

Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440

José Ignacio Jiménez, Baltasar Miñambres, José Luis García and Eduardo Díaz*

Departamento de Microbiología Molecular, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Velázquez 144, 28006 Madrid, Spain.

Summary

Analysis of the catabolic potential of *Pseudomonas putida* KT2440 against a wide range of natural aromatic compounds and sequence comparisons with the entire genome of this microorganism predicted the existence of at least four main pathways for the catabolism of central aromatic intermediates, that is, the protocatechuate (*pca* genes) and catechol (*cat* genes) branches of the β -ketoadipate pathway, the homogentisate pathway (*hmg/fah/mai* genes) and the phenylacetate pathway (*pha* genes). Two additional gene clusters that might be involved in the catabolism of N-heterocyclic aromatic compounds (*nic* cluster) and in a central *meta*-cleavage pathway (*pcm* genes) were also identified. Furthermore, the genes encoding the peripheral pathways for the catabolism of *p*-hydroxybenzoate (*pob*), benzoate (*ben*), quinate (*qui*), phenylpropenoid compounds (*fcs*, *ech*, *vdh*, *cal*, *van*, *acd* and *acs*), phenylalanine and tyrosine (*phh*, *hpd*) and *n*-phenylalkanoic acids (*fad*) were mapped in the chromosome of *P. putida* KT2440. Although a repetitive extragenic palindromic (REP) element is usually associated with the gene clusters, a supraoperonic clustering of catabolic genes that channel different aromatic compounds into a common central pathway (catabolic island) was not observed in *P. putida* KT2440. The global view on the mineralization of aromatic compounds by *P. putida* KT2440 will facilitate the rational manipulation of this strain for improving biodegradation/biotransformation processes, and reveals this bacterium as a useful model system for studying biochemical, genetic, evolutionary and ecological aspects of the catabolism of aromatic compounds.

Introduction

Pseudomonas putida, a non-pathogenic member of rRNA group I of the genus *Pseudomonas*, is able to colonize many different environments, including soil, fresh water and plant rhizosphere, and is characterized by a wide metabolic and physiologic versatility. The strain *P. putida* mt-2 (ATCC 33015) was isolated from soil by K. Hosokawa in Japan in the early 1960s by its ability to use *m*-toluate (3-methylbenzoate) as the sole carbon source, a feature later shown to result from the presence of the TOL plasmid pWW0 (Nozaki *et al.*, 1963; Assinder and Williams, 1990). *P. putida* KT2440 is a cured (Bayley *et al.*, 1977), spontaneous restriction-deficient derivative of *P. putida* mt-2 (Bagdasarian *et al.*, 1981; Franklin *et al.*, 1981) that has been used extensively as host for gene cloning and expression in *Pseudomonas* and represents the first host-vector biosafety system for cloning in Gram-negative soil bacteria. As this strain also colonizes the plant rhizosphere, it becomes useful for promoting plant growth and as a biocontrol agent for plant pathogens (O'Sullivan and O'Gara, 1992). However, *P. putida* KT2440 is mainly known for its ability to degrade aromatic compounds and as an ideal host for expanding the range of substrates that it can degrade and/or biotransform in added-value products through the recruitment of genes from other microorganisms (Rojo *et al.*, 1987; Harayama and Timmis, 1989; Ramos *et al.*, 1994).

The TOL plasmid pWW0 from *P. putida* mt-2 is a 116.5 kb catabolic plasmid that has been sequenced recently (accession no. AJ344068), and it encodes all the proteins necessary for bacterial utilization of toluene, *m*- and *p*-xylene, 3-ethyltoluene and 1,2,4-trimethylbenzene (pseudocumene), plus their alcohol, aldehyde and carboxylic acid derivatives, via a *meta*-cleavage pathway. The TOL pathway (*xyl* cluster) of *P. putida* mt-2 is among the best studied examples of aromatic hydrocarbon degradation. TOL⁺ strains, such as strain KT2440, as for any other *P. putida* strains, are still able to use benzoate as sole carbon and energy source by the chromosomally encoded β -ketoadipate pathway (Assinder and Williams, 1990). Some other natural aromatic compounds used by most *P. putida* strains are *p*-hydroxybenzoate, phenylacetate, tyrosine, phenylalanine, benzylamine and nicotinate. Some hydroaromatic compounds such as quinate are also substrates for *P. putida* strains (Stanier *et al.*, 1966).

Received 13 September, 2002; accepted 21 October, 2002. *For correspondence. E-mail ediaz@cib.csic.es; Tel. (+34) 915 611 800; Fax (+34) 915 627 518.

Despite the extensive knowledge of the TOL pathway from *P. putida* mt-2 (Kasai *et al.*, 2001), very few reports have been published on the pathways for the catabolism of other aromatic compounds in this strain. In this work, we have checked the abilities of *P. putida* KT2440 to use different aromatic compounds as sole carbon and energy source. As the entire genome (6181 kb) of *P. putida* KT2440 has been sequenced recently by the efforts of a German consortium and The Institute for Genomic Research (TIGR; <http://www.tigr.org>), we have accomplished a genomic analysis of the predicted aromatic catabolic clusters of such a strain. *P. putida* KT2440 turns out to be a very useful model system for studying biochemical, genetic, evolutionary and ecological aspects of the catabolism of aromatic compounds.

Results and discussion

Aromatic substrate range and chromosomal location of the gene clusters encoding the aromatic central pathways of P. putida KT2440

As *P. putida* KT2440 derives from a strain isolated from soil, it should be able to use several aromatic compounds, most of them proceeding from the recycling of plant-derived material, that are commonly present in the environment. Thus, in this work, we have observed that *P. putida* KT2440 is able to grow in minimal medium containing benzoate, *p*-hydroxybenzoate, benzylamine, phenylacetate, phenylalanine, tyrosine, phenylethylamine, phenylhexanoate, phenylheptanoate, phenyloctanoate, coniferyl alcohol, *p*-coumarate, ferulate, caffeate, vanillate, nicotinate and quinate (hydroaromatic compound) as sole carbon and energy source. Some other aromatic compounds, such as 2-hydroxybenzoate (salicylate), 3-hydroxybenzoate, 2,3-dihydroxybenzoate, 2-aminobenzoate (anthranilate), *p*-hydroxyphenylacetate, tyramine, aniline, atropine, 2-phenylethanol, phenol, mandelate, phenylglyoxylate, *p*-methoxybenzoate (*p*-anisate), 3,4-dimethoxybenzoate (veratrate), *p*-hydroxy-3,5-dimethoxybenzoate (syringate), cinnamate, phenylpropionate, 3-hydroxyphenylpropionate, vanillylmandelate, phthalate, pyridoxal, pyridine, isonicotinate, quinoline, isoquinoline, gallate and resorcinol, do not appear to be used by *P. putida* KT2440.

Taking into account the aromatic compounds that can be mineralized by *P. putida* KT2440 (see above), it is reasonable to predict that this strain should contain at least four different central pathways for the catabolism of these compounds, i.e. the catechol (*cat*), protocatechuate (*pca*), phenylacetate (*pha*) and homogentisate (*hmg*) pathways. As the genes responsible for such catabolic pathways have been reported in several bacteria, we have performed a sequence comparison analysis to identify the orthologue genes in *P. putida* KT2440. When the amino

acid sequences of the *cat* and *pca* gene products from *Acinetobacter* sp. ADP1 and *P. putida* PRS2000 (Harwood and Parales, 1996), *pha* gene products from *P. putida* U (Luengo *et al.*, 2001) and the *hmgA* gene product from *Sinorhizobium meliloti* (Milcamps and de Bruijn, 1999) were compared with the translated genome of strain KT2440, we were able to identify the predicted *cat*, *pca*, *pha* and *hmg* gene clusters of *P. putida* KT2440 (Fig. 1A). At positions 4441–4454 kb of the KT2440 genome, there is a 13 kb gene cluster (*nic*) that contains genes showing similarity to those encoding proteins involved in the metabolism of N-heterocyclic aromatic compounds (Fetzner, 1998). In addition, a gene cluster (*pcm*) containing genes similar to those encoding the protocatechuate 4,5-dioxygenase (*pcmA*) and oxalocitramalate aldolase (*pcmE*) from *Arthrobacter keyseri* (Eaton, 2001) was located at positions 2861–2867 kb of the KT2440 genome (Fig. 1A). Whether such gene clusters of *P. putida* KT2440 are involved in the catabolism of N-heterocyclic aromatic compounds (cluster *nic*) and in a central pathway for the degradation of aromatic compounds via a 4,5-*meta*-cleavage of the aromatic ring (cluster *pcm*) remains to be demonstrated.

The β -ketoadipate central pathway

The *pca* and *cat* gene products of *P. putida* KT2440 were significantly similar to proteins of known function from other bacteria, mainly *Acinetobacter* and *Pseudomonas* strains (Harwood and Parales, 1996) (Tables 1 and 2). The two branches of the β -ketoadipate pathway (*ortho*-cleavage pathway), i.e. the protocatechuate branch (*pca* genes) and the catechol branch (*cat* genes), will converge at β -ketoadipate enol-lactone in *P. putida*, and one set of enzymes (*pcaDIJF* gene products) will complete the conversion of the latter to the Krebs cycle intermediates, succinyl-CoA and acetyl-CoA (Harwood and Parales, 1996) (Fig. 1B).

Although the *cat* genes of the catechol branch are clustered at positions 4236–4239 kb of the *P. putida* KT2440 genome, the *pca* genes are organized in three different clusters at positions 1566–1575 kb (*pcaRKFTBDCP*), 4457–4459 kb (*pcaIJ*) and 5281–5282 kb (*pcaGH*) (Fig. 1A). The gene order within the clusters in *P. putida* KT2440 is similar to that found in *P. putida* PRS2000 (Harwood and Parales, 1996). Considering the ubiquity of the β -ketoadipate pathway, it is not surprising that this pathway is present in other species of the *Pseudomonas* genus. The *pca* orthologues from other *Pseudomonas* species of finished (*Pseudomonas aeruginosa* PAO1) (Stover *et al.*, 2000) or unfinished genomic sequence (*Pseudomonas fluorescens* Pf0-1 and *Pseudomonas syringae* pv. *tomato* DC3000) showed different chromosomal organizations from that found in *P. putida* (Fig. 2).

Phenylethylamine, benzylamine, tyramine, phenylalanine, tyrosine, nicotinate, phenylglyoxylate, resorcinol, pyridoxal HCl, quinate and phthalate were from Sigma-Aldrich and were prepared as 1 M stock solutions in water (with the exception of quinate and phthalate, which were prepared at 0.5 M and 0.8 M, respectively, and phenylalanine and tyrosine, which were prepared at 0.05 M). Benzoate, 2-hydroxybenzoate, *p*-hydroxybenzoate, 2,3-dihydroxybenzoate, tropate, atropine, mandelate, anthranilate, phenylacetate, phenylpropionate and *p*-hydroxyphenylacetate were from Sigma-Aldrich; 3-hydroxyphenylpropionate, phenylhexanoate, phenylheptanoate and phenyloctanoate were from Lancaster; 3-hydroxybenzoate was from Fluka; all these compounds were prepared as 1 M stock solutions in 2-propanol (with the exception of anthranilate and 2,3-dihydroxybenzoate, which were prepared at 0.5 M). Caffeate, ferulate, *p*-coumarate, cinnamate, vanillate, vanillylmandelate, *p*-methoxybenzoate, 3,4-dimethoxybenzoate, *p*-hydroxy-3,5-dimethoxybenzoate and gallate were from Sigma-Aldrich and were prepared as 1 M stock solutions in *N*-*N*-dimethylformamide. Neither 2-propanol nor *N*-*N*-dimethylformamide was used as a carbon source by *P. putida* KT2440. The liquid compounds phenol, 2-phenylethanol, quinoline, isoquinoline and aniline were from Sigma-Aldrich; pyridine was from Merck. Coniferyl alcohol and isonicotinate were from Sigma-Aldrich and were added directly to the growth medium at the desired concentration.

Sequence data analyses

The nucleotide sequence of the whole *P. putida* KT2440 genome was obtained from TIGR (accession no. AE015451). The complete sequence of *P. aeruginosa* PAO1 (Stover *et al.*, 2000) was obtained and analysed at the *Pseudomonas* Genome Project (<http://www.pseudomonas.com>). Nucleotide sequence analyses were done at the INFOBIOGEN server (<http://www.infobiogen.fr/services/menuserv.html#ANALN>). Open reading frame (ORF) searches were also performed with the ORF FINDER program at the National Center for Biotechnology Information (NCBI) server (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). The amino acid sequences of ORFs were compared with those present in finished and unfinished microbial genome databases using the TBLASTN algorithm (Altschul *et al.*, 1990) at the NCBI server (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/genom_table.cgi). Nucleotide and protein sequence similarity searches were also performed using BLAST programs at the BLAST server of NCBI (<http://www.ncbi.nlm.nih.gov/blast/blast.cgi>). Pairwise and multiple protein sequence alignments were made with the ALIGN (Wilbur and Lipman, 1983) and CLUSTALW (Thompson *et al.*, 1994) programs, respectively, at the INFOBIOGEN server (<http://www.infobiogen.fr/services/menuserv.html>).

Acknowledgements

We thank I. Cases and C. Weinel for their help in the analysis of the *P. putida* KT2440 genome. This work was supported by EU contract QLK3-CT2000-00170, by the Spanish Ministry of Science and Technology (MCYT) [Red del Consejo

Superior de Investigaciones Científicas (CSIC) de Biorremediación y Fitorremediación] and by grants BMC2000-0125-CO4-02 and GEN2001-4698-CO5-02 from the Comisión Interministerial de Ciencia y Tecnología. J.-I.J. was the recipient of a I3P predoctoral fellowship from the CSIC.

References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Aranda-Olmedo, I., Tobes, R., Manzanera, M., Ramos, J.L., and Marqués, S. (2002) Species-specific repetitive extragenic palindromic (REP) sequences in *Pseudomonas putida*. *Nucleic Acids Res* **30**: 1826–1833.
- Assinder, S.J., and Williams, P.A. (1990) The TOL plasmids: determinants of the catabolism of toluene and the xylenes. *Adv Microbiol Physiol* **31**: 1–69.
- Bagdasarian, M., Lurz, R., Rückert, B., Franklin, F.C.H., Bagdasarian, M.M., and Timmis, K.N. (1981) Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene* **16**: 237–247.
- Bayley, S.A., Duggleby, C.J., Worsey, M.J., Williams, P.A., Hardy, K.G., and Broda, P. (1977) Two modes of loss of the TOL function from *Pseudomonas putida* mt-2. *Mol Gen Genet* **154**: 203–204.
- Bertani, I., Kojic, M., and Venturi, V. (2001) Regulation of the *p*-hydroxybenzoic acid hydroxylase gene (*pobA*) in plant-growth-promoting *Pseudomonas putida* WCS358. *Microbiology* **147**: 1611–1620.
- Cases, I., and de Lorenzo, V. (2001) The black cat/white cat principle of signal integration in bacterial promoters. *EMBO J* **20**: 1–11.
- Collier, L.S., Gaines, G.L.I., and Neidle, E.L. (1998) Regulation of benzoate degradation in *Acinetobacter* sp. strain ADP1 by BenM, a LysR-type transcriptional activator. *J Bacteriol* **180**: 2493–2501.
- Cowles, C.E., Nichols, N.N., and Harwood, C.S. (2000) BenR, a XylS homologue, regulates three different pathways of aromatic acid degradation in *Pseudomonas putida*. *J Bacteriol* **182**: 6339–6346.
- Cuskey, S.M., Peccoraro, V., and Olsen, R.H. (1987) Initial catabolism of aromatic biogenic amines by *Pseudomonas aeruginosa* PAO: pathway description, mapping of mutations, and cloning of essential genes. *J Bacteriol* **169**: 2398–2404.
- D'Argenio, D.A., Segura, A., Coco, W.M., Bünz, P.V., and Ornstun, L.N. (1999) The physiological contribution of *Acinetobacter* PcaK, a transport system that acts upon protocatechuate, can be masked by the overlapping specificity of VanK. *J Bacteriol* **181**: 3505–3515.
- Díaz, E., Ferrández, A., Prieto, M.A., and García, J.L. (2001) Biodegradation of aromatic compounds by *Escherichia coli*. *Microbiol Mol Biol Rev* **65**: 523–569.
- Durham, D.R., and Perry, J.J. (1978) Purification and characterization of a heme-containing amine dehydrogenase from *Pseudomonas putida*. *J Bacteriol* **134**: 837–843.
- Eaton, R.W. (2001) Plasmid-encoded phthalate catabolic pathway in *Arthrobacter keyseri* 12B. *J Bacteriol* **183**: 3689–3703.

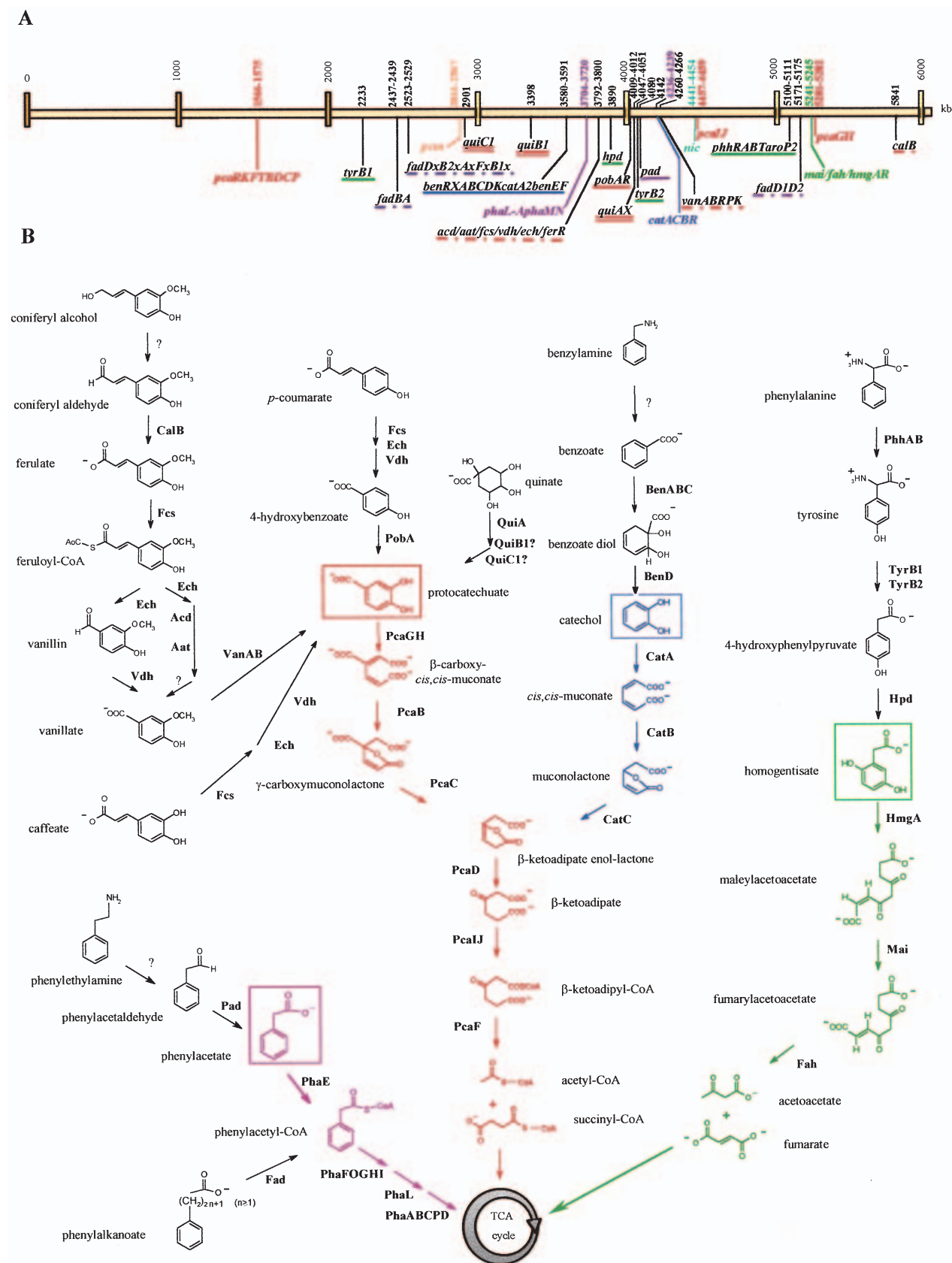


Table 1. The *pca* and *pob* genes and their products from *P. putida* KT2440.

Gene (orf no.) ^a	Gene product (aa) ^b	Related gene products				
		Name	Function	Organism	% Identity/aa	Accession no.
<i>pcaR</i> (06331)	PcaR (291)	PcaR	Transcriptional activator (IclR family)	<i>P. putida</i> PRS2000	96/291	L33795
<i>pcaK</i> (06330)	PcaK (448)	PcaK	4-Hydroxybenzoate transporter	<i>P. putida</i> PRS2000	97/448	U10895
<i>pcaF</i> (06327)	PcaF (400)	PcaF	β -Ketoacyl CoA thiolase	<i>P. putida</i> PRS2000	97/400	U10895
<i>pcaT</i> (06326)	PcaT (429)	PcaT	β -Ketoacyl CoA thiolase	<i>P. putida</i> PRS2000	98/429	U48776
<i>pcaB</i> (06324)	PcaB (450)	PcaB	β -Carboxy- <i>cis,cis</i> -muconate cycloisomerase	<i>P. putida</i> PRS2000	81/407	L17082
<i>pcaD</i> (06322)	PcaD (260)	PcaD	β -Ketoacyl CoA thiolase	<i>Acinetobacter</i> sp. ADP1	43/266	L05770
<i>pcaC</i> (06320)	PcaC (130)	PcaC	γ -carboxymuconolactone decarboxylase	<i>Acinetobacter</i> sp. ADP1	57/134	L05770
<i>pcaP</i> (06317)	PcaP (418)	PhaK	Porin protein	<i>P. putida</i> U	42/417	AF029714
<i>pcaG</i> (01496)	PcaG (201)	PcaG	Protocatechuate 3,4-dioxygenase α subunit	<i>P. putida</i> ATCC23975	98/201	L14836
<i>pcaH</i> (01495)	PcaH (239)	PcaH	Protocatechuate 3,4-dioxygenase β subunit	<i>P. putida</i> ATCC23975	97/239	L14836
<i>pcaI</i> (02302)	PcaI (231)	PcaI	β -Ketoacyl CoA thiolase	<i>P. putida</i> PRS2000	98/231	M88763
<i>pcaJ</i> (02300)	PcaJ (213)	PcaJ	β -Ketoacyl CoA thiolase	<i>P. putida</i> PRS2000	100/213	M88763
<i>pobR</i> (02933)	PobR (292)	PobC	Transcriptional activator (XylS/AraC family)	<i>P. putida</i> WCS358	87/293	AJ251792
<i>pobA</i> (02935)	PobA (395)	PobA	<i>p</i> -Hydroxybenzoate hydroxylase	<i>P. fluorescens</i>	75/394	X68438

a. Indicates the open reading frame number in the complete genome.

b. aa, number of amino acids.

Thus, although the *pca* genes from *P. aeruginosa* and *P. syringae* are arranged in three and four different clusters, respectively, the *pca* genes from *P. fluorescens* are clustered together in the same chromosomal region (Fig. 2). A similar arrangement of all *pca* genes in a single cluster was reported in *Acinetobacter* sp. ADP1 (Harwood and Parales, 1996), and it is also present in the α -proteobacteria *Agrobacterium tumefaciens* and *Caulobacter crescentus*, as well as in the Gram-positive nocardioform actinomycete *Rhodococcus opacus* (Eulberg *et al.*, 1998) (Fig. 2). However, as observed in *P. putida*, *P. aeruginosa* and *P. syringae*, the *pca* genes of some β -proteobacteria, such as *Burkholderia pseudomallei* and *Ralstonia metallidurans* (formerly *Alcaligenes eutrophus*), are arranged in several clusters (Fig. 2). The two pairs of genes *pcaGH* and *pcaIJ* encode separate subunits of a single enzyme (Table 1), and they are co-transcribed in different bacteria (Harwood and Parales, 1996). These gene products in *P.*

fluorescens, especially the PcaIJ proteins, show the lowest amino acid sequence similarity among *pca* gene products of different *Pseudomonas* strains, and this may reflect a different evolutionary origin for this pair of genes in this bacterium. Although the gene order *pcaIJF* is conserved in most of the *pca* clusters, the *pcaF* gene is not linked to *pcaIJ* in *P. putida*, *P. aeruginosa* and *P. syringae* (Fig. 2). Moreover, *pcaD* is usually contiguous to the *pcaC* gene, or fused to the latter as a *pcaL* gene in some bacteria such as *C. crescentus*, *R. metallidurans* and *R. opacus* (Eulberg *et al.*, 1998); however, in *P. syringae*, these two genes are located at different regions of the genome (Fig. 2).

The *cat* genes (Table 2) are usually organized in a single cluster (Harwood and Parales, 1996) (Fig. 3). The *catRcatBCA* gene order is maintained in the *cat* clusters of *P. putida* and *P. aeruginosa* [this arrangement differs from the *catCBA* order given by Kukor *et al.* (1988) in *P.*

Fig. 1. Pathways for the catabolism of aromatic compounds in *P. putida* KT2440.

A. The location of genes and gene clusters encoding the aromatic catabolic pathways is indicated on the complete *P. putida* KT2440 genome. Genes responsible for the four main central pathways are indicated in red (*pca*), blue (*cat*), green (*hmg/fah/mai*) and purple (*pha*). The *pcm* and *nic* clusters are also indicated in orange and light blue colours respectively. Genes encoding the peripheral pathways that lead to the central routes encoded by *pca*, *cat*, *hmg/fah/mai* and *pha* genes are underlined with red, blue, green and purple lines respectively.

B. Predicted biochemical steps for the catabolism of aromatic compounds in *P. putida* KT2440. The names of the metabolites are indicated. The enzymes involved are listed in Tables 1–5 and in the text. ? means that the enzyme encoding such biochemical step is still unknown. The four central aromatic intermediates, i.e. protocatechuate, catechol, homogentisate and phenylacetate, are shown within a red, blue, green and purple box respectively.

Table 2. The *cat* and *ben* genes and their products from *P. putida* KT2440.

Gene (orf no.) ^a	Gene product (aa) ^b	Related gene products				
		Name	Function	Organism	% Identity/aa	Accession no.
<i>catA</i> (02653)	CatA (311)	CatA	Catechol 1,2-dioxygenase	<i>P. putida</i> mt-2	100/311	D37782
<i>catB</i> (02548)	CatB (373)	CatB	<i>cis,cis</i> -muconate lactonizing enzyme (cycloisomerase)	<i>P. putida</i> PRS2000	96/374	M16236
<i>catC</i> (02650)	CatC (96)	CatC	Muconolactone isomerase	<i>P. putida</i> PRS2000	97/96	U12557
<i>catR</i> (02646)	CatR (290)	CatR	Transcriptional activator (LysR family)	<i>P. putida</i> PRS2000	92/289	M33817
<i>benR</i> (03545)	BenR (318)	BenR	Transcriptional activator (XylS/AraC family)	<i>P. putida</i> PRS2000	97/318	AF218267
<i>benX</i> ^c	BenX (313)	ORF589	Unknown	<i>C. burnetii</i>	42/138 ^d	X93204
<i>benA</i> (03542)	BenA (452)	BenA	Benzoate dioxygenase large subunit	<i>P. putida</i> PRS2000	99/452	AF218267
<i>benB</i> (03540)	BenB (161)	BenB	Benzoate dioxygenase small subunit	<i>P. putida</i> PRS2000	97/161	AF218267
<i>benC</i> (03539)	BenC (336)	BenC	Benzoate dioxygenase reductase subunit	<i>P. putida</i> PRS2000	97/336	AF218267
<i>benD</i> (03538)	BenD (253)	BenD	2-Hydro-1,2-dihydroxybenzoate dehydrogenase	<i>P. putida</i> PRS2000	98/253	AF218267
<i>benK</i> (03537)	BenK (442)	BenK	Benzoate transporter	<i>P. putida</i> PRS2000	97/443	AF218267
<i>catA2</i> (03534)	CatA2 (304)	CatA	Catechol 1,2-dioxygenase	<i>P. putida</i> mt-2	77/311	D37782
<i>benE</i> (03532)	BenE (399)	BenE	Membrane protein of unknown function	<i>P. putida</i> PRS2000	93/399	AF218267
<i>benF</i> (03530)	BenF (416)	BenF	Porin-like protein	<i>P. putida</i> PRS2000	96/397	AF218267

a. Indicates the open reading frame number in the complete genome.

b. aa, number of amino acids.

c. This gene has not an orf number in the annotated genome.

d. The % identity was calculated by comparison of a partial 138 amino acid length sequence.

aeruginosa PAO1c]. However, *P. fluorescens* has two sets of *catBCA* genes that lack the *catR* regulatory gene, and they form a pair of gene clusters that also include the *ben* genes (see below; Fig. 3). So far, the *cat* cluster has not been found in *P. syringae*. *R. metallidurans* has three isofunctional hydrolases, two *catD* and one *pcaL* gene products (Figs 2 and 3), which convert β -ketoadipate enol-lactone to β -ketoadipate, the branch convergence point of the β -ketoadipate pathway in this bacterium (Harwood and Parales, 1996). Two *catA* and *catC* genes are also found at different locations in the genome of *R. metallidurans* (Fig. 3). The catechol branch of the β -ketoadipate pathway appears to be present only in some α -proteobacteria such as *Novosphingobium aromaticivorans* (Fig. 3). In *Acinetobacter* sp. ADP1, the two branches never converge, and two independently regulated sets of genes (Figs 2 and 3) encode isofunctional enzymes for the last three steps in the pathway (Harwood and Parales, 1996).

The *catA* gene encodes the catechol 1,2-dioxygenase (pyrocatechase), an intradiol dioxygenase that catalyses the conversion of catechol to *cis,cis*-muconate (Nakai *et al.*, 1995) (Table 2 and Fig. 1B). Interestingly, a second *catA* gene (named *catA2*) that is present within the *ben* cluster for benzoate degradation (see below) has been found in *P. putida* KT2440 (Fig. 3). Although the *catA* gene product is an 311-amino-acid-long protein (Nakai *et al.*,

1995), the *catA2* gene is predicted to encode a protein of 304 amino acids (Table 2). This *catA2* gene is not found in either the *ben* cluster from *P. putida* PRS2000 (Cowles *et al.*, 2000) or the *ben* clusters from *P. aeruginosa* and *P. fluorescens* (Fig. 3). It is known that *Pseudomonas arvilla* C-1 (later reclassified as *P. putida*) has three functional isozymes ($\alpha\alpha$, $\alpha\beta$ and $\beta\beta$) of catechol 1,2-dioxygenase, being the α and β subunits encoded by the *catA_α* and *catA_β* genes respectively (Nakai *et al.*, 1990). Although the *catA_β* gene is homologous to the *catA* gene from *P. putida* mt-2, the *catA_α* gene has not yet been identified (Nakai *et al.*, 1995). It is worth noting that the N-terminal sequence of the α subunit of catechol 1,2-dioxygenase from *P. arvilla* C-1 (Nakai *et al.*, 1990) is homologous (18 identical residues and two conserved substitutions within the first 20 amino acids) to the deduced N-terminal sequence of CatA2 from *P. putida* KT2440, suggesting that *catA2* might encode an active catechol 1,2-dioxygenase not yet reported in this strain. The expression of the *catA2* gene and the physiological role of the CatA2 enzyme in *P. putida* KT2440 remain to be checked.

By analogy with the *cat* cluster in *P. putida* PRS2000, CatR (LysR-type regulatory protein) might activate the expression of *catBCA* genes in KT2440 in response to the inducer *cis,cis*-muconate. On the other hand, PcaR (IcIR-type regulatory protein) might control the β -ketoadipate-

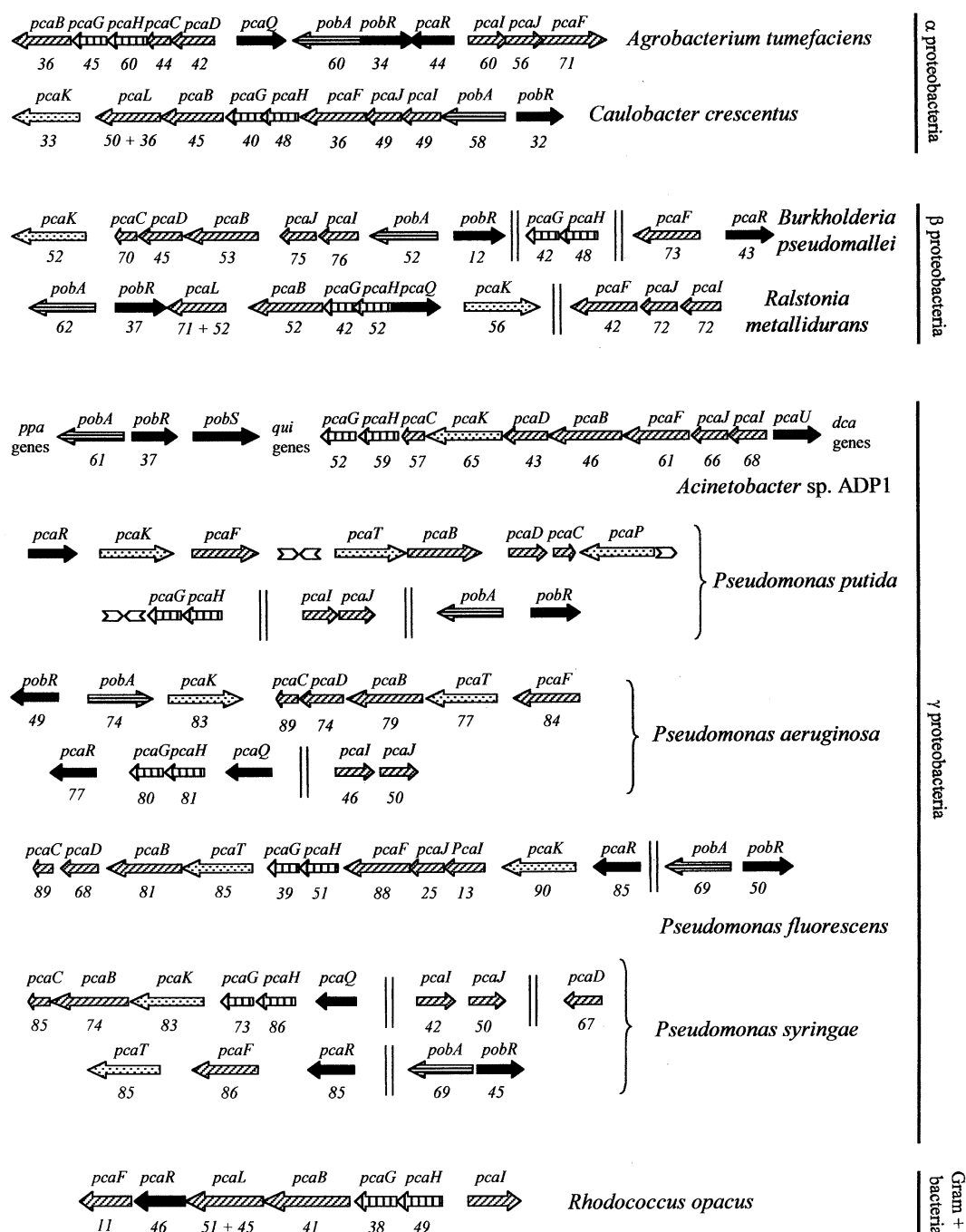


Fig. 2. Gene organization of the *pca* and *pob* clusters of *P. putida* and comparisons with equivalent clusters from other bacteria. Genes (listed in Table 1) are represented by arrows: black (regulatory genes), stippled (transport genes), vertically striped (genes encoding the protocatechuate 3,4-dioxygenase), horizontally striped (genes encoding the *p*-hydroxybenzoate hydroxylase), hatched [catabolic genes of the β -ketoadipate (protocatechuate) pathway]. Arrowheads indicate the *P. putida* REP element. Two vertical lines mean that the genes are not adjacent in the genome. Numbers underneath the arrows indicate the percentage of amino acid sequence identity between the encoded gene product and the equivalent product from *P. putida* KT2440. The two figures underneath *pcaL* genes correspond to the percentage of amino acid sequence identity with the *pcaC* and *pcaD* gene products from *P. putida* KT2440. *pcaQ*, *pobS* and *pcaU* do not have orthologues in *P. putida* KT2440. The sequence of the *pcaI* gene from *R. opacus* is not yet complete. The *ppa*, *qui* and *dca* genes in *Acinetobacter* sp. ADP1 refer to genes for phenylpropanoid/phenylpropenoid, quinate and dicarboxylic acids degradation respectively. The references of the sequences are as follows: *A. tumefaciens* strain C58 (accession no. AE008232 and AE008233) (Goodner *et al.*, 2001); *C. crescentus* strain CB15 (accession no. AE005910) (Nieman *et al.*, 2001); *Acinetobacter* sp. strain ADP1 (accession no. L05770) (Parke *et al.*, 2001); *P. aeruginosa* strain PAO1 (*Pseudomonas* Genome Project at <http://www.pseudomonas.com>) (Stover *et al.*, 2000); *R. opacus* strain 1CP (accession no. AF003947) (Eulberg *et al.*, 1998); *B. pseudomallei*, *R. metallidurans* strain CH34, *P. putida* strain KT2440, *P. fluorescens* strain Pf0-1 and *P. syringae* pv. *tomato* DC3000 (database of unfinished microbial genomes at the NCBI server: http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/genom_table.cgi).

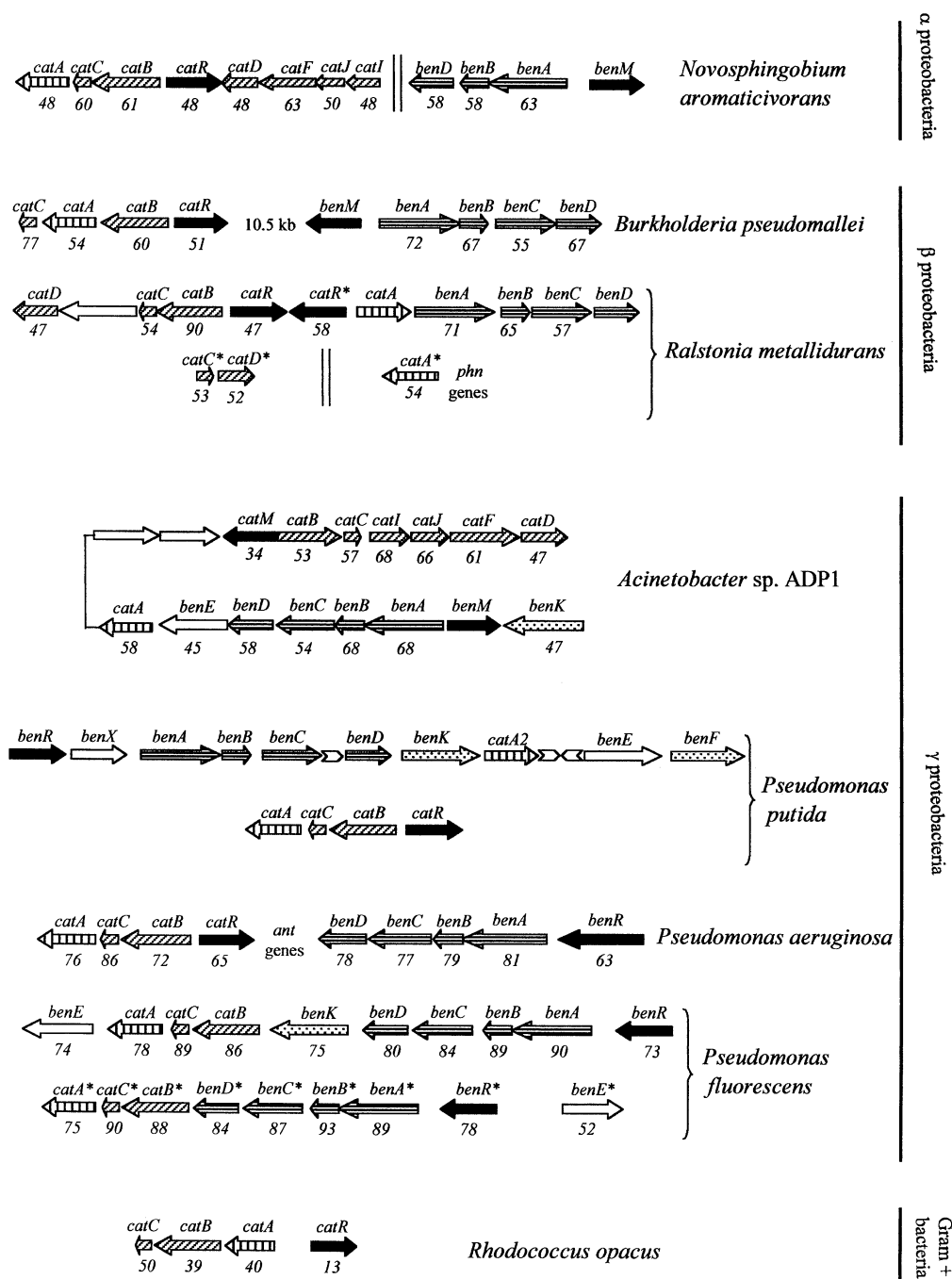


Fig. 3. Gene organization of the *cat* and *ben* clusters of *P. putida* and comparisons with equivalent clusters from other bacteria. Genes (listed in Table 2) are represented by arrows: black (regulatory genes), white (genes of unknown function), stippled (transport genes), vertically striped (genes encoding the catechol 1,2-dioxygenase), horizontally striped (genes encoding the benzoate dioxygenase and dihydrodiol dehydrogenase), hatched (catabolic genes of the catechol pathway). Arrowheads indicate the *P. putida* REP element. Two vertical lines mean that the genes are not adjacent in the genome. Numbers underneath the arrows indicate the percentage of amino acid sequence identity between the encoded gene product and the equivalent product from *P. putida* KT2440. The figures underneath *catD*, *catF*, *catI* and *catJ* in *N. aromaticivorans*, *R. metallidurans* and *Acinetobacter* sp. ADP1 were obtained by comparison with the equivalent *pca* genes from *P. putida* KT2440. *benM* genes do not have an orthologue in *P. putida* KT2440. The sequence of *catA* and *benD* from *R. metallidurans* is not yet complete. Asterisks indicate a second copy of the gene in the genome. The *phn* and *ant* genes in *R. metallidurans* and *P. aeruginosa* refer to genes for phenol and anthranilate (2-aminobenzoate) degradation respectively. The references of the sequences are as follows: *Acinetobacter* sp. strain ADP1 (accession no. AF009224) (Collier *et al.*, 1998); *P. aeruginosa* strain PAO1 (*Pseudomonas* Genome Project at <http://www.pseudomonas.com>) (Stover *et al.*, 2000); *R. opacus* strain 1CP (accession no. X99622) (Eulberg *et al.*, 1997); *N. aromaticivorans*, *B. pseudomallei*, *R. metallidurans* strain CH34, *P. putida* strain KT2440, *P. fluorescens* strain Pf0-1 and *P. syringae* pv. *tomato* DC3000 (database of unfinished microbial genomes at the NCBI server: http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/genom_table.cgi).

dependent inducible expression of genes *pcaRKFTBDCP* and *pcaIJ* required for the conversion of β -carboxy-*cis,cis*-muconate to Krebs cycle intermediates, as already shown in *P. putida* PRS2000 (Harwood and Parales, 1996). Although the regulatory protein controlling the expression of the *pcaGH* genes has not yet been identified in *P. putida*, a gene (*pcaQ*) encoding a putative LysR-type regulator homologous to the PcaQ activator from *A. tumefaciens* (Parke, 1996) is found in the vicinity of the *pcaGH* genes in other *Pseudomonas* strains, such as *P. aeruginosa* and *P. syringae* (Fig. 2).

Based on their sequence similarity to the equivalent genes in *P. putida* PRS2000, the *pcaK* and *pcaT* genes are predicted to encode the transport proteins for *p*-hydroxybenzoate and β -ketoadipate respectively (Harwood and Parales, 1996; Parke *et al.*, 2000) (Table 1). A putative porin-encoding gene (*pcaP*) is also found at the 3' end of the *pca* gene cluster (Table 1 and Fig. 2), and it is adjacent to the *tggABC* genes encoding a solvent efflux pump (Fukimori *et al.*, 1998), a similar arrangement to that observed in *P. putida* DOT-T1 (Ramos *et al.*, 1998).

Peripheral pathways leading to the β -ketoadipate central pathway

Mutants of *P. putida* mt-2 that are unable to convert benzoate into catechol have been isolated (Jeffrey *et al.*, 1992). The *ben* genes responsible for the transformation of benzoate into catechol have been reported in *P. putida* PRS2000 and *Acinetobacter* sp. ADP1 (Collier *et al.*, 1998; Cowles *et al.*, 2000), and *ben* orthologues have been identified at positions 3580–3591 kb in the genome of *P. putida* KT2440 (Fig. 1 and Table 2). Although the *ben* and *cat* clusters are distantly located in the genome of *P. putida* KT2440, they are contiguous in *P. fluorescens* (this species has two *ben*–*cat* clusters), *R. metallidurans* and *Acinetobacter* sp. ADP1 (Collier *et al.*, 1998) and closely linked in *P. aeruginosa* and *B. pseudomallei* (Fig. 3). The *ben* cluster from strain KT2440 shows two unique features that are not observed in *ben* clusters from other *Pseudomonas* strains: (i) in addition to the catabolic genes encoding the benzoate dioxygenase (*benABC*) and benzoate-dihydrodiol dehydrogenase (*benD*) enzymes that convert benzoate into catechol (Fig. 1B), there is a *catA2* gene (see above), the product of which shows significant identity (77%) to the CatA dioxygenase of the catechol branch (Fig. 3 and Table 2); (ii) a gene (*benX*) of unknown function has been inserted between the *benR* and *benA* genes (Fig. 3 and Table 2). Although the *benE* gene encodes a membrane protein of unknown function, the *benK* and *benF* genes are likely to encode a benzoate transporter and a porin respectively (Cowles *et al.*, 2000) (Table 2). *benR* (Fig. 3 and Table 2) is the orthologue of the gene encoding the transcriptional activator of the *ben*

cluster that responds to benzoate in the homologous *ben* cluster from *P. putida* PRS2000 (Cowles *et al.*, 2000). Whereas the expression of the *ben* genes in *Pseudomonas* is controlled by the *benR* gene product, expression of the *ben* genes in *Acinetobacter* is controlled by BenM, a member of the LysR family of regulatory proteins (Collier *et al.*, 1998). The *benABCD* gene order was maintained in most clusters analysed, being the *benR* and *benM* regulatory genes transcribed in the same and in the opposite direction from the catabolic genes respectively (Fig. 3).

Although the catabolism of benzylamine in *P. putida* KT2440 is still unknown, biochemical data indicate that this compound is converted to benzaldehyde and benzoate by some *Pseudomonas* strains (Cuskey *et al.*, 1987). Therefore, a peripheral pathway that oxidates benzylamine into benzoate in *P. putida* KT2440 (Fig. 1B) should be expected, and this pathway needs to be characterized.

The *pobA* and *pobR* genes encode the *p*-hydroxybenzoate hydroxylase (PobA), which converts *p*-hydroxybenzoate into protocatechuate (Fig. 1B), and the cognate transcriptional activator (PobR) in different bacteria (Harwood and Parales, 1996; Bertani *et al.*, 2001). The *pobA* and *pobR* orthologues are located at positions 4009–4012 kb of the *P. putida* KT2440 chromosome (Fig. 1A) and, as observed in most bacteria, they are divergently transcribed (Fig. 2). Unlike PobR from *Acinetobacter* sp. ADP1, which belongs to the IclR family, PobR from *P. putida* KT2440 shows similarity to regulators of the XylS/AraC family such as PobC from *P. putida* WCS358 (Table 1) that responds efficiently to *p*-hydroxybenzoate and weakly to protocatechuate (Bertani *et al.*, 2001). The corresponding PobR proteins from other *Pseudomonas* strains are also members of the XylS/AraC family of regulators (Quinn *et al.*, 2001), indicating that PobR proteins belong to either the IclR family or the XylS/AraC family. Although the *pob* genes are not linked to the *pca* genes in *P. putida*, *P. fluorescens* and *P. syringae*, they are associated in *P. aeruginosa* and other bacteria (Fig. 2).

Quinate catabolism in *Acinetobacter* sp. ADP1 requires the QuiA (quinate dehydrogenase), QuiB (type I dehydroquinate dehydratase) and QuiC (dehydroshikimate dehydratase) enzymes that transform this hydroaromatic compound into protocatechuate (Elsemore and Ornston, 1995). The *qui* cluster from *Acinetobacter* sp. ADP1 also contains a putative porin-encoding gene (*quiX*) and is located adjacent to the *pca* gene cluster (Fig. 2) (Parke *et al.*, 2000; 2001). At positions 4047–4051 kb of the KT2440 genome, we have found a gene (orf 02878) that encodes a putative QuiA-like protein (63% amino acid identity to QuiA from strain ADP1) adjacent to a gene (orf 02876) encoding a putative membrane protein that shows 30% identity to QuiX (Fig. 1A). *quiBC* genes similar to

those of *Acinetobacter* sp. ADP1 were not found in *P. putida* KT2440. However, a gene (*quiB1*, orf 03808) encoding a product that shows 53% identity to the type II dehydroquininate dehydratase (QutE) from *Emericella nidulans* (formerly *Aspergillus nidulans*) (Hawkins *et al.*, 1988) is found at position 3398 kb of the KT2440 genome in the vicinity of the *aroE* gene (encoding a putative shikimate dehydrogenase). On the other hand, a putative *quiC1* (orf 04502) gene product (635 amino acids) that shows 32% identity at its N-terminal end to the QutC 3-dehydroshikimate dehydratase (348 amino acids) from *E. nidulans* (Hawkins *et al.*, 1982) is found at position 2901 kb of the KT2440 genome. Interestingly, the *quiB1quiC1* orthologues from *P. aeruginosa* are arranged together in the vicinity of the *pca* genes. All these data suggest that *quiB1* and *quiC1* might be involved in quinate metabolism in *P. putida* KT2440, indicating that the quinate-degradative enzymes in this bacterium might be different from those reported in *Acinetobacter*, which agrees with previous reports showing differences in quinate catabolism between these two species (Ingledew and Tai, 1972).

Phenylpropenoid compounds (e.g. cinnamate, ferulate, coumarate, etc.) form a vast array of ether and ester bonds in lignin and suberin. The natural turnover of lignin and the chemically accessible suberin are the major sources of phenylpropenoids in the environment (Parke *et al.*, 2000), and these aromatic compounds therefore constitute a common carbon source for microorganisms that colonize the rhizosphere such as *P. putida*. In some bacteria, ferulic acid degradation follows a CoA-dependent non- β -oxidative pathway catalysed by the Fcs

(feruloyl-CoA synthetase) and Ech (enoyl-CoA hydratase/aldolase) proteins, producing vanillin (Overhage *et al.*, 1999; Priefert *et al.*, 2001). Vanillin is further converted to protocatechuate via an aldehyde dehydrogenase (*vdh* gene product) and a demethylase (*vanAB* gene products) (Priefert *et al.*, 1997; Segura *et al.*, 1999). Genes homologous to *fcs*, *ech* and *vdh* have been mapped at positions 3792–3800 kb of the KT2440 genome, and a gene (*ferR*) encoding a putative regulatory protein of the MarR family was also identified at the 3' end of the cluster (Fig. 1 and Table 3). The *vanAB* orthologues have been identified at positions 4260–4266 kb of the KT2440 genome, and they are clustered with a putative transcriptional repressor of the GntR family (*vanR*) (Morawski *et al.*, 2000), a transporter (*vanK*) and a porin (*vanP*) (D'Argenio *et al.*, 1999) (Fig. 1 and Table 3). In the *Pseudomonas* strains analysed, i.e. *P. putida* KT2440, *P. putida* WCS358 (Venturi *et al.*, 1998), *P. syringae* and *Pseudomonas* sp. HR199 (Overhage *et al.*, 1999), the *fcs/ech/vdh* genes form a cluster that is not linked to the *van* cluster. A similar situation is found in *Acinetobacter* sp. ADP1, and it was suggested that this gene organization would facilitate the appearance of spontaneous *van*-deficient strains in natural *Acinetobacter* populations, which might allow the production of vanillate from ferulate as a chemical signal between plants and bacteria (Segura *et al.*, 1999). Interestingly, two additional genes, *aat* (encoding a putative β -ketothiolase) and *acd* (encoding a putative acyl-CoA dehydrogenase) cluster with the *ech*, *vdh* and *fcs* genes (Fig. 1A and Table 3), and they could be responsible for a CoA-dependent β -oxidative pathway of ferulic acid degra-

Table 3. The genes and products for the catabolism of phenylpropenoid compounds in *P. putida* KT2440.

Gene (orf no.) ^a	Gene product (aa) ^b	Related gene products				
		Name	Function	Organism	% Identity/aa	Accession no.
<i>ferR</i> (03214)	FerR (156)	SlyA	Transcriptional activator (MarR family)	<i>S. typhimurium</i>	29/146	AJ010965
<i>fcs</i> (03219)	Fcs (589)	Fcs	Feruloyl-CoA synthetase	<i>Pseudomonas</i> sp. HR199	75/589	AJ238746
<i>ech</i> (03215)	Ech (276)	Ech	<i>p</i> -Hydroxycinnamoyl-CoA hydratase/lyase	<i>P. fluorescens</i> AN103	92/276	Y13067
<i>vdh</i> (03212)	Vdh (482)	Vdh	Vanillin dehydrogenase	<i>Pseudomonas</i> sp. HR199	80/481	Y11520
<i>aat</i> (03221)	Aat (431)	Aat	Putative 2-ketothiolase	<i>Pseudomonas</i> sp. HR199	65/431	AJ238746
<i>acd</i> (03223)	Acid (609)	RSc0473	Putative acyl-CoA dehydrogenase	<i>R. solanacearum</i>	45/595	AL646059
<i>vanR</i> (02608)	VanR (237)	VanR	Transcriptional repressor (GntR family)	<i>Acinetobacter</i> sp. ADP1	50/251	AF009672
<i>vanP</i> (02615)	VanP (417)	OpdK	Putative porin	<i>P. aeruginosa</i> PAO1	74/417	AE004903
<i>vanK</i> (02614)	VanK (446)	VanK	Transporter of aromatic compounds	<i>Acinetobacter</i> sp. ADP1	56/448	AF009672
<i>vanA</i> (02612)	VanA (355)	VanA	Vanillate-O-demethylase oxygenase subunit	<i>P. putida</i> WCS358	87/353	Y14759
<i>vanB</i> (02609)	VanB (316)	VanB	Vanillate-O-demethylase reductase subunit	<i>P. putida</i> WCS358	85/315	Y14759
<i>calB</i> (00705)	CalB (476)	CalB	Coniferyl aldehyde dehydrogenase	<i>Pseudomonas</i> sp. HR199	44/481	AJ006231

a. Indicates the open reading frame number in the complete genome.

b. aa, number of amino acids.

dation that has been described in some organisms (Priefert *et al.*, 2001) (Fig. 1B). Catabolism of *p*-coumaric acid and caffeic acid by *P. putida* KT2440 may proceed via *p*-hydroxybenzoate and protocatechuate, respectively, through the action of the Fcs, Ech and Vdh enzymes (Fig. 1B), as already shown in some other *Pseudomonas* strains (Venturi *et al.*, 1998; Mitra *et al.*, 1999). The catabolism of coniferyl alcohol in *Pseudomonas* sp. HR199 involves its conversion into ferulic acid by an alcohol dehydrogenase (CalA) and an aldehyde dehydrogenase (CalB) (Overhage *et al.*, 1999). Whereas a gene whose product shows 44% amino acid sequence identity to CalB from *Pseudomonas* sp. HR199 is found at position 5841 kb in the KT2440 genome (Fig. 1 and Table 3), a *calA* orthologue could not be identified in this bacterium.

The phenylacetyl-CoA catabolon

The *pha* genes from *P. putida* KT2440 are homologous to the genes involved in phenylacetate degradation in *P. putida* U, and both clusters show the same organization (Luengo *et al.*, 2001). The *pha* cluster from *P. putida* is organized in four discrete DNA segments (Fig. 4), which are predicted to encode six different functional units: *phaABCPD* and *phaE* (β -oxidation and activation of phenylacetic acid), *phaFOGHI* (hydroxylation of the aromatic ring), *phaJK* and *phaL* (phenylacetic acid transport and dearomatization of the ring) and *phaMN* (regulation of the *pha* cluster) (Fig. 1 and Table 4). The aerobic catabolism of phenylacetic acid represents a novel hybrid pathway

that does not follow the conventional routes for biodegradation of aromatic compounds and the first step in which is the activation of phenylacetic acid to phenylacetyl-CoA by the action of a phenylacetyl-CoA ligase. Then, phenylacetyl-CoA suffers an oxygenation reaction followed by cleavage of the aromatic ring and a β -oxidation-like pathway of the ring cleavage product (Ferrández *et al.*, 1998; Luengo *et al.*, 2001; Mohamed *et al.*, 2002). The gene arrangement of the *pha* cluster (also named *paa* cluster in some bacteria) differs between different bacteria (Fig. 4), suggesting that various DNA rearrangements have occurred during its evolution in each particular host. Although the *phaE* (*paaK*) and *phaFOGHI* (*paaABCDE*) genes are usually present in all bacterial species, even though they do not form a cluster with the rest of the *pha* (*paa*) genes in some bacteria, the *phaJ* and *phaK* genes of *P. putida*, encoding a permease and a specific channel-forming protein for the uptake of phenylacetic acid, respectively, are absent in the *pha* (*paa*) clusters of most bacteria (Fig. 4). The regulatory *phaN* (*paaX*) gene encoding a transcriptional repressor of the GntR family is linked to the *phaM* (*paaY*) gene of unknown function, an arrangement also observed in enteric bacteria (Díaz *et al.*, 2001). Interestingly, *Azoarcus evansii* and *B. pseudomallei* (β -proteobacteria) present a putative regulatory protein (*paaR* gene product) of the TetR family instead of a *phaN* (*paaX*) orthologue (Mohamed *et al.*, 2002) (Fig. 4). The β -oxidation-like functional unit encoded by *phaABCPD* (*paaFGHIJ*) genes shows the highest diversity (Fig. 4), suggesting that, in some bacteria, the missing gene prod-

Table 4. The *pha* genes and their products from *P. putida* KT2440.

Gene (orf no.) ^a	Gene product (aa) ^b	Homologous protein from <i>P. putida</i> U (Acc. no. AF029714)		
		Name	Function	% Identity/aa
<i>phaM</i> (03328)	PhaM (199)	PhaM	Putative regulatory protein	97/199
<i>phaN</i> (03327)	PhaN (307)	PhaN	Transcriptional repressor (GntR family)	96/307
<i>phaA</i> (03329)	PhaA (257)	PhaA ^c	Predicted enoyl-CoA hydratase/isomerase I	89/257
<i>phaB</i> (03331)	PhaB (263)	PhaB	Predicted enoyl-CoA hydratase/isomerase II	94/263
<i>phaC</i> (03333)	PhaC (505)	PhaC	Predicted hydroxyacyl-CoA dehydrogenase	91/505
<i>phaP</i> (03334)	PhaP (146)	PhaP ^{c,d}	Predicted thioesterase	93/146
<i>phaD</i> (03336)	PhaD (406)	PhaD ^c	Predicted β -ketoacyl-CoA thiolase	96/406
<i>phaE</i> (03337)	PhaE (439)	PhaE	Phenylacetyl-CoA ligase	96/439
<i>phaF</i> (03338)	PhaF (329)	PhaF	Component of a predicted oxygenation complex	98/329
<i>phaO</i> (03339)	PhaO (98)	PhaO ^e	Component of a predicted oxygenation complex	99/98
<i>phaG</i> (03340)	PhaG (252)	PhaG	Component of a predicted oxygenation complex	93/252
<i>phaH</i> (03342)	PhaH (177)	PhaH ^f	Component of a predicted oxygenation complex	81/199
<i>phaI</i> (03344)	PhaI (358)	PhaI ^f	Component of a predicted oxygenation complex	95/311
<i>phaJ</i> (03347)	PhaJ (520)	PhaJ	Phenylacetate transporter	97/520
<i>phaK</i> (03349)	PhaK (417)	PhaK	Phenylacetate porin	96/417
<i>phaL</i> (03351)	PhaL (688)	PhaL	Predicted dearomatizing protein	96/688

a. Indicates the open reading frame number in the complete genome.

b. aa, number of amino acids.

c. The length of these proteins has been reassigned in this work based on sequence comparison analyses with homologous proteins from other bacteria.

d. Díaz *et al.* (2001).

e. Luengo *et al.* (2001).

f. The region of the *pha* cluster from *P. putida* U corresponding to the 3'-end of *phaH* and the 5'-end of *phaI* differs from that of *P. putida* KT2440.

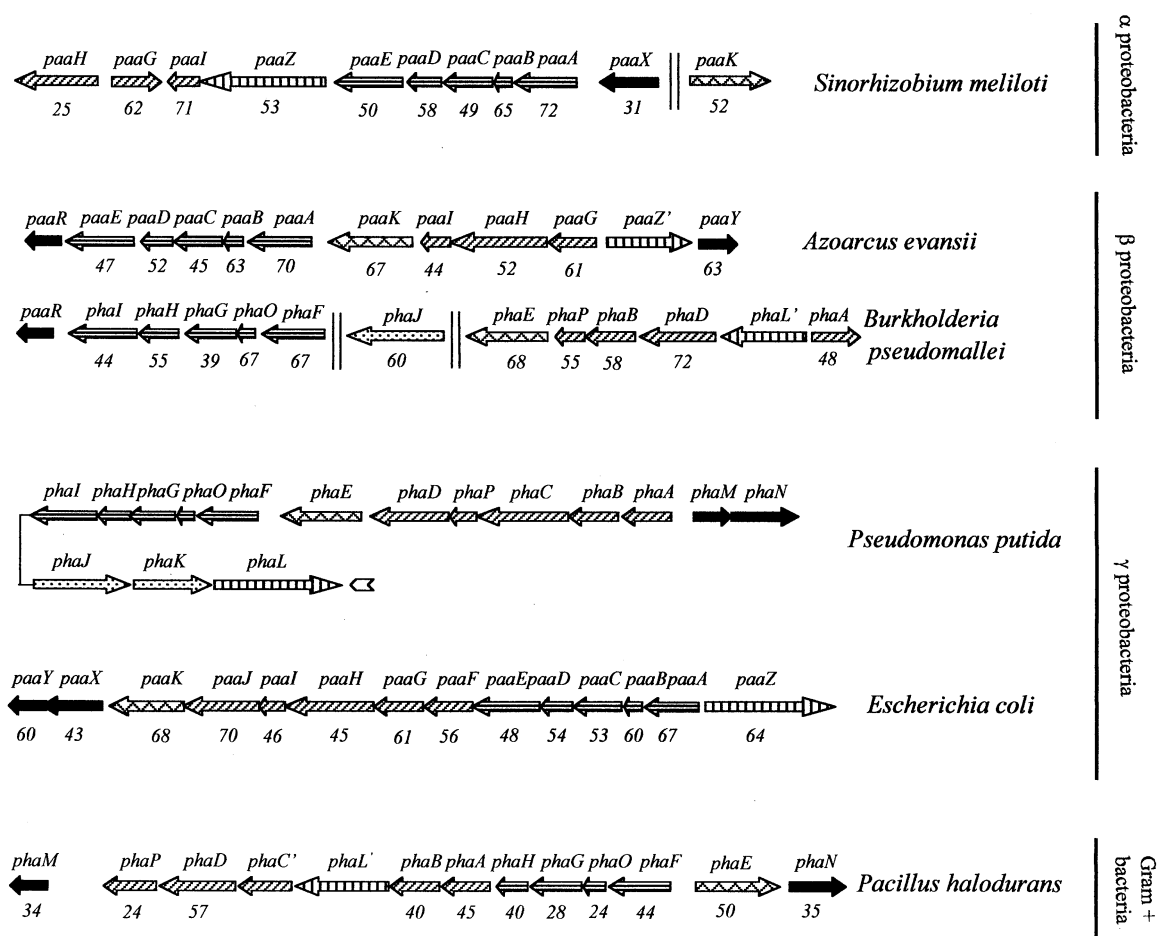


Fig. 4. Gene organization of the *pha* cluster of *P. putida* and comparisons with equivalent clusters from other bacteria. Genes (listed in Table 4) are represented by arrows: black (regulatory genes), stippled (transport genes), vertically striped (genes involved in the presumed dearomatization step), horizontally striped (genes encoding the multicomponent phenylacetyl-CoA oxygenase), hatched (genes encoding the β -oxidation-like functional unit), cross-hatched (genes encoding the phenylacetyl-CoA ligase). Arrowheads indicate the *P. putida* REP element. Two vertical lines mean that the genes are not adjacent in the genome. Numbers underneath the arrows indicate the percentage of amino acid sequence identity between the encoded gene product and the equivalent product from *P. putida* KT2440. Note the existence of two different nomenclatures (*pha* and *paa* clusters): *phaE* corresponds to *paaK*, *phaFOGHI* correspond to *paaABCDE*, *phaL* corresponds to *paaZ*, *phaABCPD* correspond to *paaFGHIJ*, and *phaMN* correspond to *paaYX* respectively. *paaZ'*, *phaL'* and *phaC'* indicate a 3' end truncated gene. *paaR* genes do not have an orthologue in *P. putida* KT2440. The references of the sequences are as follows: *S. meliloti* strain 1021 (accession no. AL603647) (Galibert *et al.*, 2001); *A. evansii* strain KB740 (accession no. AF176259, AJ278756) (Mohamed *et al.*, 2002); *E. coli* W (accession no. X97452) (Ferrández *et al.*, 1998); *B. halodurans* strain C-125 (accession no. AP001507) (Takami *et al.*, 2000); *B. pseudomallei* and *P. putida* strain KT2440 (database of unfinished microbial genomes at the NCBI server: http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/genom_table.cgi).

ucts may be replaced by similar enzymes from other β -oxidation pathways in the cell. It is worth noting that the *phaL* (*paaZ*) gene product of *A. evansii* and *B. pseudomallei* and that of *Bacillus halodurans* are C-terminally truncated compared with the orthologues from the other species (Mohamed *et al.*, 2002) (Fig. 4). Gene context conservation of a higher order than operons has been called uberoperon (Lathe *et al.*, 2000), and the genes responsible for bacterial phenylacetic acid degradation constitute a clear example of such a conserved context (Díaz *et al.*, 2001).

The term catabolon defines a complex functional unit

integrated by different catabolic pathways that catalyse the transformation of structurally related compounds into a common catabolite (Luengo *et al.*, 2001). The phenylacetyl-CoA catabolon in *P. putida* KT2440 encompasses the routes involved in the transformation of 2-phenylethylamine, phenylacetic acid and *n*-phenylalkanoic acids containing an even number of carbon atoms into phenylacetyl-CoA (Fig. 1B). Phenylethylamine is converted through phenylacetaldehyde into phenylacetic acid in different microorganisms (Hacisalihoglu *et al.*, 1997; Díaz *et al.*, 2001); however, a gene homologous to those encoding aromatic amine oxidases (enzymes that convert

phenylethylamine into phenylacetaldehyde) could not be identified in the genome of *P. putida* KT2440. Nevertheless, some *Pseudomonas* strains convert aromatic biogenic amines into the corresponding aromatic aldehydes via an amine dehydrogenase instead of via an amino oxidase (Durham and Perry, 1978; Iwaki *et al.*, 1983; Cuskey *et al.*, 1987). Whether phenylethylamine is degraded via phenylacetaldehyde by the action of an amine dehydrogenase in *P. putida* KT2440 remains to be confirmed. Phenylacetaldehyde is then oxidized to phenylacetic acid by a phenylacetaldehyde dehydrogenase (Pad) enzyme (Díaz *et al.*, 2001). Although there are several genes in the chromosome of *P. putida* KT2440 that show similarity to aryl aldehyde dehydrogenase-encoding genes, that located at position 4142 kb of the genome (orf 02745) shows the highest identity to *pad* genes from other bacteria and might encode the corresponding phenylacetaldehyde dehydrogenase from this bacterium (Fig. 1).

The degradation of *n*-phenylalkanoic acids in *P. putida* requires their activation to CoA thioesters by an acyl-CoA synthetase encoded by *fadD*. Subsequently, an acyl-CoA dehydrogenase (*fadF* gene product) catalyses the formation of a double bond at position 2 of the aliphatic chain and, finally, a protein complex (FadAB) with five enzymatic activities catalyses the release of acetyl-CoA units (Olivera *et al.*, 2001). Although the FadAB complex catalyses the formation of phenylacetyl-CoA from phenylalkanoates containing an even number of carbon atoms, the degradation of phenylalkanoates with an odd number of carbon atoms produces *trans*-cinnamoyl-CoA, which cannot be catabolized further and is excreted as cinnamic acid (Olivera *et al.*, 2001). The catabolism of *n*-phenylalkanoic

acids in *P. putida* U is carried out by two sets of β -oxidation enzymes: whereas the β_I oxidation set (*fadBA* and *fadD1fadD2* genes) is constitutive and catalyses a very efficient degradation, the β_{II} set (genes *fadDxfadB2xfadAxfadFxfadB1x*) is only expressed when some of the genes encoding the β_I enzymes are mutated, and it catabolizes *n*-phenylalkanoates with an acyl moiety longer than four carbons (Olivera *et al.*, 2001). Homologous *fadBA*, *fadD1fadD2* and *fadDxfadB2xfadAxfadFxfadB1x* genes have been identified at positions 2437–2439 kb, 5171–5175 kb and 2523–2529 kb, respectively, in the genome of *P. putida* KT2440, and they are likely to be responsible for the catabolism of *n*-phenylalkanoates in this bacterium (Fig. 1).

The homogentisate central pathway and the catabolism of phenylalanine and tyrosine

At positions 5241–5245 kb of the *P. putida* KT2440 genome, there is a cluster of genes that show similarity to those involved in homogentisic acid degradation in *S. meliloti* (Milcamps and de Bruijn (1999) and *E. nidulans* (Fernández-Cañón and Peñalva, 1998). The *hmgA*, *mai* and *fah* genes from *P. putida* KT2440 are likely to encode the homogentisate dioxygenase, maleylacetoacetate isomerase and fumarylacetoacetate hydrolase, respectively, that convert homogentisate into fumarate and acetoacetate (Fig. 1B and Table 5). A putative regulatory gene, *hmgR*, is divergently transcribed from the catabolic genes and encodes a protein from the IclR family of transcriptional regulators (Fig. 5 and Table 5). A gene arrangement similar to that found within the homogenti-

Table 5. The genes and products for the catabolism of homogentisate and aromatic amino acids in *P. putida* KT2440.

Gene (orf no.) ^a	Gene product (aa) ^b	Related gene products				
		Name	Function	Organism	% Identity/aa	Accession no.
<i>hmgR</i> (01552)	HmgR (277)	PA2010	Putative transcriptional regulator (IclR family)	<i>P. aeruginosa</i> PAO1	74/267	AE004627
<i>hmgA</i> (01553)	HmgA (433)	HmgA	Homogentisate dioxygenase	<i>S. meliloti</i> 1021	56/453	AF109131
<i>fah</i> (01554)	Fah (430)	Pha	Fumarylacetoacetate hydrolase	<i>H. sapiens</i>	47/419	M55150
<i>mai</i> (01555)	Mai (210)	Mai	Maleylacetoacetate isomerase	<i>M. musculus</i>	43/216	AF093418
<i>phhR</i> (01787)	PhhR (519)	PhhR	Transcriptional activator (NtrC family)	<i>P. aeruginosa</i> PAO1	86/518	U62581
<i>phhA</i> (01785)	PhhA (262)	PhhA	Phenylalanine hydroxylase	<i>P. aeruginosa</i> PAO1	84/262	M88627
<i>phhB</i> (01784)	PhhB (118)	PhhB	Pterin 4a-carbinolamine dehydratase	<i>P. aeruginosa</i> PAO1	86/118	M88627
<i>phhT</i> (01781)	PhhT (400)	PA1993	Putative transport protein	<i>P. aeruginosa</i> PAO1	71/402	AE004625
<i>aroP2</i> (01778)	AroP2 (478)	AroP	Aromatic amino acid permease	<i>E. coli</i> K-12	65/457	U87285
<i>hpd</i> (03099)	Hpd (358)	HPPD	<i>p</i> -hydroxyphenylpyruvate dioxygenase	<i>P. fluorescens</i> A32	88/357	1CJX_A-D
<i>tyrB1</i> (05414)	TyrB1 (398)	TyrB	Aromatic amino acid aminotransferase	<i>E. coli</i> K-12	50/397	AF029714
<i>tyrB2</i> (02841)	TyrB2 (398)	TyrB	Aromatic amino acid aminotransferase	<i>E. coli</i> K-12	70/397	AF029714

a. Indicates the open reading frame number in the complete genome.

b. aa, number of amino acids.

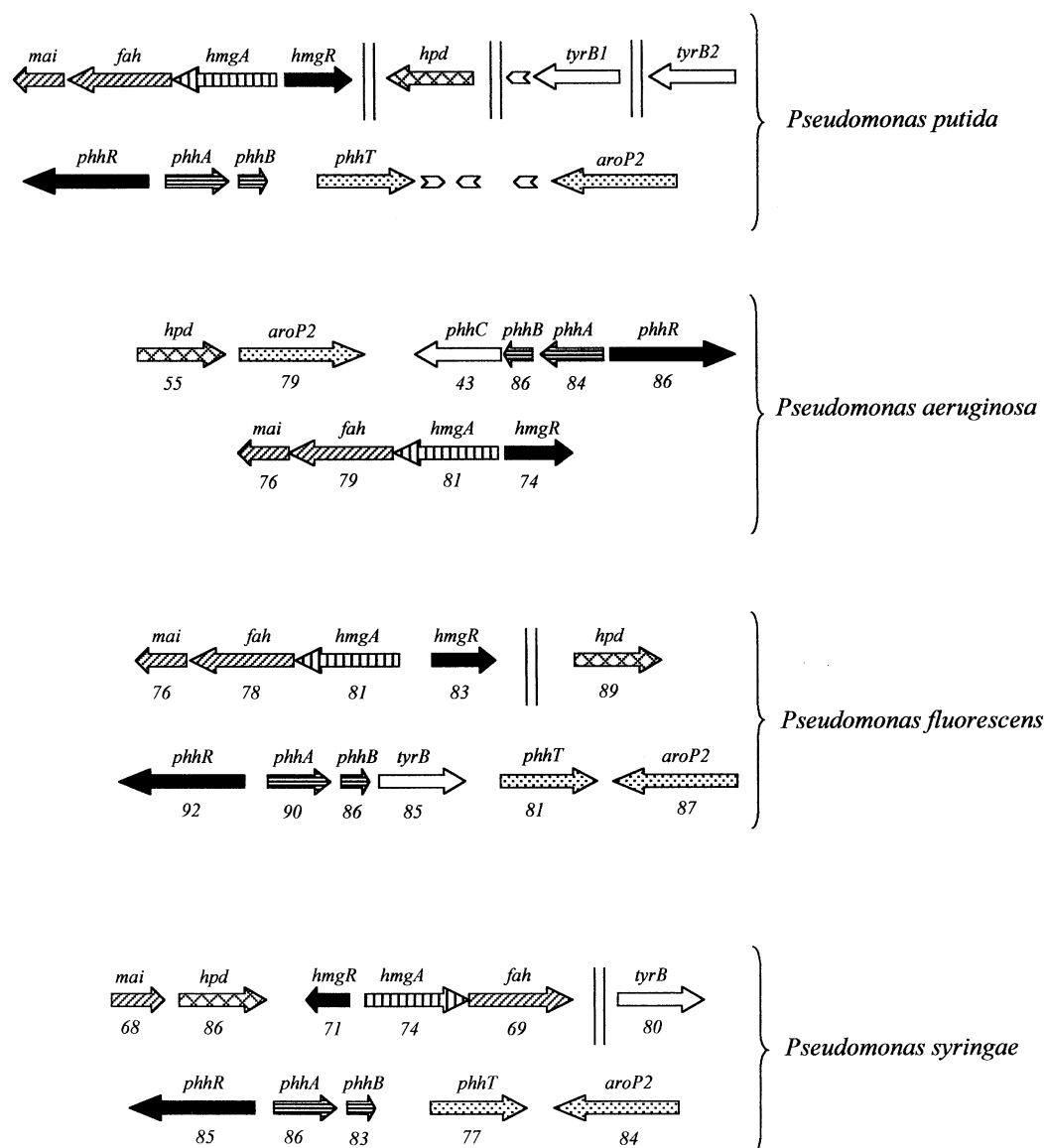


Fig. 5. Gene organization of the clusters encoding the homogentisate and phenylalanine/tyrosine catabolic pathways of *P. putida* and comparisons with equivalent clusters from other *Pseudomonas* species. Genes (listed in Table 5) are represented by arrows: black (regulatory genes), stippled (transport genes), vertically striped (genes encoding the homogentisate dioxygenase), horizontally striped (genes encoding the phenylalanine hydroxylase), hatched (catabolic genes of the homogentisate pathway), cross-hatched (genes encoding the *p*-hydroxyphenylpyruvate dioxygenase), white (genes encoding aromatic amino acid aminotransferases). Arrowheads indicate the *P. putida* REP element. Two vertical lines mean that the genes are not adjacent in the genome. Numbers underneath the arrows indicate the percentage of amino acid sequence identity between the encoded gene product and the equivalent product from *P. putida* KT2440. Values underneath *tyrB* and *phhC* genes were obtained by comparison with the *tyrB1* gene product of *P. putida* KT2440. The references of the sequences are as follows: *P. aeruginosa* strain PAO1 (*Pseudomonas* Genome Project at <http://www.pseudomonas.com>) (Stover *et al.*, 2000); *P. putida* strain KT2440, *P. fluorescens* strain Pf0-1 and *P. syringae* pv. *tomato* DC3000 (database of unfinished microbial genomes at the NCBI server: http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/genom_table.cgi).

sate cluster of *P. putida* KT2440 is also observed in the homologous clusters from *P. aeruginosa* and *P. fluorescens* (Fig. 5), as well as in the cluster from *S. melliloti* (Milcamps and de Bruijn, 1999).

The homogentisate pathway is the central route through which tyrosine and phenylalanine are mineralized in many

bacteria. The genes responsible for the peripheral pathway of the phenylalanine and tyrosine catabolism in *P. aeruginosa* are known (Song *et al.*, 1999), and we have identified the cognate orthologues in *P. putida* KT2440. Thus, at positions 5100–5111 kb of the *P. putida* KT2440 genome, there is a gene cluster (*phh*) encoding the puta-

tive pterin-dependent phenylalanine hydroxylase (*phhA*) that converts phenylalanine into tyrosine, the carbinolamine dehydratase (*phhB*) involved in regeneration of the pterin cofactor, the putative σ^{54} -dependent transcriptional activator (*phhR*) of the *phh* operon (Song and Jensen, 1996) and a potential transport protein (*phhT*) close to a gene (*aroP2*) encoding a general aromatic amino acid permease (Figs 1 and 5, and Table 5). In *P. aeruginosa*, the *phhC* gene encodes a tyrosine aminotransferase that transforms tyrosine into 4-hydroxyphenylpyruvic acid and is essential for the catabolism of either tyrosine or phenylalanine (Gu *et al.*, 1998). There is no homologue to *phhC* in *P. putida* KT2440, however, the tyrosine aminotransferase activity is likely to be accomplished by the products of the *tyrB1* and/or *tyrB2* genes located at positions 2233 kb and 4080 kb of the genome in this bacterium (Fig. 1 and Table 5). Whereas in *P. aeruginosa* and *P. fluorescens*, the *phh* genes form a cluster, the *tyrB* genes are not linked to the *phhRABT* cluster in *P. putida* and *P. syringae* (Fig. 5). On the other hand, the *hpd* gene at position 3890 kb of the *P. putida* KT2240 genome may encode the putative *p*-hydroxyphenylpyruvic dioxygenase that converts 4-hydroxyphenylpyruvate into homogentisate (Serre *et al.*, 1999) (Fig. 1 and Table 5). The role of the *hpd* gene product coupling the catabolism of aromatic amino acids with the homogentisate central pathway might explain the location of *hpd* within the homogentisate cluster, such as in *P. syringae*, or within the *phh* cluster, such as in *P. aeruginosa* (Fig. 5).

Evolutionary considerations and general conclusions

The G+C content of the four central clusters involved in the catabolism of aromatic compounds in *P. putida* KT2440, i.e. *cat*, *pca*, *pha* and *mal/fah/hmg*, averaged 64.2%, 63.1%, 63.5% and 64.6% respectively. Gene clusters encoding the peripheral pathways also show a G+C content ranging from 60% to 65%. These values are close to the mean G+C content (61%) of the genomic *P. putida* KT2440 DNA, suggesting that these sets of genes have been imprisoned within the chromosome of this bacterium over a long period of evolution.

The distribution of the aromatic catabolic clusters along the *P. putida* KT2440 chromosome reveals that the region (about 2400 kb) flanking the replication origin (position 0 kb) is almost devoid of genes related to the catabolism of aromatic compounds (Fig. 1A). This situation contrasts with that observed in other bacteria such as *Escherichia coli* (Díaz *et al.*, 2001) or *P. aeruginosa* (data not shown), in which the aromatic catabolic clusters are spread throughout the chromosome. Nevertheless, we have not observed in *P. putida* KT2440 the existence of a supraoperonic clustering of catabolic genes (catabolic island) that channel different aromatic compounds into a common

central pathway such as that reported for the *pca-qui-pob-ppa* clusters (suberon) in *Acinetobacter* sp. ADP1 (Parke *et al.*, 2001) (Fig. 2). In contrast, out of the four *Pseudomonas* species whose genomes are known (*P. aeruginosa* and *P. putida*) or being sequenced (*P. fluorescens* and *P. syringae*), *P. putida* shows the lowest level of linkage between genes involved in the same aromatic catabolic pathway. For instance, in *P. putida* KT2440, the *ben* and *cat* genes are not associated, the *pca* genes are arranged in three different clusters, and none of them is associated with the *pob* cluster, and the *phhC* and *hpd* genes are not linked to the *phh* genes (Figs 2, 3, and 5).

Some catabolic clusters from *P. putida* KT2440 show the presence of a repetitive extragenic palindromic (REP) sequence previously reported in *P. putida* strains (Houghton *et al.*, 1995; Aranda-Olmedo *et al.*, 2002). This 35 bp REP sequence is found: (i) as a single element at the 3' end of the *benC*, *phaL*, *aroP2* and *tyrB1* genes and at the 5' end of the *pcaP* gene (Figs 2–5); (ii) as pairs of convergent elements at the 3' end of the *pcaG*, *pcaF* and *phhT* genes and at the 5' end of *benE* (Figs 2, 3 and 5); (iii) as pairs of divergent elements in the *vanR-vanP* intergenic region. Although REP sequences are involved in several functions in enterobacteria, e.g. mRNA stabilization, chromosome organization, insertion of genetic elements and binding site for different proteins, the bacterial function of the REP sequence in *P. putida* has not yet been identified (Aranda-Olmedo *et al.*, 2002). It is worth noting that the location of the REP sequence associated with some of the aromatic catabolic clusters in *P. putida* is strain specific. Thus, although the *pha* cluster from strain KT2440 contains a single REP sequence downstream of the *phaL* gene, the *pha* cluster from *P. putida* U contains two convergent REP elements in the *phal-phaJ* intergenic region (data not shown). On the other hand, whereas the two inverted REP sequences upstream of the *benE* gene in *P. putida* KT2240 are not present in *P. putida* PRS2000, the latter contains a REP element at the 3' end of the *catR* gene (Houghton *et al.*, 1995) that is absent from the *cat* cluster of *P. putida* KT2440. A different *P. putida* strain (*P. putida* RB1) shows the *cat*-associated REP element in the *catB-catC* intergenic region (Houghton *et al.*, 1995). Therefore, although the genes are highly conserved among different *P. putida* strains, REP sequences appear to contribute significantly to genomic diversity within this species. Interestingly, the 35 kb region containing the *xyl* genes involved in the catabolism of toluene/xylene from plasmid pWW0 (accession no. AJ344068) of the parental *P. putida* mt-2 strain does not contain any REP element similar to the chromosomal one described above, which might be indicated that the *xyl* cluster bracketed by direct repeated copies of IS1246 (Assinder and Williams, 1990) might be originated in a different bacterial strain and then be recruited to the hypothetical pWW0 ancestor. In this

sense, the low G+C content (50.2%) of the *xylUWCMABN* upper operon from pWW0 also suggests that at least some *xyl* genes may have been recruited from outside the *P. putida* species. Despite *P. putida* mt-2 harbours two equivalent set of genes, i.e. the chromosomal *benABCD* genes and the pWW0-encoded *xylXYZL* genes (overall nucleotide sequence identity of 75%), which encode homologous dioxygenases and dihydrodiol dehydrogenases for the conversion of benzoate into catechol, both sets appear to be stably maintained in the cell, and there are no reports about genetic exchange between homologous genes. Near the 3' end of the *ben* cluster from *P. putida* KT2440, there is a gene encoding a putative maturase-related protein of group II introns that is identical to that reported previously in the vicinity of catabolic genes involved in the degradation of *p*-cresol in *Pseudomonas alcaligenes* (Yeo *et al.*, 1997), suggesting the involvement of group II introns in the evolution of catabolic functions, in much the same way as other mobile genetic elements.

The majority of the aromatic catabolic clusters from *P. putida* KT2440 contain regulatory and transport genes, suggesting that both the uptake of the compound inside the cell and the inducible expression of the catabolic genes are important control factors for the catabolism of aromatic compounds in this bacterium. Most of the predicted inner membrane transport proteins from the aromatic catabolic pathways of *P. putida* KT2440, i.e. PcaK, PcaT, BenK, VanK and PhhT, belong to the major facilitator superfamily (MFS) of transporters, being PhaJ and AroP2 members of solute:sodium symporter (SSS) and amino acid–polyamine–choline (APC) families respectively (Saier, 1998). Aromatic transporters can be accompanied by outer membrane porins such as in the *pca*-, *ben*-, *van*- and *pha*-encoded pathways (Tables 1–4). Although aromatic compounds can enter the cells by passive diffusion when present at high concentrations, active transport increases the efficiency and rate of substrate acquisition in natural environments where these compounds are present at low concentrations (Nichols and Harwood, 1997). Moreover, as already shown with the *p*-hydroxybenzoate transport protein (PcaK) from *P. putida* PRS2000, the aromatic transporters can be involved in the ability of bacteria with a motile life style to sense and swim towards the aromatic compounds (chemotaxis) (Harwood *et al.*, 1994; Harwood and Parales, 1996; Parales and Harwood, 2002).

The regulatory mechanisms that control the expression of the genes responsible for the catabolism of aromatic compounds appear to be highly diverse in *P. putida* KT2440. Thus, a global analysis of the genome allowed us to predict the existence of transcriptional activators from the XylS/AraC family (BenR, PobR), IclR family (PcaR), LysR family (CatR), NtrC family (PhhR) and MarR family (FerR), as well as transcriptional repressors from

the GntR family (VanR and PhaN) and a regulatory protein of unknown activity from the IclR family (HmgR) (see above). Moreover, there might be cross-talk between different regulatory systems. Thus, benzoate degradation in *P. putida* mt-2 can proceed via the plasmid-encoded *meta*-cleavage pathway or the chromosomally encoded *ortho*-cleavage pathway (see above). As reported in *P. putida* PRS2000, BenR participates as an activator of benzoate degradation via *ortho*-ring fission (*ben* genes), as an activator of benzoate and methylbenzoate degradation via *meta*-ring fission (*xyl* genes) and as a repressor of *p*-hydroxybenzoate degradation (*pca* genes) in response to benzoate (Cowles *et al.*, 2000). As already shown in *P. putida* KT2440 by studying the expression of the *xyl* genes from plasmid pWW0, the pathway-specific regulation will be subordinated to a more general control that adjusts the particular transcriptional output to the physiological status of the cell (Cases and de Lorenzo, 2001).

Although *P. putida* KT2440 turns out to be a very useful model system for studying biochemical, genetic, evolutionary and ecological aspects of the catabolism of aromatic compounds, our current knowledge about the overall catabolic versatility of *P. putida* towards aromatic compounds may still be far from complete. Thus, analysis of the whole *P. putida* KT2440 genome has shown the presence of several genes, e.g. the *pcm* and *nic* genes, that are likely to be involved in the degradation and/or transformation of aromatic compounds. Furthermore, the pathways for degradation of phenylacetate, quinate and aromatic amines are not yet well understood, and further work needs to be done to identify the genes and/or enzymatic steps involved in such catabolic routes. A deeper understanding of the complete set of aromatic catabolic abilities of *P. putida* KT2440 will pave the way for the rational design of more efficient and broad-range biocatalysts for many biotechnological applications.

Experimental procedures

Bacterial strains and growth conditions

The strain used in this work was *P. putida* KT2440 (Franklin *et al.*, 1981). Bacteria were cultivated in M63 minimal medium (Miller, 1972) supplemented with MgSO₄ and trace metals, with 5 mM of different carbon sources (see below), by shaking at 30°C. Cell growth in liquid media was monitored by optical density readings at 600 nm (OD₆₀₀). Compounds that did not support growth of *P. putida* KT2440 were also checked at 1 and 2 mM final concentration to rule out toxicity problems. When necessary, growth media was solidified by the addition of agar to a final concentration of 1.5% (w/v).

Carbon sources

The stock solutions of the carbon sources used were filter sterilized and added to the sterile growth medium aseptically.

- Elsemore, D.A., and Ornston, L.N. (1995) Unusual ancestry of dehydratases associated with quinate catabolism in *Acinetobacter calcoaceticus*. *J Bacteriol* **177**: 5971–5978.
- Eulberg, D., Golovleva, L.A., and Schlömann, M. (1997) Characterization of catechol catabolic genes from *Rhodococcus erythropolis* 1CP. *J Bacteriol* **179**: 370–381.
- Eulberg, D., Lakner, S., Golovleva, L.A., and Schlömann, M. (1998) Characterization of a protocatechuate catabolic gene cluster from *Rhodococcus opacus* 1CP: evidence for a merged enzyme with 4-carboxymuconolactone-decarboxylating and 3-oxoadipate enol-lactone-hydrolyzing activity. *J Bacteriol* **180**: 1072–1081.
- Fernández-Cañón, J., and Peñalva, M.A. (1998) Characterization of a fungal maleylacetoacetate isomerase gene and identification of its human homologue. *J Biol Chem* **273**: 329–337.
- Ferrández, A., Miñambres, B., García, B., Olivera, E.R., Luengo, J.M., García, J.L., and Díaz, E. (1998) Catabolism of phenylacetic acid in *Escherichia coli*. Characterization of a new aerobic hybrid pathway. *J Biol Chem* **273**: 25974–25986.
- Fetzner, S. (1998) Bacterial degradation of pyridine, indole, quinoline, and their derivatives under different redox conditions. *Appl Microbiol Biotechnol* **49**: 237–250.
- Franklin, F.C.H., Bagdasarian, M., Bagdasarian, M.M., and Timmis, K.N. (1981) Molecular and functional analysis of the TOL plasmid pWWO from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring meta cleavage pathway. *Proc Natl Acad Sci USA* **78**: 7458–7462.
- Fukimori, F., Hirayama, H., Takami, H., Inoue, A., and Horiuchi, K. (1998) Isolation and transposon mutagenesis of a *Pseudomonas putida* KT2442 toluene-resistant variant: involvement of an efflux system in solvent resistance. *Extremophiles* **2**: 395–400.
- Galibert, F., Finan, T.M., Long, S.R., Puhler, A., Abola, P., Ampe, F., et al. (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**: 668–672.
- Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., et al. (2001) Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* **294**: 2323–2328.
- Gu, W., Song, J., Bonner, C.A., Xie, G., and Jensen, R.A. (1998) PhhC is an essential aminotransferase for aromatic amino acid catabolism in *Pseudomonas aeruginosa*. *Microbiology* **144**: 3127–3134.
- Hacisalihoglu, A., Jongejan, J.A., and Duine, J.A. (1997) Distribution of amine oxidases and amine dehydrogenases in bacteria grown on primary amines and characterization of the amine oxidase from *Klebsiella oxytoca*. *Microbiology* **143**: 505–512.
- Harayama, S., and Timmis, K.N. (1989) Catabolism of aromatic hydrocarbons by *Pseudomonas*. In *Genetics of Bacterial Diversity*. Hopwood, A., and Chater, K.F. (eds). London: Academic Press, pp. 151–174.
- Harwood, C.S., and Parales, R.E. (1996) The β -ketoadipate pathway and the biology of self-identity. *Annu Rev Microbiol* **50**: 553–590.
- Harwood, C.S., Nichols, N.N., Kim, M.-K., Ditty, J.L., and Parales, R.E. (1994) Identification of the *pcaRKf* gene cluster from *Pseudomonas putida*: involvement in chemotaxis, biodegradation, and transport of 4-hydroxybenzoate. *J Bacteriol* **176**: 6479–6488.
- Hawkins, A.R., Giles, N.H., and Kinghorn, J.R. (1982) Genetical and biochemical aspects of quinate breakdown in the filamentous fungus *Aspergillus nidulans*. *Biochem Genet* **20**: 271–286.
- Hawkins, A.R., Lamb, H.K., Smith, M., Keyte, J.W., and Roberts, C.F. (1988) Molecular organisation of the quinic acid utilization (*QUT*) gene cluster in *Aspergillus nidulans*. *Mol Gen Genet* **214**: 224–231.
- Houghton, J.E., Brown, T.M., Appel, A.J., Hughes, E.J., and Ornston, L.N. (1995) Discontinuities in the evolution of *Pseudomonas putida* cat genes. *J Bacteriol* **177**: 401–412.
- Inglede, W.M., and Tai, C.C. (1972) Quinate metabolism in *Pseudomonas aeruginosa*. *Can J Microbiol* **18**: 1817–1824.
- Iwaki, M., Yagi, T., Horiike, K., Saeki, Y., Ushijima, T., and Nozaki, M. (1983) Crystallization and properties of aromatic amine dehydrogenase from *Pseudomonas* sp. *Arch Biochem Biophys* **220**: 253–262.
- Jeffrey, W.H., Cuskey, S.M., Chapman, P.J., Resnick, S., and Olsen, R.H. (1992) Characterization of *Pseudomonas putida* mutants unable to catabolize benzoate: cloning and characterization of *Pseudomonas* genes involved in benzoate catabolism and isolation of a chromosomal DNA fragment able to substitute for *xylS* in activation of the TOL lower-pathway promoter. *J Bacteriol* **174**: 4986–4996.
- Kasai, Y., Inoue, J., and Harayama, S. (2001) The TOL plasmid pWWO *xylN* gene product from *Pseudomonas putida* is involved in *m*-xylene uptake. *J Bacteriol* **183**: 6662–6666.
- Kukor, J.J., Olsen, R.H., and Ballou, D.P. (1988) Cloning and expression of the *catA* and *catBC* gene clusters from *Pseudomonas aeruginosa* PAO. *J Bacteriol* **170**: 4458–4465.
- Lathe, W.C., III, Snel, B., and Bork, P. (2000) Gene context conservation of a higher order than operons. *Trends Biochem Sci* **25**: 474–479.
- Luengo, J.M., García, J.L., and Olivera, E.R. (2001) The phenylacetyl-CoA catabolon: a complex catabolic unit with broad biotechnological applications. *Mol Microbiol* **39**: 1434–1442.
- Milcamps, A., and de Bruijn, F.J. (1999) Identification of a novel nutrient-deprivation-induced *Sinorhizobium meliloti* gene (*hmgA*) involved in the degradation of tyrosine. *Microbiology* **145**: 935–947.
- Miller, J. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Mitra, A., Kitamura, Y., Gasson, M.J., Narbad, A., Parr, A.J., Payne, J., et al. (1999) 4-Hydroxycinnamoyl-CoA hydratase/lyase (HCHL) – an enzyme of phenylpropanoid chain cleavage from *Pseudomonas*. *Arch Biochem Biophys* **365**: 10–16.
- Mohamed, M.E., Ismail, W., Heider, J., and Fuchs, G. (2002) Aerobic metabolism of phenylacetic acids in *Azoarcus evansii*. *Arch Microbiol* **178**: 180–192.
- Morawski, B., Segura, A., and Ornston, L.N. (2000) Repression of *Acinetobacter* vanillate demethylase synthesis by VanR, a member of the GntR family of transcriptional regulators. *FEMS Microbiol Lett* **187**: 65–68.

- Nakai, C., Horiike, K., Kuramitsu, S., Kagamiyama, H., and Nozaki, M. (1990) Three isoenzymes of catechol 1,2-dioxygenase (pyrocatechase), $\alpha\alpha$, $\alpha\beta$, and $\beta\beta$, from *Pseudomonas arvilla* C-1. *J Biol Chem* **265**: 660–665.
- Nakai, C., Uyeyama, H., Kagamiyama, H., Nakazawa, T., Inouye, S., Kishi, F., *et al.* (1995) Cloning, DNA sequencing, and amino acid sequencing of catechol 1,2-dioxygenase (pyrocatechase) from *Pseudomonas putida* mt-2 and *Pseudomonas arvilla* C-1. *Arch Biochem Biophys* **321**: 353–362.
- Nichols, N.N., and Harwood, C.S. (1997) PcaK, a high-affinity permease for the aromatic compounds 4-hydroxybenzoate and protocatechuate from *Pseudomonas putida*. *J Bacteriol* **179**: 5056–5061.
- Nierman, W.C., Feldblyum, T.V., Laub, M.T., Paulsen, I.T., Nelson, K.E., Eisen, J.A., *et al.* (2001) Complete genome sequence of *Caulobacter crescentus*. *Proc Natl Acad Sci USA* **98**: 4136–4141.
- Nozaki, M., Kagamiyama, H., and Hayaishi, O. (1963) Crystallization and some properties of metapyrocatechase. *Biochem Biophys Res Commun* **11**: 65–70.
- Olivera, E.R., Carnicero, D., García, B., Miñambres, B., Moreno, M.A., Cañedo, L., *et al.* (2001) Two different pathways are involved in the β -oxidation of *n*-alkanoic and *n*-phenylalkanoic acids in *Pseudomonas putida* U: genetic studies and biotechnological applications. *Mol Microbiol* **39**: 863–874.
- O'Sullivan, D.J., and O'Gara, F. (1992) Traits of fluorescent *Pseudomonas* spp. involved in suppression of plant root pathogens. *Microbiol Rev* **56**: 662–676.
- Overhage, J., Priefert, H., and Steinbüchel, A. (1999) Biochemical and genetic analyses of ferulic acid catabolism in *Pseudomonas* sp. strain HR199. *Appl Environ Microbiol* **65**: 4837–4847.
- Parales, R.E., and Harwood, C.S. (2002) Bacterial chemotaxis to pollutants and plant-derived aromatic molecules. *Curr Opin Microbiol* **5**: 266–273.
- Parke, D. (1996) Characterization of PcaQ, a LysR-type transcriptional activator required for catabolism of phenolic compounds, from *Agrobacterium tumefaciens*. *J Bacteriol* **178**: 266–272.
- Parke, D., D'Argenio, D.A., and Ornston, L.N. (2000) Bacteria are not what they eat: that is why they are so diverse. *J Bacteriol* **182**: 257–263.
- Parke, D., García, M.A., and Ornston, L.N. (2001) Cloning and genetic characterization of *dca* genes required for β -oxidation of straight-chain dicarboxylic acids in *Acinetobacter* sp. strain ADP1. *Appl Environ Microbiol* **67**: 4817–4827.
- Priefert, H., Rabenhorst, J., and Steinbüchel, A. (1997) Molecular characterization of genes of *Pseudomonas* sp. strain HR199 involved in bioconversion of vanillin to protocatechuate. *J Bacteriol* **179**: 2595–2607.
- Priefert, H., Rabenhorst, J., and Steinbüchel, A. (2001) Biotechnological production of vanillin. *Appl Microbiol Biotechnol* **56**: 296–314.
- Quinn, J.A., McKay, D.B., and Entsch, B. (2001) Analysis of the *pobA* and *pobR* genes controlling expression of *p*-hydroxybenzoate hydroxylase in *Azotobacter chroococcum*. *Gene* **264**: 77–85.
- Ramos, J.L., Díaz, E., Dowling, D., de Lorenzo, V., Molin, S., O'Gara, F., *et al.* (1994) The behaviour of bacteria designed for biodegradation. *Biotechnology* **12**: 1349–1356.
- Ramos, J.L., Duque, E., Godoy, P., and Segura, A. (1998) Efflux pumps involved in toluene tolerance in *Pseudomonas putida* DOT-T1E. *J Bacteriol* **180**: 3323–3329.
- Rojo, F., Pieper, D.H., Engesser, K.-H., Knackmuss, H.-J., and Timmis, K.N. (1987) Assemblage of *ortho* cleavage route for simultaneous degradation of chloro- and methylaromatics. *Science* **238**: 1395–1398.
- Saier, M.H., Jr (1998) Molecular phylogeny as a basis for the classification of transport proteins from bacteria, archaea, and eukarya. *Adv Microb Physiol* **40**: 81–136.
- Segura, A., Bünz, P.V., D'Argenio, D.A., and Ornston, L.N. (1999) Genetic analysis of a chromosomal region containing *vanA* and *vanB*, genes required for conversion of either ferulate or vanillate to protocatechuate in *Acinetobacter*. *J Bacteriol* **181**: 3494–3504.
- Serre, L., Sailland, A., Sy, D., Boudec, P., Rolland, A., Pebay-Peyroula, E., and Cohen-Addad, C. (1999) Crystal structure of *Pseudomonas fluorescens* 4-hydroxyphenylpyruvate dioxygenase: an enzyme involved in the tyrosine degradation pathway. *Structure* **7**: 977–988.
- Song, J., and Jensen, R.A. (1996) PhhR, a divergently transcribed activator of the phenylalanine hydroxylase gene cluster of *Pseudomonas aeruginosa*. *Mol Microbiol* **22**: 497–507.
- Song, J., Xia, T., and Jensen, R.A. (1999) PhhB, a *Pseudomonas aeruginosa* homolog of mammalian pterin 4a-carbinolamine dehydratase/DcoH, does not regulate expression of phenylalanine hydroxylase at the transcriptional level. *J Bacteriol* **181**: 2789–2796.
- Stanier, R.Y., Palleroni, N.J., and Doudoroff, M. (1966) The aerobic *Pseudomonads*: a taxonomic study. *J Gen Microbiol* **43**: 159–271.
- Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., *et al.* (2000) Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature* **406**: 959–964.
- Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, Y., Masui, N., *et al.* (2000) Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. *Nucleic Acids Res* **28**: 4317–4331.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.
- Venturi, V., Zennaro, F., Degraffi, G., Okeke, B.C., and Bruschi, C.V. (1998) Genetics of ferulic acid bioconversion to protocatechuic acid in plant-growth-promoting *Pseudomonas putida* WCS358. *Microbiology* **144**: 965–973.
- Wilbur, W.J., and Lipman, D.J. (1983) Rapid similarity searches of nucleic acid and protein data banks. *Proc Natl Acad Sci USA* **80**: 726–730.
- Yeo, C.C., Tham, J.M., Yap, M.W.-C., and Poh, C.L. (1997) Group II intron from *Pseudomonas alcaligenes* NCIB 9867 (P25X): entrapment in plasmid RP4 and sequence analysis. *Microbiology* **143**: 2833–2840.