sequences of the *meta*-cleavage pathway of TOL plasmid pWW0 from *Pseudomonas putida* with other *meta*-cleavage genes suggests that both single and multiple nucleotide substitutions contribute to enzyme evolution. *Mol Gen Genet* **239**, 81–89.

Harayama, S., Rekik, M., Wasserfallen, A. & Bairoch, A. (1987). Evolutionary relationship between catabolic pathways for aromatics: conservation of gene order and nucleotide sequences for catechol oxidation genes of pWW0 and NAH7 plasmids. *Mol Gen Genet* **210**, 241–247.

Harayama, S., Rekik, M., Wubbolts, M., Rose, K., Leppik, R. A. & Timmis, K. N. (1989). Characterization of 5 genes in the upper pathway operon of TOL plasmid pWW0 from *Pseudomonas putida* and identification of the gene products. *J Bacteriol* **171**, 5048–5055.

Hofer, B., Backhaus, S. & Timmis, K. N. (1994). The biphenyl/polychlorinated biphenyl-degradation locus (*bph*) of *Pseudomonas* sp. LB400 encodes four additional metabolic enzymes. *Gene* **144**, 9–16.

Holmes, D. S. & Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Anal Biochem* **114**, 193–197.

Keil, H. & Williams, P. A. (1985). A new class of TOL plasmid deletion mutants in *Pseudomonas putida* MT15 and their reversion by tandem gene amplification. *J Gen Microbiol* **131**, 1023–1033.

Keil, H., Keil, S., Pickup, R. W. & Williams, P. A. (1985). Evolutionary conservation of genes coding for *meta* pathway enzymes within TOL plasmids pWW0 and pWW53. *J Bacteriol* **164**, 887–895.

Keil, H., Saint, C. M. & Williams, P. A. (1987). Gene organization of the first catabolic operon of TOL plasmid pWW53: production of indigo by the *xylA* gene product. *J Bacteriol* **169**, 764–770.

Kikuchi, Y., Yasukochi, Y., Nagata, Y., Fukuda, M. & Takagi, M. (1994). Nucleotide sequence and functional analysis of the *meta*-cleavage pathway involved in biphenyl and polychlorinated biphenyl degradation in *Pseudomonas* sp. strain KKS102. *J Bacteriol* **176**, 4269–4276.

Kimura, N., Nishi, A., Goto, M. & Furukawa, K. (1997). Functional analyses of a variety of chimeric dioxygenases constructed from two biphenyl dioxygenases that are structurally similar but different functionally. *J Bacteriol* **179**, 3936–3943.

King, E. D., Ward, M. K. & Raney, D. C. (1954). Simple method for the demonstration of pyocyanin and fluorescein. *J Lab Clin Methods* **44**, 301–307.

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

van der Meer, J. R., De Vos, W. M., Harayama, S. & Zehnder, A. B. (1992). Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiol Rev* **56**, 677–694.

Mermod, N., Harayama, S. & Timmis, K. N. (1986). New route for the bacterial production of indigo. *Biotechnology* **4**, 321–324.

Neidle, E. L., Hartnett, C., Ornston, L. N., Bairoch, A., Rekik, M. & Harayama, S. (1991). Nucleotide sequences of the *Acinetobacter calcoaceticus benABC* genes for benzoate dioxygenase reveal evolutionary relationships among multicomponent oxygenases. *J Bacteriol* **173**, 5385–5395.

Neidle, E. L., Hartnett, C., Ornston, L. N., Bairoch, A., Rekik, M. & Harayama, S. (1992). *Cis*-diol dehydrogenases encoded by the TOL (pWW0) plasmid and the *Acinetobacter calcoaceticus* chromosomal *benD* gene are members of the short-chain alcohol dehydrogenase superfamily. *Eur J Biochem* **204**, 113–120.

O'Donnell, K. & Williams, P. A. (1991). Duplication of both *xyl* operons on TOL plasmid pWW15. *J Gen Microbiol* **137**, 2831–2838.

Osborne, D. J., Pickup, R. W. & Williams, P. A. (1988). The presence of two complete homologous *meta* pathway operons on TOL plasmid pWW53. *J Bacteriol* **134**, 2965–2975.

Pickup, R. W. & Williams, P. A. (1982). Spontaneous deletions in the TOL plasmid pWW20 which give rise to the B3 mutants of *Pseudomonas putida* MT20. *J Gen Microbiol* **128**, 1385–1390.

Platt, A., Shingler, V., Taylor, S. C. & Williams, P. A. (1995). The 4-hydroxy-2-oxovalerate aldolase and acetaldehyde dehydrogenase (acylating) encoded by the *nahM* and *nahO* genes of the naphthalene catabolic plasmid pWW60-22 provide further evidence of conservation of *meta*-cleavage pathway gene sequences. *Microbiology* **141**, 2223–2233.

Polissi, A. & Harayama, S. (1993). *In vivo* reactivation of catechol 2,3-dioxygenase mediated by a chloroplast-type ferredoxin – a bacterial strategy to expand the substrate specificity of aromatic degradative pathways. *EMBO J* **12**, 3339–3347.

Saitou, N. & Nei, M. (1987). The neighbour-joining method – a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 426–444.

Sala-Trepat, J. M. & Evans, W. C. (1971). The *meta* cleavage of catechols by *Azotobacter* species. *Eur J Biochem* **20**, 400–413.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Shingler, V., Powlowski, J. & Marklund, U. (1992). Nucleotide sequence and functional analysis of the complete phenol/3,4-dimethylphenol catabolic pathway of *Pseudomonas* sp. strain CF600. *J Bacteriol* **174**, 711–724.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Microbiol* **98**, 503–517.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W – improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

Vieira, J. & Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259–268.

Wheatcroft, R. & Williams, P. A. (1981). Rapid methods for the study of both stable and unstable plasmids in *Pseudomonas*. *J Gen Microbiol* **124**, 433–437.

Williams, P. A. & Murray, K. (1974). Metabolism of benzoate and the methylbenzoates by *Pseudomonas putida* (*arvilla*) mt-2: evidence for the existence of a TOL plasmid. *J Bacteriol* **120**, 416–423.

Williams, P. A. & Sayers, J. R. (1994). The evolution of pathways for aromatic hydrocarbon oxidation in *Pseudomonas*. *Biodegradation* **5**, 195–217.

Williams, P. A. & Worsey, M. J. (1976). Ubiquity of plasmids in coding for toluene and xylene metabolism in soil bacteria: evidence for the existence of new TOL plasmids. *J Bacteriol* **125**, 818–828.

Williams, P. A., Shaw, L. E., Pitt, C. W. & Vrecl, M. (1997). *xylUW*, two genes at the start of the upper pathway operon of TOL plasmid pWW0, appear to play no essential part in determining its catabolic phenotype. *Microbiology* **143**, 101–107.

Worsey, M. J. & Williams, P. A. (1975). Metabolism of toluene and xylenes by *Pseudomonas putida* (*arvilla*) mt-2: evidence for a new function of the toluene plasmid. *J Bacteriol* **124**, 7–13.

Worsey, M. J., Franklin, F. C. H. & Williams, P. A. (1978). Regulation of the degradative pathway enzymes coded for by the TOL plasmid (pWW0) from *Pseudomonas putida* mt-2. *J Bacteriol* **134**, 757–764.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.

Received 14 November 1997; revised 26 January 1998; accepted 10 February 1998.