environmental microbiology

Environmental Microbiology (2016) **18**(12), 4888–4906 doi:10.1111/1462-2920.13413

# Comparative genomic analysis of isoproturon-mineralizing sphingomonads reveals the isoproturon catabolic mechanism

**Xin Yan,[1] Tao Gu,[1] Zhongquan Yi,[1] Junwei Huang,[1] Xiaowei Liu,[1] Ji Zhang,[1] Xihui Xu,[1] Zhihong Xin,[2] Qing Hong,[1] Jian He,[1] Jim C. Spain,[3] Shunpeng Li[1] and Jiandong Jiang[1]***

[1]*Department of Microbiology, Key Lab of Microbiological Engineering of Agricultural Environment, Ministry of Agriculture, College of Life Sciences, Nanjing Agricultural University, Nanjing 210095, People's Republic of China.*

[2]*College of Food Science and Technology, Nanjing Agricultural University, Nanjing 210095, People's Republic of China.*

[3]*School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0512, USA.*

## Summary

**The worldwide use of the phenylurea herbicide, isoproturon (IPU), has resulted in considerable concern about its environmental fate. Although many microbial metabolites of IPU are known and IPU-mineralizing bacteria have been isolated, the molecular mechanism of IPU catabolism has not been elucidated yet. In this study, complete genes that encode the conserved IPU catabolic pathway were revealed, based on comparative analysis of the genomes of three IPU-mineralizing sphingomonads and subsequent experimental validation. The complete genes included a novel hydrolase gene *ddhA*, which is responsible for the cleavage of the urea side chain of the IPU demethylated products; a distinct aniline dioxygenase gene cluster *adoQTA1A2BR*, which has a broad substrate range; and an inducible catechol *meta*-cleavage pathway gene cluster *adoXEGKLIJC*. Furthermore, the initial mono-*N*-demethylation genes *pdmAB* were further confirmed to be involved in the successive *N*-demethylation of the IPU mono-*N*-demethylated product. These IPU-catabolic genes were organized into four transcription units and distributed on three plasmids. They were flanked by multiple mobile genetic elements and highly conserved among IPU-mineralizing sphingomonads. The elucidation of the molecular mechanism of IPU catabolism will enhance our understanding of the microbial mineralization of IPU and provide insights into the evolutionary scenario of the conserved IPU-catabolic pathway.**

Received 11 January, 2016; revised 15 March, 2016; accepted 17 March, 2016. *For correspondence. E-mail jiang_jjd@njau.edu.cn; Tel./Fax: (+86) 025 84399726.

## Introduction

The introduction of xenobiotic chemicals, such as pesticides, into the environment drives microorganisms to evolve new catabolic pathways either to exploit new carbon and energy sources or to detoxify these toxic compounds. Mineralization of pesticides is considerably more difficult than detoxification because microorganisms must recruit multiple enzymes to convert pesticides into the specific intermediates which are involved in a central metabolic pathway (Copley, 2009). However, it is likely that, given a sufficient time of exposure, microbial cells can evolve new pathways to mineralize most newly introduced pesticides (Lal *et al.*, 2006; Shapir *et al.*, 2007).

The phenylurea herbicide (PUH) family, first discovered and marketed in the mid-20th century, is one of the most important and extensively used herbicides in different regions of the world, including China, Europe and the United States (Fabbri *et al.*, 2015). Ecotoxicological data have suggested that PUHs and their intermediates including aniline derivatives are harmful to animals, plants, aquatic invertebrates, algae and microbes, in addition to humans (Oturan *et al.*, 2008; Mosleh, 2009; Da Rocha *et al.*, 2013). Isoproturon (IPU), first released in 1972, is the most extensively used PUH in conventional agriculture particularly in Europe and China. Due to its widespread usage, IPU and its metabolites including 3-(4-isopropylphenyl)−1-methylurea (MDIPU), *N*-(4-isopropylphenyl)urea (DDIPU) and 4-isopropylaniline (4-IA) have been commonly detected as contaminants in rivers, streams, lakes, marine water and groundwater around the world. Therefore, much attention has been paid to the environmental fate of IPU for more

than 40 years (Sørensen *et al*., 2003; Hussain *et al*., 2015).

Microbial degradation is considered to be the primary mechanism for the dissipation of IPU from the environment (Mudd *et al*., 1983; Gaillardon and Sabar, 1994; Cox *et al*., 1996). Previous studies have indicated that an IPU-catabolic pathway has evolved in bacteria during several decades of widespread application of PUHs. In 1998, some pure cultures of bacteria including a particularly active strain tentatively identified as *Pseudomonas fluorescens* that catabolized IPU as the sole carbon and nitrogen source, were obtained (Roberts *et al*., 1998). In 2001, *Sphingomonas* sp. strain SRS2 was isolated from an IPU-treated agricultural soil in the United Kingdom based on its ability to mineralize IPU (Sørensen *et al*., 2001). Subsequently, about ten IPU-mineralizing bacteria were isolated and characterized from different regions of Europe and Asia (Fig. 1A. *Pseudomonas fluorescens* is not included due to the unavailability of its 16S rRNA gene sequence). Among these strains, eight are sphingomonads (Bending *et al*., 2003; Sun *et al*., 2006; 2009; Hussain *et al*., 2011) and other two are *Methylopila* sp. strain TES (El Sebai *et al*., 2004) and *Pseudomonas aeruginosa* strain JS-11 (Dwivedi *et al*., 2011). Because similar intermediates have been detected during the IPU-mineralizing process in sphingomonads, a common catabolic pathway has been proposed (Sørensen *et al*., 2003; Hussain *et al*., 2015). In this pathway, two successive *N*-demethylations are followed by the cleavage of the urea side chain, generating 4-IA; 4-IA is transformed into 4-isopropylcatechol (4-IPC), which is the substrate for ring-cleavage (Fig. 1B).

Though the metabolic pathway of IPU has been proposed, the molecular mechanism of IPU mineralization is far from clear. The *pdmAB* genes encode the Rieske non-heme iron oxygenase (RO), which were responsible for the initial *N*-demethylation step (IPU to MDIPU), were the only identified genes for IPU catabolism and were found to be highly conserved (99% similarity) among the sphingomonads YBL1, YBL2, YBL3, Y57 and SRS2 (Gu *et al*., 2013). Considering the facts that the *linABCDEF* genes in hexachlorocyclohexane (HCH)-mineralizing sphingomonads are highly conserved and the first initial demethylation genes *pdmAB* in IPU-mineralizing sphingomonads are also conserved (Lal *et al*., 2006; Gu *et al*., 2013), we hypothesized that the other genes involved in the IPU catabolism might also be conserved (likely with $>$ 95% similarity).

In this study, the genomes of the three IPU-mineralizing sphingomonads *Sphingobium* sp. strain YBL2 (isolated in Changzhou, China; Sun *et al*., 2009), *Sphingomonas* sp. strain Y57 (isolated in Suzhou, China; Sun *et al*., 2006) and *Sphingomonas* sp. strain SRS2 (isolated from Wellesbourne, UK; Sørensen *et al*., 2001) were sequenced and compared. Strain YBL2 is capable of utilizing IPU as the sole carbon and nitrogen sources for growth and degrading IPU and its intermediates (Fig. 2). The highly conserved ORFs ($>$ 95% similarity) among the three strains were chosen and subjected to bioinformatics analysis and experimental validation. Genes involved in the IPU-catabolic pathway were identified through recombinant expression, gene knockout and complementation studies. Furthermore, the expression regulation, transcriptional response, substrate range and organization structure of IPU-catabolic genes were also investigated. This work will elucidate the molecular mechanism of IPU mineralization and enhance our understanding of the evolutionary mechanism of the IPU-mineralizing pathway.

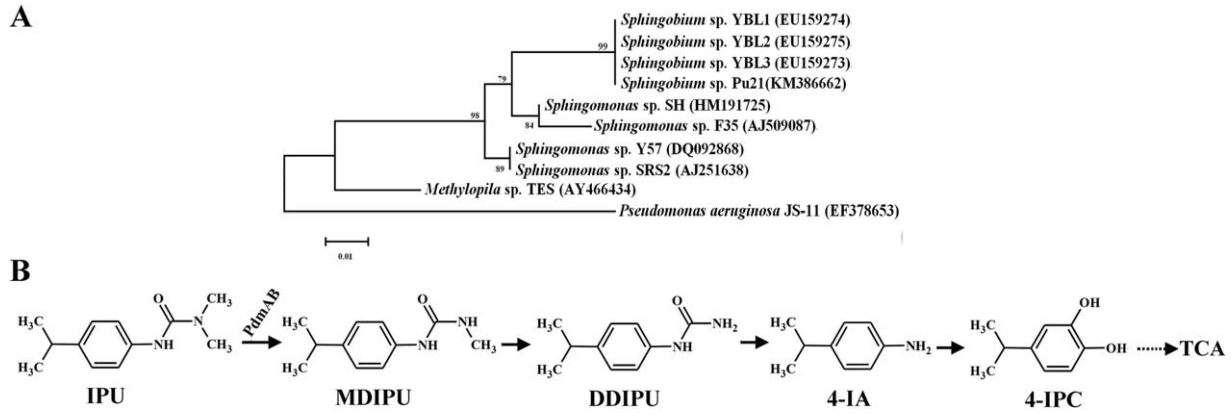


**Fig. 1.** IPU-mineralizing bacteria and proposed IPU-catabolic pathway in sphingomonads.
(A) Phylogenetic analysis of IPU-mineralizing bacteria. *Pseudomonas fluorescens* is not included due to the unavailability of its 16S rRNA gene sequence (Roberts *et al*., 1998).
(B) Proposed IPU-catabolic pathway based on the intermediates detected in the IPU degradation process. PdmAB has been previously demonstrated to be responsible for the first *N*-demethylation step (Gu *et al*., 2013).

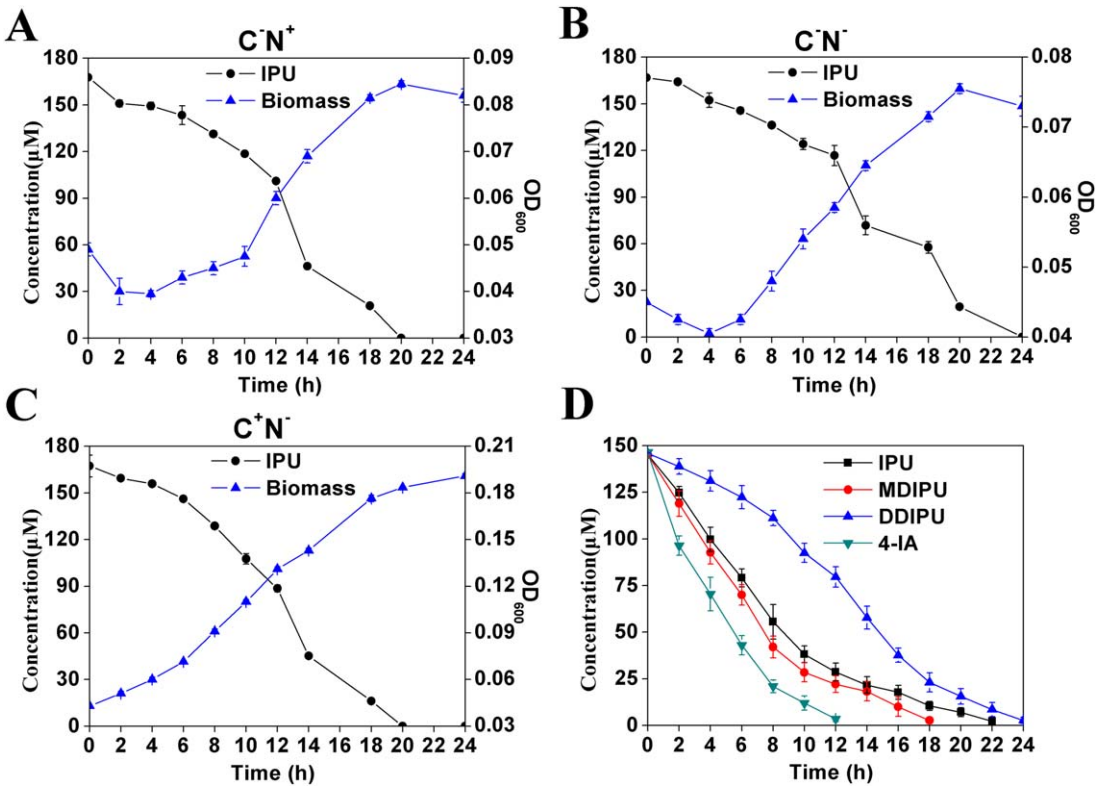


**Fig. 2.** Degradation of IPU and its intermediates by strain YBL2. Growth of strain YBL2 using IPU as the sole carbon source (A), as the sole carbon and nitrogen sources (B), or as the sole nitrogen source (C). Degradation kinetics of IPU and its metabolites in MSM by the cell suspension of strain YBL2 (D). Detailed method was specified in the experimental procedures section. All experiments were carried out in triplicate.

## Results

### *Genome features of three IPU-mineralizing sphingomonads and prediction of IPU-catabolic genes based on comparative genomic analysis*

The complete genome sequence of strain YBL2 revealed seven replicons, consisting of one circular chromosome (4 766 421 bp, 64.78% G + C, 4480 ORFs) and six circular plasmids: pYBL2-1 (27 036 bp, 60.55% G + C, 34 ORFs), pYBL2-2 (227 503 bp, 62.83% G + C, 244 ORFs), pYBL2-3 (322 226 bp, 61.14% G + C, 399 ORFs), pYBL2-4 (44 342 bp, 62.24% G + C, 58 ORFs), pYBL2-5 (20 552 bp, 60.72% G + C, 23 ORFs) and pYBL2-6 (19 339 bp, 60.85% G + C, 17 ORFs). The whole genome has an average G + C content of 64.42%, and it possesses a total of three copies of *rrn* operons located on the chromosome. A total of 5768 protein-coding genes were predicted. The draft genome sequence of strain SRS2 contains 168 contigs, constituting a total size of 4.63 Mb, of which the G + C content is 63.9%. The draft genome sequence of strain Y57 is 5.61 Mb in length (64.42% G + C content) and consists of 195 contigs. The general features of these three genomes are summarized in Supporting Information Table S1.

Strains YBL2 (5.43 Mb) and Y57 (5.61 Mb) had similar genome sizes, which were larger than that of strain SRS2 (4.63 Mb). However, a pairwise sequence comparison of the genome sequences revealed high degrees of macrosynteny (i.e., chromosomes with similar gene contents, orders and orientations) (Hane *et al.*, 2011) between strains Y57 and SRS2, with an average amino acid sequence similarity of more than 80% (Supporting Information Fig. S1). The macrosynteny results were in agreement with the phylogenetic relationships of strains Y57 and SRS2 (Fig. 1A). In contrast, only mesosynteny (i.e., chromosomes with similar gene contents but with different orders and orientations of genes) was observed between strains YBL2 and Y57 or between strains YBL2 and SRS2, and most of the similarities of the orthologous proteins were 60%–80%.

A total of 84 ORFs sharing more than 95% similarities among the genomes of the three sphingomonads were found (Supporting Information Table S2). Most of these conserved ORFs were located on the six plasmids of strain YBL2, with only three genes (encoding CoA reductase, endonuclease and flavin monoamine oxidase) on the chromosome. The location of the 84 ORFs on the genomes of strains Y57 and SRS2 is unclear due to the lack of

complete genome sequences. Given the pivotal roles of plasmids in xenobiotics degradation, these 84 highly conserved ORFs are likely to contain genes involved in IPU catabolism.

Based on the IPU catabolic pathway (Fig. 1B), the candidates responsible for each reaction step were predicted from these 84 ORFs. (i) For the initial *N*-demethylation of IPU to MDIPU, *pdmAB* genes were identified based on our previous study (Gu *et al*., 2013). (ii) Because the demethylation of MDIPU by PdmAB was not identified in our previous study (Gu *et al*., 2013), the successive *N*-demethylation of MDIPU to DDIPU was originally believed to be catalyzed by another demethylase. According to the current understanding, the *O*/*N*-demethylation reaction can be catalyzed by cytochrome P450 monooxygenase (CYP) (Wang *et al*., 2015), RO (Summers *et al*., 2012) or tetrahydrofolate-dependent demethylase (Abe *et al*., 2005). However, no appropriate candidates were found among the 84 conserved ORFs. (iii) The enzyme catalyzing the cleavage of the urea side chain of DDIPU to 4-IA is proposed to be a hydrolase. Although no likely candidate was found among these conserved ORFs, two ORFs on the plasmids of strain YBL2 warranted further investigation. One candidate is ORF228 on pYBL2-2. Its deduced protein sequence shows 56% and 41% similarity to known β-lactamase and esterase, respectively. Another candidate is ORF4 on pYBL2-5. Its deduced protein sequence exhibits 21% similarity (coverage of 51%) to the carbaryl hydrolase CehA from *Rhizobium* sp. strain AC100 (Hashimoto *et al*., 2002). CehA was able to cleave the ester bond of carbaryl, a reaction similar to the cleavage of the urea side chain of DDIPU. (iv) For the degradation of 4-IA, all of the currently known gene clusters involved in the transformation of aniline/its derivatives into TCA-cycle intermediates consist of the multi-component aniline dioxygenase gene cluster and the catechol-catabolic gene cluster. ORF228-243 on plasmid pYBL2-2 was found to be a typical aniline catabolic gene cluster. Specifically, ORF238-243 encodes the aniline dioxygenase, and ORF228-235 encodes the catechol *meta*-cleavage pathway. Then, all of the candidate genes predicted for IPU catabolism (Table 1) were subjected to experimental validation.

### *ORF4 on pYBL2-5 encodes a novel hydrolase (DdhA) that converts both MDIPU and DDIPU into 4-IA*

The candidates for urea side chain cleavage were validated first because of the uncertainty regarding the second *N*-demethylase. ORF4 and ORF228 were separately expressed in both *Escherichia coli* BL21(DE3) and *Sphingomonas wittichii* RW1. Both recombinant strains containing ORF4 converted DDIPU into 4-IA. In contrast, neither strain expressing ORF228 showed activity toward DDIPU. Therefore, ORF4 was designated as the urea side chain cleavage gene *ddhA*. Although recombinant strain BL21(DE3)(pET-ddhA) exhibited obvious activity against DDIPU when induced with isopropyl β-D-thiogalactopyranoside (IPTG), SDS-PAGE results showed that DdhA was expressed at an extremely low level, even using different hosts or growth conditions. Therefore, cell-free extract of strain BL21(DE3)(pET-ddhA) after IPTG induction was used to test the activity of DdhA, and strain BL21(DE3) containing the empty vector was used as the control. The cell-free extract of strain BL21(DE3)(pET-ddhA) was able to transform both DDIPU and MDIPU into 4-IA (Fig. 3). For MDIPU and DDIPU, the specific activities of DdhA were 11.7 U/mg protein and 4.0 U/mg protein, respectively. In addition, the extract showed no activity toward IPU.

To further identify the *in vivo* role of *ddhA* in strain YBL2, *ddhA* was inactivated through a single-crossover insertion, generating the mutant YBL2-ddhA. The mutant failed to convert either MDIPU or DDIPU into 4-IA, while complementation by the *ddhA*-containing plasmid pBBRddhA restored its ability, showing that *ddhA* is the only gene responsible for cleavage of the urea side chain in both MDIPU and DDIPU. Therefore, these results demonstrated that MDIPU could be directly transformed into 4-IA with higher efficiency than DDIPU by DdhA in strain YBL2, which was different from previous speculation that MDIPU should be transformed into DDIPU first and then the urea side chain was cleaved (Sørensen *et al*., 2003; Hussain *et al*., 2015).

### *p*dmAB *genes are also involved in the successive* N*-demethylation and aromatic ring hydroxylation of MDIPU*

Although the mutant YBL2-ddhA failed to cleave the urea side chains of MDIPU and DDIPU, it was able to convert MDIPU into DDIPU and a new compound at a low efficiency (Fig. 4). The new compound was presumed to be the aromatic-ring-hydroxylated MDIPU (called HO-MDIPU), which was identified through LC-MS/MS analysis (Supporting Information Fig. S2). Moreover, the *pdmAB/ddhA* double mutant YBL2-ABA showed no detectable activity toward IPU, MDIPU or DDIPU. Therefore, these data indicated that *pdmAB* genes are also involved in the transformation of MDIPU into DDIPU and HO-MDIPU. To further confirm this, plasmid pBBRPAB containing *pdmAB* genes was introduced into strain RW1. Same with the mutant YBL2-ddhA, strain RW1(pBBRPAB) could convert MDIPU into DDIPU and HO-MDIPU within 36 h. By contrast, the control strain RW1(pBBR1MCS-5) showed no detectable activity toward MDIPU. Taken together, these results suggested that PdmAB is responsible for the conversion of IPU into MDIPU, MDIPU into DDIPU as well as MDIPU into HO-MDIPU. Previous reports have shown that

**Table 1.** Candidate genes involved in IPU catabolism in strain YBL2.

| ORF no. (Location) | Protein length (aa[a]) (Accession no.) | Closest relative (Accession no.) | Identity (aa[a]) | Function |
|---|---|---|---|---|
| 4 (pYBL2-4) | 324 (AGS18337.1) | *Sphingobium* sp. strain YBL2 PdmA (AGS18337.1) | 100% | IPU to MDIPU, MDIPU to DDIPU, and MDIPU to HO-MDIPU |
| 5 (pYBL2-4) | 176 (AGS18336.1) | *Sphingobium* sp. strain YBL2 PdmB (AGS18336.1) | 100% | |
| 228 (pYBL2-2) | 412 (WP_044663302.1) | *Caulobacter* sp. strain K31 beta-lactamase (WP_012284656.1) | 58% | DDIPU to 4-IA |
| 4 (pYBL2-5) | 712 (WP_044663665.1) | *Pseudomonas* sp. strain C5pp hypothetical protein (WP_039615392.1) | 43% | |
| 238 (pYBL2-2) | 114 (WP_044663309.1) | uncultured bacterium GntR family-transcriptional regulator (BAH90337.1) | 62% | Aniline dioxygenase; 4-IA to 4-IPC |
| 239 (pYBL2-2) | 335 (WP_044663310.1) | *Pseudomonas putida* electron transfer protein (WP_032490114.1) | 60% | |
| 240 (pYBL2-2) | 213 (WP_044663335.1) | *Burkholderia* sp. strain K24 hypothetical protein, partial (WP_035520062.1) | 71% | |
| 241 (pYBL2-2) | 444 (WP_044663311.1) | *Delftia acidovorans* DcaA1 (WP_015060610.1) | 72% | |
| 242 (pYBL2-2) | 242 (WP_044663312.1) | *Burkholderia fungorum* hypothetical protein (WP_028196319.1) | 55% | |
| 243 (pYBL2-2) | 499 (WP_044663313.1) | *Comamonas testosterone* CaoQ (WP_014386258.1) | 66% | |
| 227 (pYBL2-2) | 251 (WP_044663301.1) | *Dechloromonas aromatic* transcriptional regulator (WP_011289515.1) | 61% | Catechol degradation; 4-IPC to TCA-cycle intermediates |
| 229 (pYBL2-2) | 508 (WP_044663334.1) | *Hydrocarboniphaga effuse* 2-hydroxymuconic semialdehyde dehydrogenase (WP_007184067.1) | 76% | |
| 230 (pYBL2-2) | 262 (WP_044663303.1) | *Azoarcus toluclasticus* 2-keto-4-pentenoate hydratase (WP_018992902.1) | 64% | |
| 231 (pYBL2-2) | 267 (WP_044663304.1) | *Polycyclovorans algicola* 4-oxalocrotonate decarboxylase (WP_029889205.1) | 54% | |
| 232 (pYBL2-2) | 77 (WP_044663305.1) | *Aquaspirillum serpens* 4-oxalocrotonate tautomerase (WP_022653256.1) | 54% | |
| 233 (pYBL2-2) | 305 (WP_044663306.1) | *Polycyclovorans algicola* acetaldehyde dehydrogenase (WP_029890418.1) | 76% | |
| 234 (pYBL2-2) | 346 (WP_044663307.1) | *Methyloversatilis universalis* 4-hyroxy-2-oxovalerate aldolase (WP_018230569.1) | 81% | |
| 235 (pYBL2-2) | 295 (WP_044663308.1) | *Methyloversatilis universalis* catechol 2,3-dioxygenase (WP_018230568.1) | 72% | |

**a.** Amino acid.

trace DDIPU was detected during the process of IPU degradation by different sphingomonads (Sørensen *et al.*, 2001; Sun *et al.*, 2006), which was in agreement with our finding of low transformation efficiency of MDIPU into DDIPU by PdmAB. Moreover, strain RW1(pBBRPAB) could not hydroxylate DDIPU or 4-IA (data not shown), indicating PdmAB could only hydroxylate MDIPU. Additionally, the cell-free extract of IPTG-induced strain BL21(DE3)(pET-ddhA) showed no detectable activity toward HO-MDIPU.

### *ORF238-243* (adoQTA1A2BR) *on pYBL2-2 encodes the aniline dioxygenase with a broad substrate range*

The downstream genes responsible for the catabolism of 4-IA and aniline derivatives were further identified. As shown in Fig. 5, the aniline catabolic genes are typically composed of two parts: the aniline dioxygenase gene cluster and the catechol *meta*/*ortho*-cleavage pathway gene cluster. Because conversion of aniline into catechol is the key step in the aniline catabolism, this work focused mainly

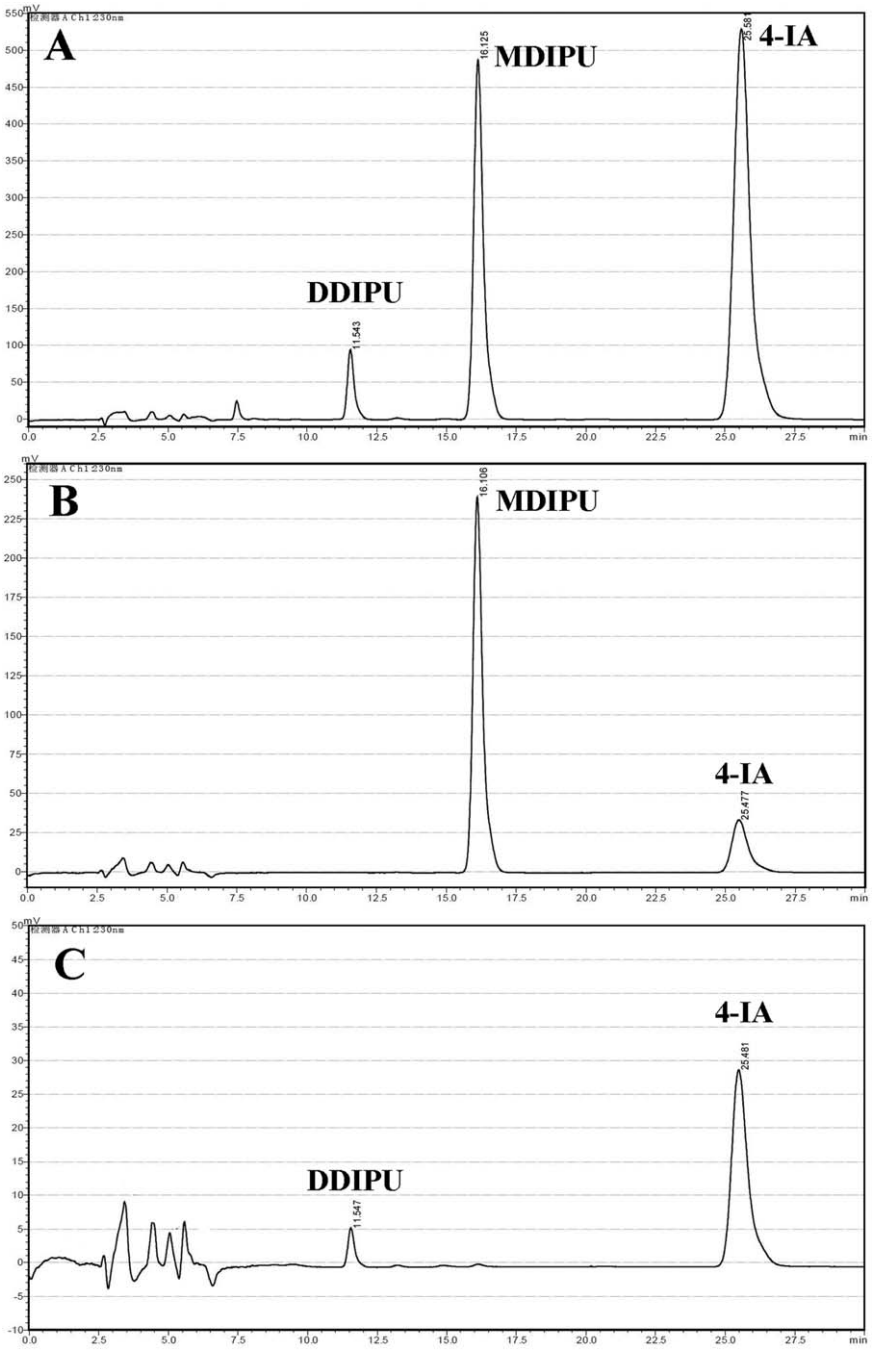


**Fig. 3.** HPLC analysis of the products of MDIPU and DDIPU treated with DdhA.
(A) Chromatogram of authentic DDIPU, MDIPU and 4-IA.
(B) Chromatogram of the products of MDIPU treated with the cell-free extract of IPTG-induced BL21(DE3)(pET-ddhA).
(C) Chromatogram of the products of DDIPU treated with the cell-free extract of IPTG-induced BL21(DE3)(pET-ddhA). The assay was carried out as specified in the experimental procedure section. The reaction times for MDIPU (B) and DDIPU (C) were 5 and 12 h, respectively.

on characterizing the aniline dioxygenase gene cluster *adoQTA1A2BR* from two aspects: sequence analysis and substrate range. Urata *et al*. (2004) suggested that the currently characterized aniline dioxygenase gene clusters from Gram-negative bacteria could be divided into two distinct groups, *tdn* (Fukumori and Saint, 1997) and *atd* (Fujii *et al*., 1997), by comparison with glutamine synthetase (GS)-like protein (TdnQ/AtdA1). Within each group, the GS-like proteins shared strong similarities (86.3%–99.2%). The GS-like proteins in the *tdn* group were only 60%–66% similar to those of the *atd* group. In addition, the regulators of both groups (TdnR/AtdR) were LysR-type transcriptional activators in the presence of aniline (Geng *et al*., 2009).

The six genes in *adoQTA1A2BR* cluster were expected to encode a GS-like protein (AdoQ, corresponding to TdnQ/AtdA1), a glutamine amidotransferase-like protein (AdoT, corresponding to TdnT/AtdA2), large and small subunits of an aromatic compound dioxygenase (AdoA1A2, corresponding to TdnA1A2/AtdA3A4), a reductase (AdoB, corresponding to TdnB/AtdA5) and a regulator

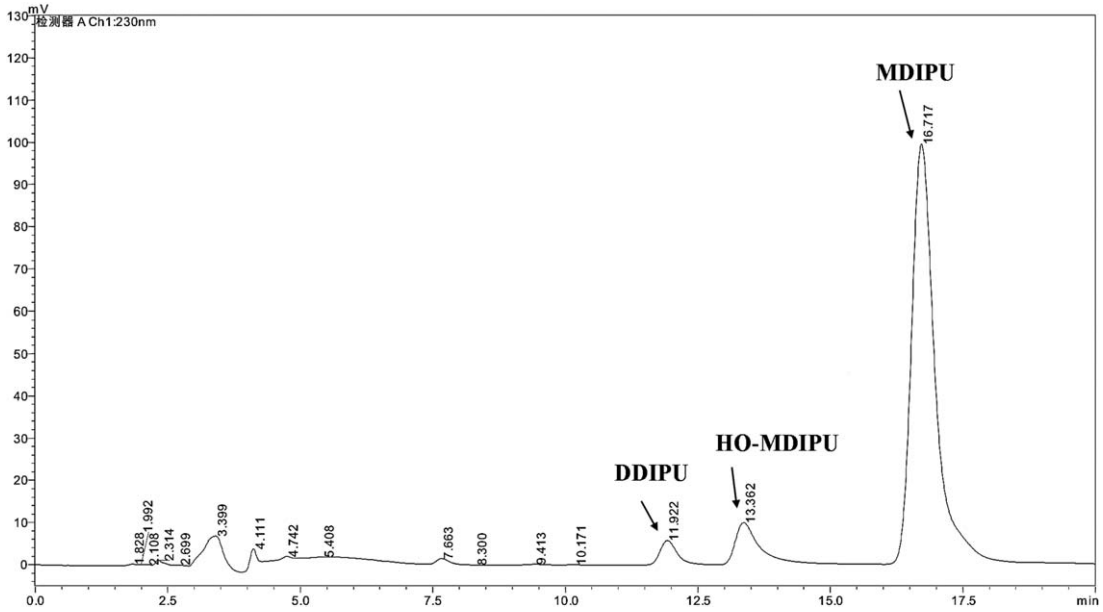


**Fig. 4.** HPLC analysis of the products of MDIPU treated with the *ddhA*-disrupted mutant YBL2-ddhA. The cell suspension assay was carried out as specified in the experimental procedure section and the reaction time was 24 h. The products were further identified as DDIPU and HO-MDIPU by LC-MS/MS analysis (shown in Supporting Information Fig. S2).

protein (AdoR, corresponding to TdnR/AtdR). Sequence analysis indicated that the *ado* cluster is distinct from the *tdn* and *atd* clusters. First, AdoQTA1A2B was 50%–70% similar to its counterparts in the Tdn group and 30%–57% similar to its counterparts in the Atd group. Second, AdoQ only showed 65% and 57% similarity to TdnQ and AtdQ, respectively, which was significantly less than the similarity for assigning *ado* cluster into the *tdn* or *atd* group (Urata *et al.*, 2004). Third, unlike LysR-type TdnR and AtdR, AdoR is a truncated GntR-type regulator (Fig. 5C). Finally, different from the organization of *tdn* and *atd* groups, *adoQTA1A2BR* and *adoXEGKLIJC* (the following catechol *meta*-cleavage pathway gene cluster) were transcribed in the opposite direction rather than organized into one transcription unit.

To test its function, the plasmid pBBRQR harbouring *adoQTA1A2BR* was introduced into *Pseudomonas putida* KT2440-Δ*catA* (in which both catechol dioxygenase genes were deleted) (Takeo *et al.*, 2013). Gas chromatography-mass spectrometry (GC-MS) analysis showed that 4-IA was transformed into 4-IPC (Supporting Information Fig. S3). In addition, all of the aniline derivatives shown in Fig. 6 could be transformed into the corresponding catechols by strain KT2440-Δ*catA*(pBBRQR) (Supporting Information Fig. S3). Interestingly, all of the aniline derivatives related to the commonly used PUHs could be transformed. Moreover, strain KT2440-Δ*catA*(pBBRQR) was able to efficiently transform halogen-substituted anilines, such as 3,4-dichloroaniline (Supporting Information Fig. S3). More than 90% of 3,4-dichloroaniline (123 μM) was transformed into the corresponding catechol within 12 h (data not shown). However, aniline derivatives containing halo-substitutions at position 2 or 5 could not be transformed. Whether strain YBL2 could use these aniline derivatives as the sole carbon and energy source for growth was further investigated. As shown in Supporting Information Fig. S4, strain YBL2 could grow on 4-IA, 3-methylaniline (3-MA) and 4-methylaniline (4-MA), indicating the produced catechol derivatives could be further transformed through aromatic ring-cleavage pathway in strain YBL2. Although halo-substituted anilines were degraded, no growth of strain YBL2 was observed, probably due to the toxicity of the halo-substituted anilines or their metabolites to cells.

To confirm whether ORF235-228 (*adoXEGKLIJC*) on pYBL2-2 are involved in the catabolism of 4-IPC in strain YBL2, the *adoC* gene encoding the catechol 2,3-dioxygenase was inactivated. Ultraviolet scanning and GC-MS analyses showed that inactivation of *adoC* resulted in the accumulation of 4-IPC during the degradation of IPU, indicating that *adoXEGKLIJC* are responsible for the further catabolism of 4-IPC (Supporting Information Fig. S5).

### *Transcription of IPU-catabolic genes in response to IPU*

The genes *pdmAB*, *ddhA*, *adoQTA1A2B* and *adoXEGKLIJC* were proved to be organized in four transcriptional units via RT-PCR (Supporting Information Fig. S6). Transcription of these catabolic genes in strain YBL2 with or without IPU was determined using both RT-PCR and RT-qPCR. As shown in Fig. 7A and C, the transcription of the

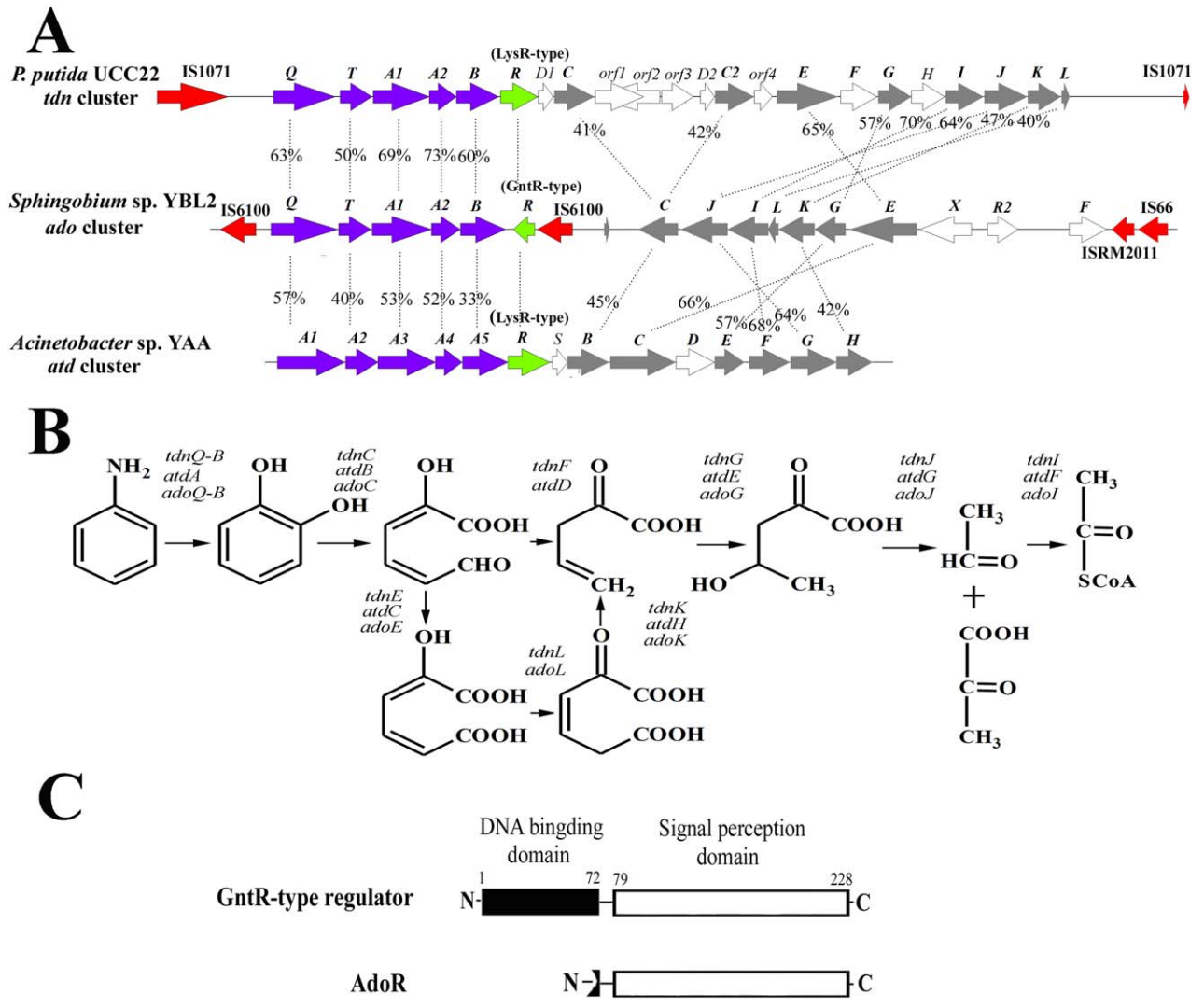


**Fig. 5.** Comparison of the three aniline degradation gene clusters *ado*, *tdn* and *atd*.
(A) Genetic organization of *ado*, *tdn* and *atd*. The genes encoding transposase, aniline dioxygenase, catechol cleavage pathway and a transcription regulator are shown as red, blue, grey and green arrows, respectively. The homologous genes are connected by dotted lines and the amino acid similarities are shown above the dotted lines.
(B) The proposed catabolic pathway of aniline. The genes involved in each step are shown.
(C) Schematic domain organization of GntR-type regulator and AdoR. The domain containing the HTH DNA binding motif is indicated by a black box and the domain bound to the chemical inducer is indicated by a white box.

genes *pdmAB*, *ddhA* and *adoQTA1A2B* was constitutive and not enhanced by IPU, which was in agreement with the degradation dynamics of IPU and its intermediates by strain YBL2 (Fig. 2D). No obvious lag was observed at the beginning of the degradation process of any of the tested substrates (IPU, MDIPU, DDIPU and 4-IA). In addition, compared with MDIPU, DDIPU was degraded with a lower efficiency, which could be attributed to the low catalytic efficiency of DdhA to DDIPU. The *adoR* in the *adoQTA1A2BR* cluster was presumed to encode a transcriptional repressor belonging to the GntR family. However, it is a truncated gene (losing the N-terminal DNA binding domain) and allows the constitutive expression of *adoQTA1A2B* (Fig. 7A and C), which has also been confirmed by the *lacZ-adoQ* promoter fusion assay (data not shown).

The *adoC* and *adoE* genes in the catechol *meta*-cleavage pathway showed low transcriptional levels in the absence of IPU, but their transcriptional levels increased by 5.5- to 6.5-fold in the presence of IPU (Fig. 7A and C). Because the *adoXEGKLIJC* gene cluster was organized in one transcription unit, the transcriptional levels of *adoC* and *adoE* as two representative genes indicated that the expression of the entire cluster was induced by IPU or its intermediates. Upstream of *adoXEGKLIJC*, there is a

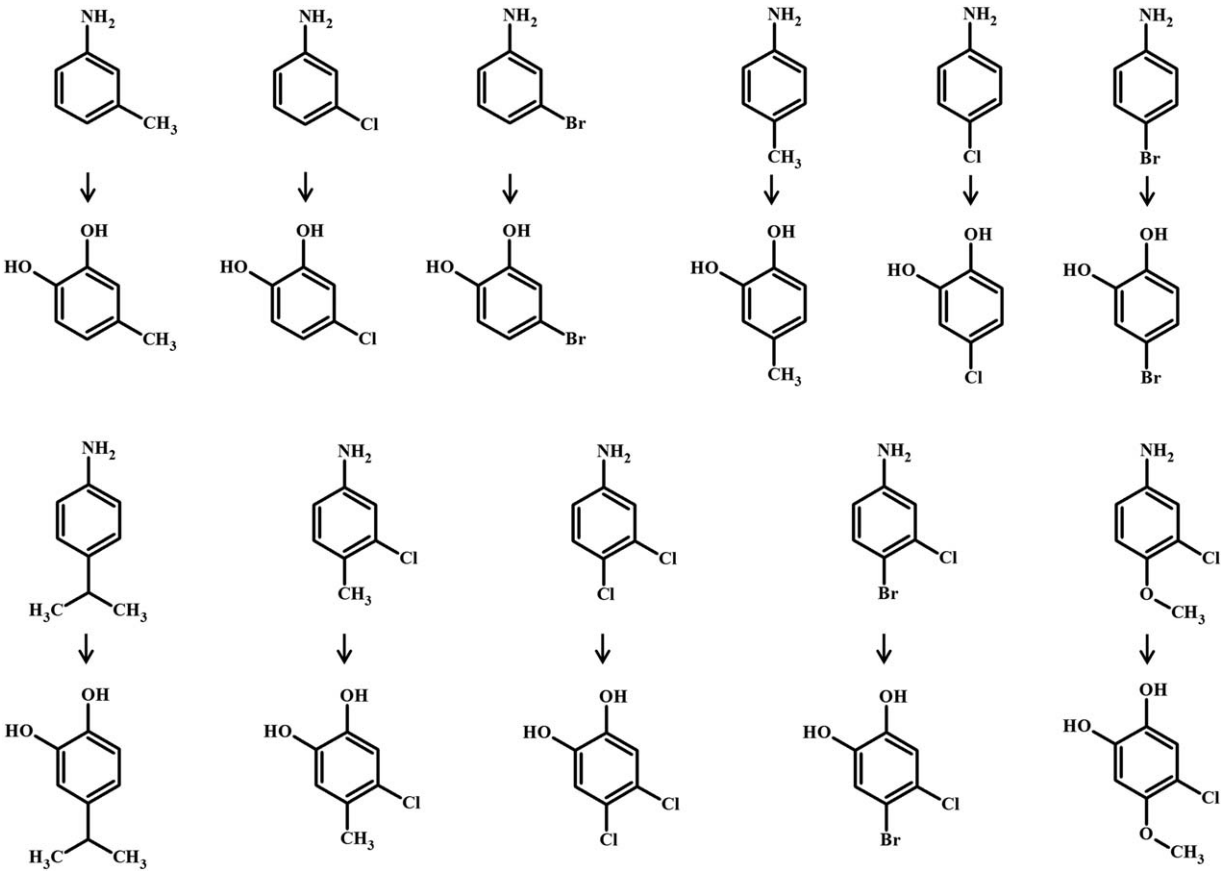

**Fig. 6.** Conversion of aniline derivatives to corresponding catechols by AdoQTA1A2B. The cell suspension assay was carried out as specified in the experimental procedure section. Substrate (30 mg/l) was added to the cell suspension of strain KT2440-Δ*catA*(pBBRQR) in MSM ($1.3 \times 10^8$ cells/ml). Cultures were incubated at 30°C and 200 rpm for 24 h. The products were analyzed by GC-MS (shown in Supporting Information Fig. S3). Strain KT2440-Δ*catA*(pBBR1MCS-5) was set as the control and the control experiments were carried out in parallel.

divergent GntR-type regulator gene *adoR2*. Inactivation of *adoR2* resulted in the enhanced transcription of *adoC* and *adoE* by approximately 5-fold compared with the wild-type in the absence of IPU (Fig. 7B), suggesting that AdoR2 is a transcription repressor of operon *adoXEGKLIJC*.

### *Distribution of IPU-catabolic genes in the genome of strain YBL2*

The key genes involved in IPU catabolism were all encoded on the plasmids in strain YBL2, specifically *pdmAB* on pYBL2-4, *ddhA* on pYBL2-5 and *adoQTA1A2BR*/*adoXEGKLIJC* on pYBL2-2 (Fig. 8), suggesting that these plasmids played important roles in the evolution of IPU-catabolic pathway in strain YBL2 and also indicating that IPU-catabolic genes were not simply acquired all at once. Interestingly, as shown in Fig. 8, among the six replication initiation genes (*repAs*) of plasmids, only the *repAs* of pYBL2-2, pYBL2-4 and pYBL2-5, which harbour IPU-catabolic genes, contained highly conserved regions with both strains Y57 and SRS2. Moreover, a major part of pYBL2-4 was highly conserved with strain Y57 and approximately half of pYBL2-4 was highly conserved with strain SRS2; further, the major part of pYBL2-5 was highly conserved with both strains Y57 and SRS2. Therefore, it seems likely that plasmids pYBL2-2, 4 and 5, particularly pYBL2-4 and −5, played key roles in the recruitment of the IPU catabolic genes in sphingomonads.

BLASTp results showed that the RepAs of the six plasmids in strain YBL2 have 94%–100% similarities to different replicons originating from sphingomonads, indicating that these replicons are specifically distributed in sphingomonads. Like many catabolic genes of xenobiotic compounds, all of the IPU-catabolic genes were found to be associated with mobile genetic elements (MGEs). *pdmAB* was flanked by a single copy of IS*Spwi2* and a Tn*Ase*-like transposon, as well as four copies of *tnpX*, which is a site-specific recombinase gene responsible for the excision and insertion of the transposons Tn*4451/53* (Lyras *et al.*, 2004) (Fig. 9). *ddhA* was surrounded by a single copy of *tnpX* and a Tn*As2*-like transposon. *adoQTA1A2BR* and *adoXEGKLIJC* were associated with

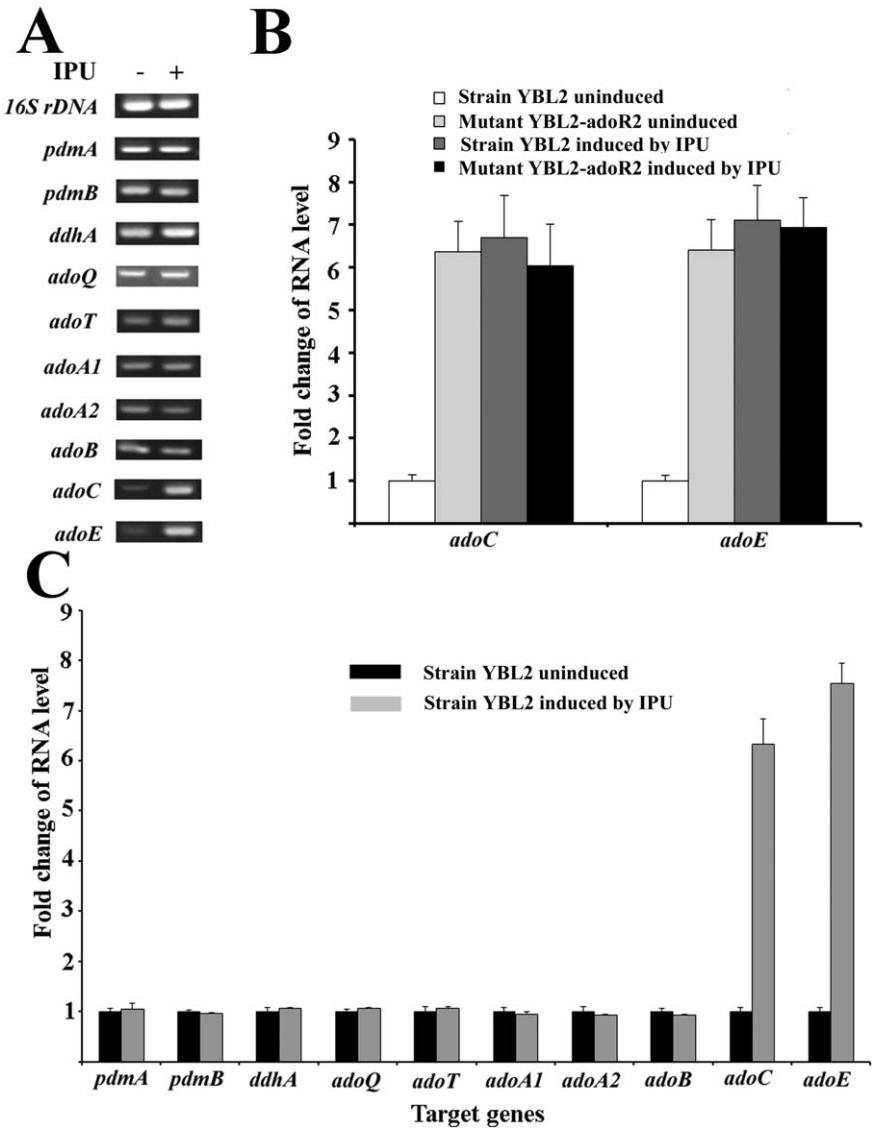


**Fig. 7.** Transcription of IPU-catabolic genes in response to IPU. RT-PCR (A) and RT-qPCR (C) analyses of the transcription of IPU-catabolic genes in strain YBL2 grown with or without IPU. (B) RT-qPCR analyses of the transcription of *adoC* and *adoE* in strain YBL2 and mutant YBL2-adoR2 grown with or without IPU. Detailed method was specified in the experimental procedures section.

IS*Spwi2*, IS*Ssp2*-like insertion sequence (IS) and multiple copies of IS*6100* (Fig. 9). IS*6100* has been demonstrated to play an important role in the spread and reorganization of the *lin* catabolic genes in sphingomonads (Lal *et al*., 2006). Thus, in addition to the mobile plasmids pYBL2-2, pYBL2-4 and pYBL2-5, MGEs such as transposons and ISs, could also play important roles in the evolution of the IPU-catabolic pathway.

## Discussion

### *The evolution of the conserved IPU-catabolic pathway in sphingomonads occurred in less than 50 years*

Microbes usually require decades to evolve a complete mineralizing pathway for a newly introduced pesticide. For the recalcitrant HCH, the first mineralizing strain was isolated in the 1980s, approximately 40 years after the beginning of widespread use of HCH (Bachmann *et al*., 1988; Ryozo *et al*., 1989). Similarly, the first strain capable of mineralizing atrazine was isolated in 1994, nearly 40 years after its first application (Yanze-Kontchou and Gschwind, 1994). While, the evolution of the mineralizing pathway for 2,4-dichlorophenoxyacetic acid (2,4-D) seemed faster. The first 2,4-D mineralizing strain was isolated approximately 26 years after its first use (Bollag *et al*., 1968).

IPU was first introduced into environment in 1972. However, approximately 20 types of PUHs with similar structures have entered the market since the 1950s. For example, monuron, the first herbicide among the PUHs, came into the market in 1951. On the other hand, both the IPU-mineralizing sphingomonads (strains YBL1, YBL2, YBL3, Y57 and SRS2) and PdmAB could also degrade other PUHs, such as diuron, monuron and chlorotoluron. In addition, AdoQTA1A2T showed activity toward all of the aniline derivatives related to the commonly used PUHs.

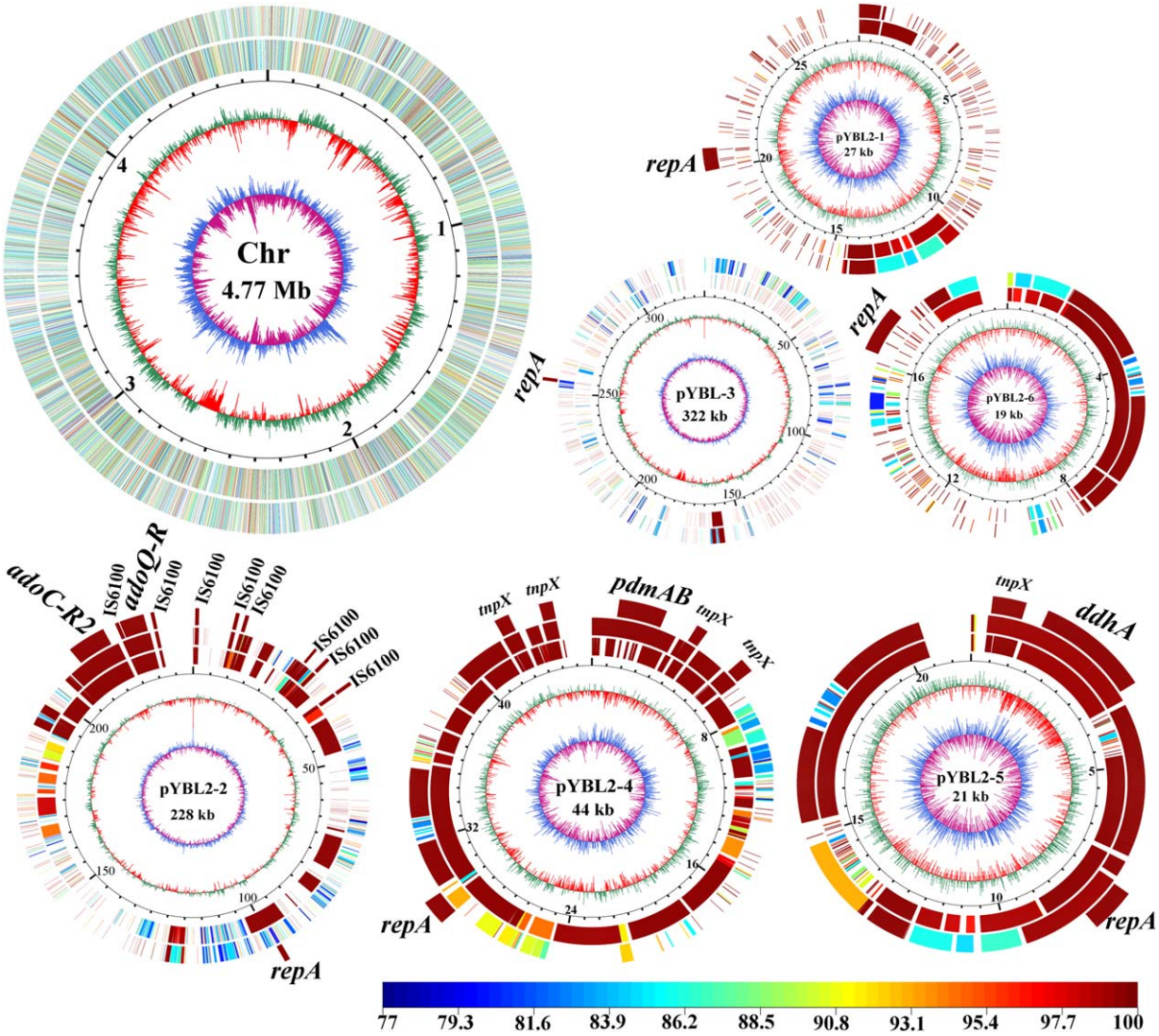


**Fig. 8.** Comparative genomic analysis of the three IPU-mineralizing sphingomonads. The genome of strain YBL2 contains one circular chromosome and six circular plasmids. Note that the replicons are not arranged in order and are not drawn to scale. The size scale for each replicon is marked on the black scale circle. BLASTn searches for the strain YBL2 genome sequence toward the genome sequences of strains Y57 (the first cycle outside of the black scale circle) and SRS2 (the second cycle outside of the black scale circle) are shown. The positions of *repA*, transposon, IS and IPU-catabolic gene on the plasmids are displayed in the outer ring. The region containing the sequence that is homologous to the other two IPU-mineralizing sphingomonads is coloured as a gradient to indicate the level of similarity. The results of GC skew and G + C content are shown inside the black scale circles: GC skew (inside), the parts greater than and less than zero are coloured cyan and magenta, respectively; G + C content (outside), the parts greater than and less than the average of each replicon are coloured green and red, respectively. Drawn by MATLAB (http://www.mathworks.com/).

Considering these facts, the IPU-catabolic pathway might have evolved under the pressures of different PUHs and not only IPU alone. Because we are not sure whether the first isolated IPU-mineralizing strain of *Pseudomonas fluorescens* (Roberts *et al.*, 1998) has the same IPU-catabolic pathway with sphingomonads, we can only speculate that the time for the evolution of the conserved IPU-catabolic pathway in sphingomonads was less than 50 years, from the launch of the first xenobiotic monuron in 1951 to the isolation of the first IPU-mineralizing strain SRS2 in 2001 (Sørensen *et al.*, 2001).

Some microorganisms belonging to specific taxonomies have been reported to show unique abilities to mineralize pesticides. For instance, most HCH-mineralizing strains are sphingomonads (Lal *et al.*, 2006). It is interesting that the majority of the strains reported to be able to mineralize IPU are also sphingomonads, and they share the conserved pathway for IPU catabolism (Fig. 10). It has been

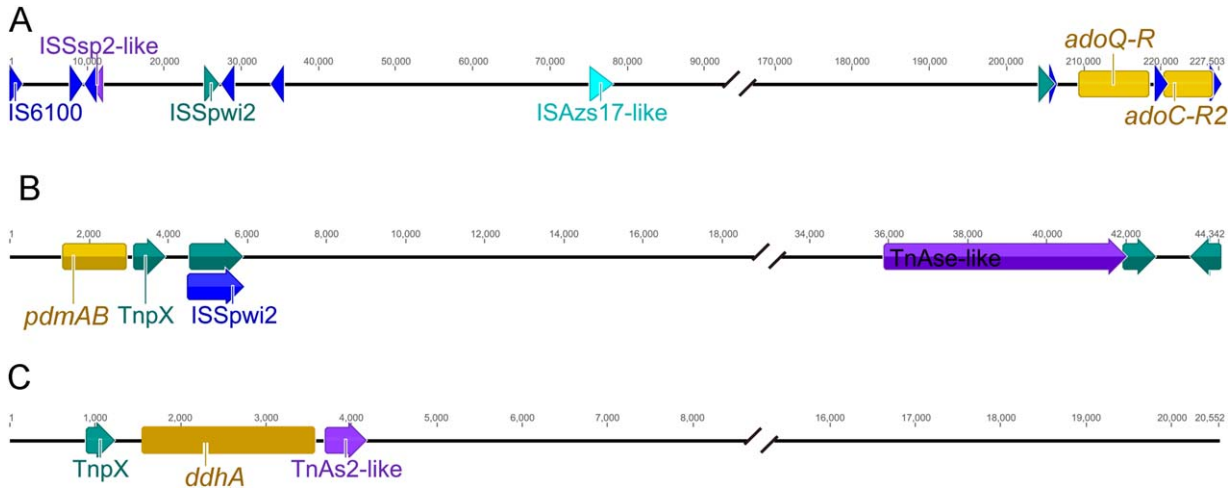


**Fig. 9.** The distributions of transposons, ISs and IPU-catabolic genes in the plasmids of strain YBL2.
(A) pYBL2-2.
(B) pYBL2-4.
(C) pYBL2-5. Each type of transposon/IS is shown in the same colour. The IPU-catabolic genes are indicated with yellow rectangles.

reported that genes involved in the mineralization of HCH, atrazine, 2,4-D are highly conserved among hosts that have been isolated from geographically distinct locations (Liang *et al*., 2012). Among IPU-mineralizing sphingomonads, including strains YBL1, YBL2, YBL3, Y57, pu21 and SRS2, which were isolated from China and Europe, the IPU-catabolic genes are also highly conserved. Therefore, it is reasonable to speculate that this pathway is also conserved in other IPU-catabolic sphingomonads.

### *The genes involved in the IPU-mineralizing pathway are novel and distinct*

In our previous study, PdmAB was confirmed to be a novel bacterial *N*-demethylase belonging to the RO system (Gu *et al*., 2013). ROs are multicomponent enzyme systems that are remarkably diverse and are found in bacteria isolated from various habitats (Ferraro *et al*., 2005). The majority of the characterized ROs are involved in hydroxylating aromatic ring substrates, while several ROs have been identified as *O*-demethylases or *N*-demethylases. It has been suggested that divergent evolution has resulted in two groups of ROs: one for catalyzing aromatic ring hydroxylation and one for a *C—O*/*C—N* bond-cleaving reaction (Summers *et al*., 2012). Here, PdmAB was able not only to *N*-demethylate IPU and MDIPU but also to hydroxylate MDIPU. Therefore, it is likely that, although the ROs have evolved into two groups, some ROs continue to possess the functions of both aromatic ring hydroxylation and *C—O*/*C—N* bond-cleaving demethylation.

Sequence analysis provided little information about the function or even the type of DdhA, indicating that it is a

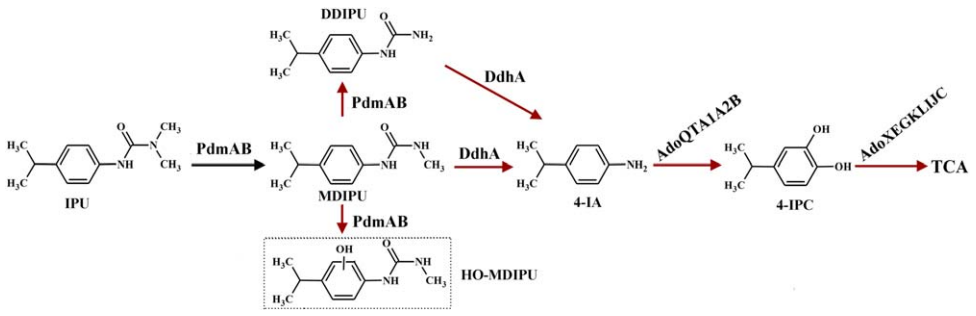


**Fig. 10.** The IPU-mineralizing pathway in strain YBL2 and the enzymes involved. IPU was first converted into MDIPU and trace amounts of DDIPU and HO-MDIPU by PdmAB; then, MDIPU and DDIPU were transformed into 4-IA by DdhA; finally, 4-IA was converted into 4-IPC by AdoQTA1A2BR, followed by the catechol *meta*-cleavage pathway and the TCA cycle. The previously characterized reaction is shown by black arrows and the reactions investigated in this work are shown by red arrows. The dashed box indicates that the structure is predicted based on LC-MS/MS analysis.

novel enzyme. Most of the BLASTp hits in GenBank were hypothetical proteins and showed minimal similarity with DdhA, and even no putative conserved domain was detected. The only clue is that the C-terminal half of DdhA (80 kDa) showed 21% similarity to that of CehA (87 kDa), which is a carbaryl hydrolase (Hashimoto *et al.*, 2002). However, CehA is also a rather uncommon enzyme, and sequence analysis provided no information about its classification. Interestingly, DdhA was also found to be able to hydrolyze carbaryl (data not shown).

The sphingomonads are among the most prominent groups for the degradation of aromatic compounds. Strangely, few sphingomonads were reported to be aniline-degraders (Liu *et al.*, 2005), and no aniline-degrading genes have been identified in sphingomonads. Therefore, to the best of our knowledge, *adoQTA1A2BR* is the first aniline dioxygenase gene cluster identified in sphingomonads. Furthermore, both the sequence comparison and function identification showed that the *ado* cluster is clearly distinct from the *tdn* (Fukumori and Saint, 1997) and *atd* (Fujii *et al.*, 1997) groups. First, many previous sequence comparisons have shown low similarity of xenobiotic-degrading genes between sphingomonads and non-sphingomonads (Stolz, 2009). In agreement with these results, the sequence and organization of *adoQTA1A2BR* were clearly distinct from those of the *tdn* and *atd* groups. Due to the low similarities of AdoQ to TdnQ/AtdQ, *adoQTA1A2BR* should be classified as the third group of the aniline dioxygenase gene cluster according to the classification method suggested by Urata *et al.* (2004). TdnR and AtdR, from both of the existing groups, are LysR-type positive regulators, but the truncated AdoR is a GntR-type, and our data indicated that it is a negative regulator. Second, according to our results and previous reports, the substrate specificity of AdoQTA1A2B is also different from that of the other two groups. Both AtdA and TdnQTA1A2B showed low activity toward halo-substituted anilines, particularly dihalo-substituted anilines (Fujii *et al.*, 1997; Fukumori and Saint, 1997). However, AdoQTA1A2B exhibited strong activity toward some halo-substituted anilines.

### *MGEs played key roles in the recruitment and refinement of the IPU-mineralizing pathway*

The evolution of key catabolic genes usually includes mutation, transposition, DNA rearrangement, horizontal gene transfer (HGT) and evolutionary refinement (Ochman *et al.*, 2000). The G + C% contents of *pdmAB* (55.37%), *ddhA* (46.89%), *adoQTA1A2BR* (60.76%) and *adoXEGKLIJC* (61.62%) are obviously lower than those of the genomes of their host strains YBL2 (64.42%), Y57 (65.96%) and SRS2 (63.83%), showing the features of HGT. HGT is usually mediated by MGEs, such as conjugative plasmids and ISs (Van der Meer *et al.*, 1992; Top and Springael, 2003; Juhas *et al.*, 2009).

Plasmids are among the most important elements in the evolution of prokaryotes and in their adaptations to changing environments (Brown Kav *et al.*, 2012; Frost *et al.*, 2005). The recently sequenced genome of the HCH-degrader *Sphingomonas* sp. strain MM-1 showed that the *lin* genes were located on five plasmids. For example, the *linF* was found on pISP0; *linA*, *linC* and a truncated *linF* on pISP1; *linDER* on pISP3; *linB*, *linC* and another truncated *linF* on pISP4; and *linGHIJ* on pISP0 (Tabata *et al.*, 2013). Similar phenomena were observed in strain YBL2. To recruit the genes for the IPU-catabolic pathway, it appears that strain YBL2 accepted plasmids pYBL2-2, pYBL2-4 and pYBL2-5. Compared with other three plasmids, pYBL2-2, −4 and −5 contained more regions that were highly conserved with the genomes of strains Y57 and SRS2, indicating pYBL2-2, pYBL2-4 and pYBL2-5 played key roles in the recruitment of the IPU-catabolic genes in strains YBL2, Y57 and SRS2.

Other important MGEs in the evolution of the IPU-catabolic pathway are transposons and ISs. All four transcription units were found to be flanked by ISs or *tnpX*. The IS*6100* element belongs to the IS*6* family and has an extremely broad host range (Mahillon and Chandler, 1998), while *tnpX* was only found in IPU-degrading sphingomonads. These MGEs probably facilitated the evolution of the IPU-catabolic pathway by transferring genes from chromosomes to plasmids or altering native regulators. For example, theoretically, the expression of the genes encoding ROs should be inducible and should be suppressed without the corresponding substrates. However, the regulators of *pdmAB* and *adoQTA1A2B* were destroyed by *tnpX* and IS*6100*, respectively, allowing for their constitutive expression. All these data indicated that MGEs played important roles during the evolution of the IPU-catabolic pathway.

### *The IPU-catabolic pathway is probably in its early stage of evolution*

Like other xenobiotic-catabolic pathways, this newly evolved IPU-catabolic pathway is still far from ideal. At least several aspects are clearly not optimal. The first deficiency is the regulation of gene expression. Constitutive expression of *pdmAB*, *ddhA* and *adoQTA1A2BR* consumes large amounts of energy, causing a heavy burden on cells in the absence of IPU. The second issue is the organization of the catabolic genes. The catabolic genes were organized into four transcription units and were distributed on three plasmids, which is obviously undesirable for coordinating regulation. Additionally, aniline is a naturally existing compound, and all of the reported catabolic gene clusters have been well evolved (Fujii *et al.*, 1997; Geng *et al.*, 2009). However, compared with other aniline

catabolic gene clusters, *ado* is clearly immature. The aniline dioxygenase genes and the catechol *meta*-cleavage pathway genes were neither organized into one transcription unit nor coordinately regulated. All these indicated that the IPU-catabolic pathway is in its early stage of evolution.

## *Experimental procedures*

*Chemicals and media.* IPU (99% purity), MDIPU (99% purity) and DDIPU (99% purity) were purchased from Dr. Ehrenstorfer-Schafers (Augsburg, Germany). 4-IA (98% purity) and other aniline derivatives (98% purity) were purchased from Sigma (Munich, Germany). Mineral salt medium (MSM) contained (g/L) NaCl 1.0, $NH_4NO_3$ 1.0, $K_2HPO_4$ 1.5, $KH_2PO_4$ 0.5, $MgSO_4$ 0.1, $CaCl_2{\cdot}2H_2O$ 0.1 and 1 ml trace element solution (Pfennig and Lippert, 1966) at pH 7.0. The concentrations of various carbon sources supplemented in MSM were indicated in experimental procedures below.

*Bacterial strains, plasmids and culture conditions.* The bacterial strains and plasmids used in this study are listed in Supporting Information Table S3. The *E. coli* strains were grown at 37°C in Luria–Bertani (LB) medium. Other bacterial strains were grown aerobically at 30°C in LB medium, R2A medium (Difco) or MSM medium supplemented with various carbon sources. Antibiotics were added as follows (mg/l): ampicillin (Amp), 100; spectinomycin (Spe), 100; gentamycin (Gm), 30; chloramphenicol (Cm), 20; kanamycin (Km), 50; and streptomycin (Str), 100.

*Expression of DdhA in sphingomonads and* E. coli. The fragment containing *ddhA* and its 220-bp upstream region was amplified from strain YBL2 using the primer pair pBBRddhAF/pBBRddhAF (Supporting Information Table S4). The PCR product was then ligated with a *BamHI*-digested broad-host-range plasmid pBBR1MCS-2 (Kovach *et al*., 1995), yielding pBBRddhA, which was transferred into *S. wittichii* RW1 or strain YBL2 through triparental mating.

To express DdhA in *E. coli*, *ddhA* was amplified using the PCR primer pair petddhaF/petddhaB. The PCR product was then ligated with pET-29a(+)to generate pETddhA. For DdhA expression, *E. coli* BL21(DE3) harbouring pETddhA was grown at 37°C to $OD_{600nm} = 0.5$, and then 0.2 mM IPTG was added. The induction was performed at 16°C for 12 h.

*Expression of AdoQTA1A2B in* P. putida *KT2440.* The fragment containing the 300-bp upstream region of the start codon of *adoQ*, together with *adoQTA1A2BR*, was amplified with the primer pair pBBRQRF/pBBRQRB and was ligated with pBBR1MCS-5 (Kovach *et al*., 1995) to generate pBBRQR, which was introduced into strain KT2440 via triparental mating.

*Disruption of the IPU-catabolic genes in strain YBL2.* To disrupt *ddhA* through a single crossover event, a 580-bp DNA fragment (in the middle of *ddhA*) (amplified with the primer pair pJQ200ddhAF/pJQ200ddhAF) was cloned at the *BamHI* site of the suicide plasmid pJQ200SK (Quandt and Hynes, 1993) to generate pJQddhA. pJQddhA was delivered into strain YBL2 from *E. coli* via triparental mating, and the transconjugants were selected on LB plates supplemented with Str and Gm. Because *ddhA* was located on the plasmid in strain YBL2, the transconjugant was sub-cultured with Gm for more than ten generations to obtain the pure *ddhA*-disrupted mutant YBL2-ddhA. The gene *ddhA* was also inactivated in the *pdmAB*-disrupted strain YBL2-1017, yielding the *pdmAB*/*ddhA* double mutant YBL2-ABA. Similarly, *adoC* and *adoR2* were inactivated in strain YBL2, generating mutants YBL2-adoC and YBL2-adoR2, respectively.

*DNA manipulation, sequencing and analysis.* Oligonucleotide synthesis and DNA sequencing were performed by Invitrogen Technology (Shanghai, China) Co., Ltd. BLASTn and BLASTp were used for the nucleotide sequence and deduced amino acid identity searches, respectively. For phylogenetic analysis, Clustal_X was used to align all of the sequences (Thompson *et al*., 1997). The multiple sequence alignment was then imported into MEGA software, version 5.0 (Tamura *et al*., 2011), and the phylogenetic tree was constructed by the Neighbour-Joining method.

Genome sequencing was performed by Shanghai Biozeron Biotechnology Co., Ltd. (Shanghai, China). The complete genome sequence of strain YBL2 was assembled using a method combining Illumina MiSeq sequencing technology, Pacific Biosciences platform and PCR validation (Luo *et al*., 2012; Koren *et al*., 2012; Chin *et al*., 2013). The draft genome sequence of strain Y57 was generated using the Illumina MiSeq platform. The draft genome sequence of strain SRS2 was kindly provided by Sebastian R. Sørensen (Nielsen *et al*., 2015). Annotation was performed using Glimmer 3.02 (Delcher *et al*., 2007), tRNAscan-SE version 1.3.1 (Lowe and Eddy, 1997) and Barrnap 0.4.2 (Lagesen *et al*., 2007). The genome map of strain YBL2 was drawn by MATLAB (http://www.mathworks.com/). The blastn search of the genome of strain YBL2 was conducted with the genome sequences of strains Y57 and SRS2, and the results are shown on the map. The whole-genome alignment tool MUMmer (Kurtz *et al*., 2004) was used to analyze the genome synteny among the three genomes. Transposons and insertion sequences were identified using ISfinder (Kurtz *et al*., 2004).

To find the highly conserved genes among strains YBL2, Y57 and SRS2, the genome sequences of strains Y57 and SRS2 were searched against the genome of strain YBL2, using BLASTn with default parameters; the BLASTn output

was further filtered by an in-house PERL script using the strict criteria of alignment length $\geq$ 100 bp, E-value $<$ 1e-5 and similarity $\geq$ 95%; then, the highly conserved sequences were retrieved with a filtered BLASTn output.

*Biotransformation and biodegradation experiments.* To determine the growth of strain YBL2 on IPU, MSM without $NH_4NO_3$ but supplemented with 170 μM IPU was used as the basal medium. When IPU was used as the sole carbon source, $NH_4NO_3$ (1 g/l) was added. When IPU was used the sole nitrogen source, glucose (5 g/l) was added. The overnight-precultured strain YBL2 cells (in LB broth) were harvested, washed twice with basal medium and then diluted to a final $OD_{600nm}$ of about 0.04 ($4.8 \times 10^6$ cells/ml). Cultures (400 ml) were incubated in 1000-ml Erlenmeyer flasks on a rotatory shaker (200 rpm) at 30°C and aliquots from each suspension were taken regularly during the incubation period. Cell density was monitored by measuring the turbidity at 600 nm in a 1-cm cuvette with UV–visible light spectrophotometer. The samples were extracted with an equal volume of dichloromethane. The extracts were dried over anhydrous $Na_2SO_4$ and evaporated at room temperature. IPU was detected with HPLC as described below. Three replicates were performed for each treatment.

To determine the degradation kinetics of IPU, DMIPU, DDIPU and 4-IA by strain YBL2, the overnight-precultured strain YBL2 cells (in LB broth) were harvested, washed twice with MSM and then suspended in MSM containing 145 μM substrate to a final $OD_{600nm}$ of about 1.0 ($1.2 \times 10^8$ cells/ml). Cultures (400 ml) were incubated in 1000-ml Erlenmeyer flasks on a rotatory shaker (200 rpm) at 30°C. Substrate was detected with HPLC method as described below.

To validate the function of PdmAB, pBBRPAB (Gu *et al.*, 2013) which contained *pdmAB* under the control of its native promoter, was introduced into strain RW1 through triparental mating. Strain RW1(pBBRPAB) was grown overnight in LB broth supplemented 30 mg/l Gm. The cells were harvested, washed twice with MSM and then suspended in MSM containing 170 μM IPU to a final $OD_{600nm}$ of about 3.0 ($3.3 \times 10^8$ cells/ml). Cultures (400 ml) were incubated in 1000-ml Erlenmeyer flasks at 30°C and 200 rpm for 48 h. Substrate was analyzed with HPLC and LC-MS/MS as described below. Strain RW1 containing the empty vector pBBR1MCS-5 was used as the control.

To analyze the products of MDIPU treated by the *ddhA*-disrupted mutant YBL2-ddhA, the mutant was grown overnight in LB broth. The cells were harvested, washed twice with MSM and then suspended in MSM containing 170 μM MDIPU to a final $OD_{600nm}$ of about 3.0 ($3.6 \times 10^8$ cells/ml). Cultures (100 ml) were incubated in 500-ml Erlenmeyer flasks at 30°C and 200 rpm for 24 h. Substrate was analyzed with HPLC and LC-MS/MS as described below.

To investigate the substrate range of AdoQTA1A2B, the overnight-precultured strain KT2440-Δ*catA*(pBBRQR) (in LB broth) were harvested, washed twice with MSM and then suspended in MSM containing 30 mg/L substrate to a final $OD_{600nm}$ of about 1.0 ($1.3 \times 10^8$ cells/ml). Cultures (20 ml) were incubated in 100-ml Erlenmeyer flasks at 30°C and 200 rpm for 24 h. Strain KT2440-Δ*catA*(pBBR1MCS-5) was set as the control. Substrate was detected by GC-MS as described below.

*Enzyme assay.* To test the activity of DdhA, *E. coli* BL21(DE3)(pET-ddhA) was grown to $OD_{600nm} = 0.5$ in LB broth, and then the cells were incubated with 0.2 mM IPTG at 200 rpm and 16°C for 12 h. The cells (100 ml) were harvested, suspended in 15 ml Tris-HCl buffer (20 mM, pH 7.0) and lysed using a French press, followed by centrifugation at 15,000*g* for 30 min. Protein concentration in cell-free extract was determined by the Bradford method (Bradford, 1976) with bovine serum albumin as the standard. The reaction mixture (5 ml) contained 20 mM Tris-HCl buffer (pH 7.0), 170 μM substrate and 100 μl cell-free extract (10.3 μg crude protein). The reaction was performed at 30°C without shaking. The reaction was stopped by incubation at 70°C for 10 min. Substrate concentration was analyzed by HPLC as described below. One unit of enzyme activity was defined as the amount of enzyme required to hydrolyze 1 μmol of substrate per hour at 30°C. Specific activities are expressed as units per milligram of crude protein. *E. coli* BL21(DE3) containing the empty vector was used as the control. All assays were performed independently in triplicate.

### *RT-PCR and RT-qPCR*

Strain YBL2 or mutant YBL2-adoR2 was grown in R2A broth until $OD_{600nm} = 0.5$. IPU was added to the culture at a final concentration of 145 μM. Cells were obtained for RNA extraction at different time points after the addition of IPU (0.5, 1, 2 and 3 h). RNA was extracted from approximately $2 \times 10^9$ cells using the Spin Column Bacterial Total RNA Purification Kit (Sangon Biotech, Shanghai). The RNA concentration was first determined using a Nanodrop and then was diluted to equal concentrations. Reverse transcription was performed with the PrimeScript RT reagent kit with gDNA Eraser (Takara, Dalian). qPCR was performed using the CFX96 Real-Time PCR Detection system (Bio-Rad) with SYBR *Premix Ex Taq* (Tli RNaseH Plus) (Takara, Dalian) and qPCR primers (Supporting Information Table S4). The experiments were performed with controls and IPU-induced cells of strain YBL2 in triplicate. The threshold cycle ($C_T$) value of the target gene was normalized to the reference gene (16S rRNA gene). The $2^{\Delta\Delta CT}$ method was used to calculate the relative expression level (Livak and Schmittgen, 2001).

*Analytical methods*

For HPLC analysis, a separation column (internal diameter, 4.6 mm; length, 250 mm) filled with Kromasil 100-5 $C_{18}$ was used. The mobile phase was acetonitrile:water (40:60, V:V), and the flow rate was 0.8 ml/min. The detection wavelength was 230 nm and the injection volume was 20 μl.

For LC-MS/MS analysis, the products were detected by a UHPLC system (Dionex, Thermo, The United States) connected to a LTQ-Orbitrap XLhybrid mass spectrometer (Thermo Fisher Scientific). The instrument was controlled using the Tune program (version 2.6.0) and the Chromeleon program. The UHPLC column was Hypersil GOLD $C_{18}$ (100 mm × 2.10 mm, 3 μm particle size, Thermo Fisher Scientific). The mobile phase was solvent A (0.2% formic acid in water) and solvent B (acetonitrile) at a ratio of 60:40 (V:V). The data-dependent scan mode was applied for MS/MS fragmentation. The ion source was equipped with a HESI source. The product ions were generated by the HCD ion trap at a normalized collision energy of 40% and a q-activation of 0.25, using an isolation width of 1.0 Da. The data were processed using the Xcalibur software.

To detect aniline derivatives and corresponding catechols, GC-MS analysis was performed on a Thermo Trace DSQ mass spectrometer under the following conditions. Helium was used as a carrier gas at a flow rate of 1.2 ml/min. GC was conducted using an RTX-5MS column (15 m × 0.25 mm × 0.25 mm, Restek Corp., The United States). The column temperature was programmed from 50°C (1.5 min hold) to 150°C (1 min hold) at 40°C/min and then from 150°C to 200°C (6 min hold) at 10°C/min and finally from 200°C to 260°C (6 min hold) at 50°C/min. The injector temperature was set at 220°C with a split ratio of 20:1. The interface temperature and ion source temperature were set to 250°C. The column outlet was inserted directly into the electron ionization source block, and it operated at 70 eV.

*Nucleotide sequence accession number*

The genome sequences of strains YBL2, Y57 and SRS2 have been submitted to the GenBank database under accession no. CP010954-CP010960, LDES01000000 and LARW00000000, respectively.

## Acknowledgments

We appreciate Dr. Sebastian R. Sørensen from Geological Survey of Denmark and Greenland for providing the genome sequence of strain SRS2 and Dr. Masahiro Takeo from University of Hyogo for providing the strain of *P. putida* KT2440-Δ*catA*. This work was supported by grants from the National Key Basic Research Program of China (2015CB150505), the Chinese National Science Foundation for Excellent Young Scholars (31222003), the Chinese National Natural Science Foundation (31000060 and 31470225), the Program for New Century Excellent Talents in University (NCET-12-0892) and the Outstanding Youth Foundation of Jiangsu Province (BK20130029). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## References

Abe, T., Masai, E., Miyauchi, K., Katayama, Y., and Fukuda, M. (2005) A tetrahydrofolate-dependent *O*-demethylase, LigM, is crucial for catabolism of vanillate and syringate in *sphingomonas paucimobilis* SYK-6. *J Bacteriol* **187:** 2030–2037.

Bachmann, A., de Bruin, W., Jumelet, J.C., Rijnaarts, H.H., and Zehnder, A.J. (1988) Aerobic biomineralization of alpha-hexachlorocyclohexane in contaminated soil. *Appl Environ Microbiol* **54:** 548–554.

Bending, G.D., Lincoln, S.D., Sørensen, S.R., Morgan, J.A.W., Aamand, J., and Walker, A. (2003) In-field spatial variability in the degradation of the phenyl-urea herbicide isoproturon is the result of interactions between degradative *sphingomonas* spp. And soil pH. *Appl Environ Microbiol* **69:** 827–834.

Bollag, J.M., Helling, C.S., and Alexander, M. (1968) 2,4-D metabolism. Enzymic hydroxylation of chlorinated phenols. *J Agric Food Chem* **16:** 826–828.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **529:** 248–254.

Brown Kav, A., Sasson, G., Jami, E., Doron-Faigenboim, A., Benhar, I., and Mizrahi, I. (2012) Insights into the bovine rumen plasmidome. *Proc Natl Acad Sci USA* **109:** 5452–5457.

Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., *et al.* (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* **10:** 563–569.

Copley, S.D. (2009) Evolution of efficient pathways for degradation of anthropogenic chemicals. *Nat Chem Biol* **5:** 559–566.

Cox, L., Walker, A., and Welch, S.J. (1996) Evidence for accelerated degradation of isoproturon in soils. *Pestic Sci* **48:** 253–260.

Da Rocha, M.S., Arnold, L.L., Dodmane, P.R., Pennington, K.L., Qiu, F., De Camargo, J.L., *et al.* (2013) Diuron metabolites and urothelial cytotoxicity: in vivo, in vitro and molecular approaches. *Toxicology* **314:** 238–246.

Delcher, A.L., Bratke, K.A., Powers, E.C., and Salzberg, S.L. (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **23:** 673–679.

Dwivedi, S., Singh, B.R., Al-Khedhairy, A.A., and Musarrat, J. (2011) Biodegradation of isoproturon using a novel *pseudomonas aeruginosa* strain JS-11 as a multi-functional bioinoculant of environmental significance. *J Hazard Mater* **185:** 938–944.

El Sebai, T., Lagacherie, B., Soulas, G., and Martin-Laurent, F. (2004) Isolation and characterisation of an

isoproturon-mineralising *methylopila* sp. Strain TES from french agricultural soil. *FEMS Microbiol Lett* **239:** 103–110.

Fabbri, D., Minella, M., Maurino, V., Minero, C., and Vione, D. (2015) Photochemical transformation of phenylurea herbicides in surface waters: a model assessment of persistence, and implications for the possible generation of hazardous intermediates. *Chemosphere* **119:** 601–607.

Ferraro, D.J., Gakhar, L., and Ramaswamy, S. (2005) Rieske business: structure-function of rieske non-heme oxygenases. *Biochem Biophys Res Commun* **338:** 175–190.

Frost, L.S., Leplae, R., Summers, A.O., and Toussaint, A. (2005) Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* **3:** 722–732.

Fujii, T., Takeo, M., and Maeda, Y. (1997) Plasmid-encoded genes specifying aniline oxidation from *acinetobacter* sp. Strain YAA. *Microbiology* **143:** 93–99.

Fukumori, F., and Saint, C.P. (1997) Nucleotide sequences and regulational analysis of genes involved in conversion of aniline to catechol in *pseudomonas putida* UCC22 (pTDN1). *J Bacteriol* **179:** 399–408.

Gaillardon, P., and Sabar, M. (1994) Changes in the concentration of isoproturon and its degradation products in soil and soil solution during incubation at two temperatures. *Weed Res* **34:** 243–251.

Geng, L., Chen, M., Liang, Q., Liu, W., Zhang, W., Ping, S., *et al*. (2009) Functional analysis of a putative regulatory gene, *tadR*, involved in aniline degradation in *delftia tsuruhatensis* AD9. *Arch Microbiol* **191:** 603–614.

Gu, T., Zhou, C., Sørensen, S.R., Zhang, J., He, J., Yu, P., *et al*. (2013) The novel bacterial *N*-demethylase PdmAB is responsible for the initial step of *N,N*-dimethyl-substituted phenylurea herbicide degradation. *Appl Environ Microbiol* **79:** 7846–7856.

Hane, J.K., Rouxel, T., Howlett, B.J., Kema, G.H., Goodwin, S.B., and Oliver, R.P. (2011) A novel mode of chromosomal evolution peculiar to filamentous ascomycete fungi. *Genome Biol* **12:** R45.

Hashimoto, M., Fukui, M., Hayano, K., and Hayatsu, M. (2002) Nucleotide sequence and genetic structure of a novel carbaryl hydrolase gene (*cehA*) from *rhizobium* sp. Strain AC100. *Appl Environ Microbiol* **68:** 1220–1227.

Hussain, S., Arshad, M., Springael, D., Sørensen, S.R., Bending, G.D., Devers-Lamrani, M., *et al*. (2015) Abiotic and biotic processes governing the fate of phenylurea herbicides in soils: a review. *Crit Rev Environ Sci Technol* **45:** 1947–1998.

Hussain, S., Devers-Lamrani, M., El Azhari, N., and Martin-Laurent, F. (2011) Isolation and characterization of an isoproturon mineralizing *sphingomonas* sp. Strain SH from a french agricultural soil. *Biodegradation* **22:** 637–650.

Juhas, M., van der Meer, J.R., Gaillard, M., Harding, R.M., Hood, D.W., and Crook, D.W. (2009) Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev* **33:** 376–393.

Koren, S., Schatz, M.C., Walenz, B.P., Martin, J., Howard, J.T., Ganapathy, G., *et al*. (2012) Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nat Biotechnol* **30:** 693–700.

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M. 2nd, *et al*. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166:** 175–176.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., *et al*. (2004) Versatile and open software for comparing large genomes. *Genome Biol* **5:** R12.

Lagesen, K., Hallin, P., Rødland, E.A., Stærfeldt, H.-H., Rognes, T., and Ussery, D.W. (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* **35:** 3100–3108.

Lal, R., Dogra, C., Malhotra, S., Sharma, P., and Pal, R. (2006) Diversity, distribution and divergence of lin genes in hexachlorocyclohexane-degrading sphingomonads. *Trends Biotechnol* **24:** 121–130.

Liang, B., Jiang, J., Zhang, J., Zhao, Y., and Li, S. (2012) Horizontal transfer of dehalogenase genes involved in the catalysis of chlorinated compounds: evidence and ecological role. *Crit Rev Microbiol* **38:** 95–110.

Liu, Z.P., Wang, B.J., Liu, Y.H., and Liu, S.J. (2005) *Novosphingobium taihuense* sp. Nov., a novel aromatic-compound-degrading bacterium isolated from taihu lake, China. *Int J Syst Evol Microbiol* **55:** 1229–1232.

Livak, K.J., and Schmittgen, T.D. (2001) Analysis of relative gene expression data using Rreal-time quantitative PCR and the $2^{\Delta\Delta CT}$ method. *Methods* **25:** 402–408.

Lowe, T.M., and Eddy, S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic Sequence. *Nucleic Acids Res* **25:** 0955–0964.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., *et al*. (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* **1:** 18.

Lyras, D., Adams, V., Lucet, I., and Rood, J.I. (2004) The large resolvase TnpX is the only transposon-encoded protein required for transposition of the Tn4451/3 family of integrative mobilizable elements. *Mol Microbiol* **51:** 1787–1800.

Mahillon, J., and Chandler, M. (1998) Insertion sequences. *Microbiol Mol Biol Rev* **62:** 725–774.

Mosleh, Y.Y. (2009) Assessing the toxicity of herbicide isoproturon on *aporrectodea caliginosa* (oligochaeta) and its fate in soil ecosystem. *Environ Toxicol* **24:** 396–403.

Mudd, P.J., Hance, R.J., and Wright, S.J.L. (1983) The persistenceand metabolism of isoproturon in soil. *Weed Res* **23:** 239–247.

Nielsen, T.K., Sørensen, S.R., and Hansen, L.H. (2015) Draft genome sequence of isoproturon-mineralizing *Sphingomonas* sp. SRS2, isolated from an agricultural field in the United Kingdom. *Genome Announc* **3:** e00569–15.

Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405:** 299–304.

Oturan, N., Trajkovska, S., Oturan, M.A., Couderchet, M., and Aaron, J.J. (2008) Study of the toxicity of diuron and its metabolites formed in aqueous medium during application of the electrochemical advanced oxidation process "electro-fenton". *Chemosphere* **73:** 1550–1556.

Pfennig, N., and Lippert, K.D. (1966) Über das Vitamin-B-12-bedürfnis phototropher Schwefelbakterien. *Arch Microbiol* **55:** 245–256.

Quandt, J., and Hynes, M.F. (1993) Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. *Gene* **127:** 15–21.

Roberts, S.J., Walker, A., Cox, L., and Welch, S.J. (1998) Isolation of isoproturon-degrading bacteria from treated soil via three different routes. *J Appl Microbiol* **85:** 309–316.

Ryozo, I., Yuji, N., Keishi, S., Hidenori, W., Masao, F., Masamichi, T., *et al*. (1989) Dehydrochlorination of γ-hexachlorocyclohexane (γ-BHC) by γ-BHC-assimilating *pseudomonas paucimobilis*. *Agric Biol Chem* **53:** 2015–2017.

Shapir, N., Mongodin, E.F., Sadowsky, M.J., Daugherty, S.C., Nelson, K.E., and Wackett, L.P. (2007) Evolution of catabolic pathways: genomic insights into microbial s-triazine metabolism. *J Bacteriol* **189:** 674–682.

Sørensen, S.R., Ronen, Z., and Aamand, J. (2001) Isolation from agricultural soil and characterization of a *sphingomonas* sp. Able to mineralize the phenylurea herbicide isoproturon. *Appl Environ Microbiol* **67:** 5403–5409.

Sørensen, S.R., Bending, G.D., Jacobsen, C.S., Walker, A., and Aamand, J. (2003) Microbial degradation of isoproturon and related phenylurea herbicides in and below agricultural fields. *FEMS Microbiol Ecol* **45:** 1–11.

Stolz, A. (2009) Molecular characteristics of xenobiotic-degrading sphingomonads. *Appl Microbiol Biotechnol* **81:** 793–811.

Summers, R.M., Louie, T.M., Yu, C.L., Gakhar, L., Louie, K.C., and Subramanian, M. (2012) Novel, highly specific *N*-demethylases enable bacteria to live on caffeine and related purine alkaloids. *J Bacteriol* **194:** 2041–2049.

Sun, J.Q., Huang, X., He, J., Xu, J.L., and Li, S.P. (2006) Isolation identification of isoproturon degradation bacterium Y57 and its degradation characteristic. *China Environ Sci* **26:** 315–319.

Sun, J.Q., Huang, X., Chen, Q.L., Liang, B., Qiu, J.G., Ali, S., *et al*. (2009) Isolation and characterization of three *sphingobium* sp. Strains capable of degrading isoproturon and cloning of the catechol 1,2-dioxygenase gene from these strains. *World J Microbiol Biotechnol* **25:** 259–268.

Tabata, M., Ohtsubo, Y., Ohhata, S., Tsuda, M., and Nagata, Y. (2013) Complete genome sequence of the γ-hexachlorocyclohexane-degrading bacterium *sphingomonas* sp. Strain MM-1. *Genome Announc* **1:** e00247–e00213.

Takeo, M., Ohara, A., Sakae, S., Okamoto, Y., Kitamura, C., Kato, D., *et al*. (2013) Function of a glutamine synthetase-like protein in bacterial aniline oxidation via γ-Glutamylanilide. *J Bacteriol* **195:** 4406–4414.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28:** 2731–2739.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25:** 4876–4882.

Top, E.M., and Springael, D. (2003) The role of mobile genetic elements in bacterial adaptation to xenobiotic organic compounds. *Curr Opin Biotechnol* **14:** 262–269.

Urata, M., Uchida, E., Nojiri, H., Omori, T., Obo, R., Miyaura, N., *et al*. (2004) Genes involved in aniline degradation by *delftia acidovorans* strain 7N and its distribution in the natural environment. *Biosci Biotechnol Biochem* **68:** 2457–2465.

Van der Meer, J.R., de Vos, W.M., Harayama, S., and Zehnder, A.J. (1992) Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiol Rev* **56:** 677–694.

Wang, F., Zhou, J., Li, Z., Dong, W., Hou, Y., Huang, Y., *et al*. (2015) Involvement of the cytochrome P450 system EthBAD in the *N*-deethoxymethylation of acetochlor by *rhodococcus* sp. Strain T3-1. *Appl Environ Microbiol* **81:** 2182–2188.

Yanze-Kontchou, C., and Gschwind, N. (1994) Mineralization of the herbicide atrazine as a carbon source by a *pseudomonas* strain. *Appl Environ Microbiol* **60:** 4297–4302.

## Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Fig. S1.** Genome synteny of strains BYL2, Y57 and SRS2. The x axis refers to the subject genome sequence with the query genome sequence shown by the y axis. The coordinates of the best hit protein between the two compared genomes are plotted. The colours represent the level of similarity of the match. For the genomes of strains Y57 and SRS2, only contigs longer than 50 kb were analyzed.

**Fig. S2.** Analysis of HO-MDIPU by LC-MS/MS. (A) The HPLC spectrum of the separated HO-MDIPU. (B) The standard MS spectrum of HO-MDIPU. (C) The second-order MS spectrum of HO-MDIPU. MDIPU was treated with the cell suspension of *S. wittichii* RW1 harbouring *pdmAB*. Strain RW1(pBBRPAB) was grown overnight in LB broth supplemented with 30 mg/L Gm. The cells were harvested, washed twice with MSM and then suspended in MSM containing 170 μM IPU to a final $OD_{600nm}$ of about 3.0 ($3.3 \times 10^8$ cells/ml). Cultures (400 ml) were incubated in 1000-ml Erlenmeyer flasks at 30°C and 200 rpm for 48 h. HO-MDIPU was separated by HPLC and was then subjected to LC-MS/MS analysis with a UHPLC system (Dionex, Thermo, USA), connected to a LTQ OrbitrapXL hybrid mass spectrometer (Thermo Fisher Scientific). The possible compounds corresponding to each peak were predicted, supporting that the separated compound is HO-MDIPU, and the hydroxyl group is most likely located on the aromatic ring.

**Fig. S3.** GC-MS analysis of the degradation products of aniline derivatives by strain KT2440-Δ*catA* harbouring *adoQTA1A2BR*. (A) - (K) show the GC-MS analysis of different aniline derivatives and their products. The overnight-precultured strain KT2440-Δ*catA*(pBBRQR) cells (in LB broth) were harvested, washed twice with MSM and then suspended in MSM containing 30 mg/L substrate to a final $OD_{600nm}$ of about 1.0 ($1.3 \times 10^8$ cells/ml). Cultures (20 ml) were incubated in 100-ml Erlenmeyer flasks at 30°C and 200 rpm for 24 h. Strain KT2440-Δ*catA*(pBBR1MCS-5) was set as the control and the control experiments were carried out in parallel. GC-MS analysis was performed as described in experimental procedure section.

**Fig. S4.** Growth of strain YBL2 using different aniline derivatives as the sole carbon source in MSM. The overnight-precultured strain YBL2 cells (in LB broth) were harvested, washed twice with MSM and then diluted to the final $OD_{600nm}$ of about 0.04 ($4.8 \times 10^6$ cells/ml) or 0.1 ($1.2 \times 10^7$ cells/ml). Due to the toxicity of halogenated anilines, lower concentration of substrate and higher inoculation

amount of strain YBL2 were used. Cultures (400 ml) were incubated in 1000-ml Erlenmeyer flasks on a rotatory shaker (200 rpm) at 30°C and aliquots from each suspension were taken regularly during the incubation period. Cell density was monitored by measuring the turbidity at 600 nm a 1-cm cuvette with UV-visible light spectrophotometer. Aniline derivatives were detected with HPLC. All experiments were carried out in triplicate.

**Fig. S5.** UV scanning of the products of IPU degraded by strain YBL2 and mutant YBL2-adoC. The tested cells were washed twice with sterilized MSM and suspended in MSM at a final $OD_{600nm}$ of about 1.0. IPU was added to a final concentration of 145 μM. Cultures (20 ml) were incubated in 100-ml Erlenmeyer flasks at 30°C and 200 rpm for 12 h. The products were detected by UV-visible light spectrophotometer. The accumulated 4-IPC was further confirmed by GC-MS (data not shown).

**Fig. S6.** Analysis of the transcription of operons containing the IPU-catabolic genes in strain YBL2 by RT-PCR. (A) *ORF3*, *ddhA* and *ORF5* on plasmid pYBL2-5. (B) *tnpX*, *ORF1*, *pdmA* and *pdmB* on plasmid pYBL2-4. (C) *adoQ-TA1A2BR* on plasmid pYBL2-2. (D) *adoR2XEGKLIJC* on plasmid pYBL2-2. Total RNA of strain YBL2 induced by IPU was used as the template for RT-PCR analysis and the reaction performed without reverse transcription was used as the negative control (CK). Lane M, molecular marker (bp), 2000, 1000, 750, 500, 250, 100.

**Table S1.** Genome information of strains YBL2, Y57 and SRS2.

**Table S2.** Highly conserved ORFs (≥95% nucleotide similarity) among IPU-mineralizing sphingomonads YBL2, Y57 and SRS2.

**Table S3.** Strains and plasmids used in this study.

**Table S4.** Primers used in this study.