Microbiology (2011), **157**, 2891–2903 DOI 10.1099/mic.0.049387-0

Gentisate 1,2-dioxygenase, in the third naphthalene catabolic gene cluster of *Polaromonas naphthalenivorans* CJ2, has a role in naphthalene degradation

Hyo Jung Lee,[1] Jeong Myeong Kim,[1] Se Hee Lee,[1] Minjeong Park,[2] Kangseok Lee,[1] Eugene L. Madsen[3] and Che Ok Jeon[1]

Correspondence
Che Ok Jeon
cojeon@cau.ac.kr

[1]Schools of Biological Sciences and Research Center for Biomolecules and Biosystems, Chung-Ang University, Seoul 156-756, Republic of Korea

[2]Environmental Biotechnology National Core Research Center, Gyeongsang National University, Jinju 660-701, Republic of Korea

[3]Department of Microbiology, Cornell University, Ithaca, NY 14853-8101, USA

Received 1 March 2011
Revised 31 May 2011
Accepted 4 July 2011

Polaromonas naphthalenivorans strain CJ2 metabolizes naphthalene via the gentisate pathway and has recently been shown to carry a third copy of gentisate 1,2-dioxygenase (GDO), encoded by *nagI3,* within a previously uncharacterized naphthalene catabolic gene cluster. The role of this cluster (especially *nagI3*) in naphthalene metabolism of strain CJ2 was investigated by documenting patterns in regulation, transcription and enzyme activity. Transcriptional analysis of wild-type cells showed the third cluster to be polycistronic and that *nagI3* was expressed at a relatively high level. Individual knockout mutants of all three *nagI* genes were constructed and their influence on both GDO activity and cell growth was evaluated. Of the three knockout strains, CJ2Δ*nagI3* showed severely diminished GDO activity and grew slowest on aromatic substrates. These observations are consistent with the hypothesis that *nagI3* may prevent toxic intracellular levels of gentisate from accumulating in CJ2 cells. All three *nagI* genes from strain CJ2 were cloned into *Escherichia coli*: the *nagI2* and *nagI3* genes were successfully overexpressed. The subunit mass of the GDOs were ~36–39 kDa, and their structures were deduced to be dimeric. The K_m values of NagI2 and NagI3 were 31 and 10 μM, respectively, indicating that the higher affinity of NagI3 for gentisate may protect the wild-type cells from gentisate toxicity. These results provide clues for explaining why the third gene cluster, particularly the *nagI3* gene, is important in strain CJ2. The organization of genes in the third gene cluster matched that of clusters in *Polaromonas* sp. JS666 and *Leptothrix cholodnii* SP-6. While horizontal gene transfer (HGT) is one hypothesis for explaining this genetic motif, gene duplication within the ancestral lineage is equally valid. The HGT hypothesis was discounted by noting that the *nagI3* allele of strain CJ2 did not share high sequence identity with its homologues in *Polaromonas* sp. JS666 and *L. cholodnii* SP-6.

INTRODUCTION

Gentisic acid (2,5-dihydroxybenzoate) is a key intermediary metabolite in the biodegradative pathways responsible for the metabolism of many abundant aromatic compounds (Bayly & Barbour, 1984; Chapman, 1972). For example, one series of bacterially mediated reactions that both produce and consume gentisate is the gentisate biochemical pathway of naphthalene degradation (Fuenmayor *et al.*, 1998; Grund *et al.*, 1992), in which the metabolite 3-hydroxybenzoate is converted to pyruvate (Supplementary Fig. S1, available with the online version of this paper). Characteristics of the reaction catalysed by gentisate 1,2-dioxygenase (GDO; EC 1.13.11.4) resemble those of extradiol dioxygenases (Chen *et al.*, 2008). GDO catalyses the oxygen-dependent ring fission of gentisate between the carboxyl and proximal hydroxyl groups at positions 1 and 2 of the aromatic ring to form maleylpyruvate (Adams *et al.*, 2006; Kivisaar, 2009). Because ring fission is the critical step in the gentisate

Abbreviation: GDO, gentisate 1,2-dioxygenase.

A supplementary table and four supplementary figures are available with the online version of this paper.

pathway, GDO has been widely studied (Chen *et al.*, 2008; Feng *et al.*, 1999; Hirano *et al.*, 2007; Ohmoto *et al.*, 1991).

Polaromonas naphthalenivorans CJ2, previously found to be responsible for the degradation of naphthalene *in situ* at a coal-tar-waste-contaminated site (Jeon *et al.*, 2003; Jeon *et al.*, 2004), carries naphthalene-catabolic (*nag*) genes that catalyse the gentisate metabolic pathway, converting naphthalene into fumarate and pyruvate via salicylate (2-hydroxybenzoate) and gentisate (Fuenmayor *et al.*, 1998). The *nag* genes in *P. naphthalenivorans* CJ2 are homologous to those of *Ralstonia* sp. U2 which features a single continuous operon, but the metabolic genes of strain CJ2 are split into two clusters, comprising *nagRAdGHAb AcAdBFCQEDJI1–orf1–tnpA* (first gene cluster) and *nagR2–XI2KL* (second gene cluster) (Jeon *et al.*, 2006). Each cluster carries a gene (*nagI1* or *nagI2* for the respective clusters) encoding GDO. The second catabolic gene cluster is also essential for the degradation of 3-hydroxybenzoate (Park *et al.*, 2007a). These findings have been confirmed biochemically, and other physiological traits of strain CJ2, particularly the accumulation of toxic metabolites during naphthalene metabolism, have been elucidated (Park *et al.*, 2007b; Pumphrey & Madsen, 2007). The genome of *P. naphthalenivorans* CJ2 contains a third copy of the GDO gene (Yagi *et al.*, 2009) separated from *nagI1* by approximately 16.5 kb on the chromosome; however, the physiological role and genetic regulation of this third GDO catabolic gene in naphthalene catabolism by strain CJ2 have not been investigated to our knowledge. In this study, we analysed the organization and regulatory control of the cluster of genes containing the third copy of *nagI*, assessed patterns of the GDO transcription, characterized properties of two GDO proteins and delved into the possible origins of the third cluster.

METHODS

Bacterial strains, plasmids and growth conditions. All bacterial strains, vectors, plasmids and PCR primers used in the present study are listed in Table 1. *P. naphthalenivorans* CJ2 was grown at 20 °C and was maintained on mineral salt basal medium (MSB) (Stanier *et al.*, 1966) agar with either naphthalene vapour (MSB-N), 0.3 % (w/v) pyruvate (MSB-P), 0.3 % (w/v) gentisate (MSB-G), 0.3 % (w/v) 3-hydroxybenzoate (MSB-H) or 0.3 % (w/v) salicylate (MSB-S) as the sole carbon source. All *Escherichia coli* strains were grown at 37 °C in Luria-Bertani (LB) medium on a rotary incubator (180 r.p.m.). When required, the appropriate antibiotics and reagents were added to the medium: X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; 30 μg ml^{-1}), IPTG (40 μg ml^{-1}), kanamycin (20 μg ml^{-1}), rifampicin (200 μg ml^{-1}), ampicillin (50 μg ml^{-1}) and salicylate (1 mM).

DNA manipulation and sequence analysis. DNA manipulation and other molecular biology techniques were carried out using established standard protocols (Sambrook & Janssen, 2001). DNA sequencing was performed using an ABI model 3700 instrument at Macrogen (Korea). Transcription promoters and termination sequences of the third catabolic gene cluster were analysed using web-based programs (http://www.softberry.com/). BLASTX and BLASTN were used, respectively, to determine the deduced amino acid and nucleotide identity of the third catabolic gene cluster (Altschul *et al.*, 1990). Isoelectric points (pI) and molecular masses of GDO enzymes were calculated theoretically with the aid of the Lasergene software package (DNASTAR). The G+C content and the tetranucleotide frequency of the CJ2 chromosome were calculated using the Oligoweb interface (http://insilico.ehu.es/oligoweb/). The phylogeny of the *nagI* genes was analysed by using the web-based set of tools at Phylogeny.fr (http://www.phylogeny.fr/version2_cgi/index.cgi; Dereeper *et al.*, 2008): multiple alignments were prepared with MUSCLE 3.7, alignments were curated using Gblock settings, phylogeny was constructed using PhyML 3.0 aLRT with 1000 bootstraps, and the final tree was rendered using TreeView.

RT-PCR. For RT-PCR experiments, *P. naphthalenivorans* CJ2 was grown on MSB-N agar plates, and cells were collected from the resulting colonies. Total RNA was prepared using the RNeasy Mini kit (Qiagen) according to the manufacturer's instructions. The total RNA was treated with RNase-free DNase I (Qiagen) and RNase inhibitor (Takara) in RDD buffer (Qiagen) for 30 min at room temperature. The RNA preparation was cleaned via passage through the RNeasy spin column (Qiagen). Next, RT-PCRs were carried out using the One Step RT-PCR kit (Qiagen) based on the manufacturer's instructions, with nine primer pairs, RT1-F/R to RT9-F/R (Table 1), to amplify the intergenic regions in the third catabolic gene cluster. To confirm that cDNA synthesis occurred and that RNA preparation was free of genomic DNA, a negative control RT-PCR was performed with *Taq* polymerase (omitting the reverse transcriptase).

Real-time RT-PCR for gene expression analysis of the third catabolic gene cluster. For gene expression analysis of the third gene cluster in *nagR* knockout mutant CJN110 and wild-type CJ2, real-time RT-PCR was used. Cells of wild-type CJ2 and mutant CJN110 were grown on MSB-P agar for 8 days at 20 °C for their initial cell growth. The cell volume was split in two transferred onto MSB-P agar or MSB-P agar containing 1 mM salicylate as an inducer, respectively, and the agar plates were further incubated for 24 h at 20 °C for gene induction. Cells were collected from colonies on agar plates and were resuspended in PBS solution (137 mM NaCl, 8.1 mM Na_2HPO_4, 2.68 mM KCl, 1.47 mM KH_2PO_4, pH 7.2). The suspended cells were measured with a UV-Vis spectrophotometer (SynergyMX) at 600 nm and were harvested via centrifugation. Total RNA from the harvested cells was extracted using the RNeasy Mini kit (Qiagen) based on the manufacturer's instructions.

Real-time RT-PCR primer pairs, rt-orf2-F/rt-orf2-R and rt-nagI3-F/rt-nagI3-R, targeting *orf2* and *nagI3* were designed (Table 1). Real-time RT-PCR using the total RNA as template was performed by using the iScript One-Step RT-PCR kit with SYBR Green (Bio-Rad) in a C1000 Thermal Cycler (Bio-Rad) under the following conditions. Each 20 μl reaction mixture contained 0.3 μM rt-orf2-F/rt-orf2-R or rt-nagI3-F/rt-nagI3-R primers, 1 × SYBR Green RT-PCR mix, 1 × iScript reverse transcriptase and 100 ng total RNA template. The reverse transcriptase step for cDNA synthesis was 10 min at 50 °C. Cycling conditions for real-time PCR analysis were as follows: 95 °C for 5 min, 40 cycles at 95 °C for 30 s, 59 °C for 30 s and 72 °C for 30 s. Fluorescence data were acquired at the end of each annealing step. The relative expression values of the third gene cluster in *nagR* knockout mutant CJN110 (Jeon *et al.*, 2006) and wild-type CJ2 were calculated as percentages of the expression levels of wild-type strain CJ2 on MSB-P agar without salicylate.

Real-time RT-PCR to analyse the expression of three *nagI* genes in strain CJ2. To analyse transcriptional expression of three *nagI* genes in wild-type CJ2, real-time RT-PCR was used. Cells of wild-type CJ2 were grown on R2A, MSB-P and MSB-N agar plates. Cells were harvested from the resulting colonies, and total RNA was

extracted from cells as described above. Three specific primer pairs, rt-nagI1-F/rt-nagI1-R, rt-nagI2-F/rt-nagI2-R, and rt-nagI3-F/rt-nagI3-R, targeting *nagI1*, *nagI2* and *nagI3* were designed, respectively (Table 1), and specific amplification of their respective *nagI* genes was confirmed by sequencing their PCR products. Real-time RT-PCR using the total RNA extracts as template was carried out as described above. Real-time RT-PCR using the 16S rRNA gene primers, Bac1055YF and Bac1392R (Table 1; Ritalahti *et al.*, 2006), was used for the normalization of total RNA templates. Standard curves for the calculation of cDNA copy numbers were generated using the three *nagI* cloned plasmids pET21b*nagI1*, pET21b*nagI2* and pET21b*nagI3* (Table 1). The relative expression values of the *nagI* genes were calculated as percentages of the *nagI2* gene expression level on MSB-P agar.

Construction and growth of three individual *nagI* knockout mutants in strain CJ2. Plasmid pK19*mobsacB* (Schäfer *et al.*, 1994), containing *oriT* for conjugative mobilization and *sacB* (encoding the *Bacillus subtilis* levansucrase) as a counter-selectable marker, was used to construct three individual *nagI* knockout mutants in strain CJ2. Mutagenic plasmids pK19*mobsacB*Δ*nagI1*, pK19*mobsacB*Δ*nagI2* and pK19*mobsacB*Δ*nagI3* with 2 bp deletions in the *nagI* genes were constructed via a three-step round-PCR approach (Lee *et al.*, 2011) as shown in Fig. 1a. Briefly, *nagI* gene PCR amplicons with a 2 bp deletion and *Hin*dIII and *Eco*RI sites at the flanking regions were generated by PCR using two primer sets each, nagIx1F (*Hin*dIII)/nagIx1R and nagIx2F/nagIx2R (*Eco*RI) shown in Table 1, where 'x' indicates 1, 2 or 3 for *nagI1*, *nagI2* and *nagI3*, respectively, and then were digested with *Hin*dIII and *Eco*RI (New England Biolabs). The digested products were ligated into the *Hin*dIII–*Eco*RI site of pK19*mobsacB* to produce three versions of pK19*mobsacB*Δ*nagI*x. These three plasmids were introduced via electroporation into *E. coli* S17-1/*λpir* cells carrying the *tra* region of RP4 (Fig. 1a).

Then, to construct the *nagI* knockout mutants, pK19*mobsacB*△*nagI*x was mobilized from *E. coli* S17-1/*λpir* cells containing pK19*mobsacB*Δ*nagI*x (~5.0×10^8–1.0×10^9, grown overnight in LB broth with 20 μg kanamycin ml^{-1}) into rifampicin-resistant *P. naphthalenivorans* CJ2 cells (~1.0×10^9 to 2.0×10^9, grown overnight in MSB-P broth with 200 μg rifampicin ml^{-1}) through mating on R2A agar medium at 25 °C for 8–16 h as described by Hohnstock *et al.* (2000) and Park *et al.* (2007a). Three transconjugants (strains CJ2Δ*nagI*xkm^r) were selected from each of the R2A plates containing kanamycin (20 μg ml^{-1}) and rifampicin (200 μg ml^{-1}) at 25 °C. Three *nagI* gene knockout mutants with a 2 bp deletion were selected via growth of strains CJ2△*nagI*xkm^r in R2A broth containing 10 % sucrose and subsequent plating onto R2A agar medium (Fig. 1b). The Suc^r colonies obtained were replica plated onto R2A agar supplemented with kanamycin. The *nagI* gene knockout mutants with 2 bp deletion in Suc^r colonies, which lost their kanamycin resistance (Suc^r/Kan^s), were confirmed via colony PCR and sequencing (using outer primer pairs nagI11-F/nagI12-R, nagI21-F/nagI22R, and nagI31-F/nagI32-R; Table 1). For the growth tests, strains CJ2 and three *nagI* knockout mutants, CJ2Δ*nagI1*, CJ2Δ*nagI2* and CJ2Δ*nagI3*, were streaked onto R2A, MSB-P, MSB-N, MSB-S, MSB-G and MSB-H agar media and were incubated at 20 °C. Their growth rate was assessed based on the number of days of incubation required before the appearance of visible colonies on the agar plates.

GDO activity of wild-type CJ2 and *nagI* knockout mutants. To measure the GDO activity of wild-type strain CJ2 and three *nagI* knockout mutants, cells of wild-type CJ2 and those of CJ2Δ*nagI1*, CJ2Δ*nagI2* and CJ2Δ*nagI3* were grown on R2A agar for 4 days at 20 °C to establish initial cell growth. The grown cells were transferred onto MSB-N agar, and the agar plates were further incubated for 2 days at 20 °C. Cells were collected from the colonies on the agar plates and were resuspended in ice-cold 100 mM potassium phosphate buffer (pH 7.4). The cell densities were measured with a UV-Vis spectrophotometer at 600 nm. Crude cell extracts were prepared by disrupting resuspended cells using sonication in an ice-water bath for three periods of 30 s with 30 s intervals, after which cell debris was removed by centrifugation at 12 000 r.p.m. for 30 min at 4 °C, and 100 mM ferrous ammonium sulfate was added to a concentration of 0.1 mM Fe^{2+} in the enzyme solution to bring about enzyme stabilization. Protein concentration of the crude cell extracts was determined using the Bradford assay (Bradford, 1976). GDO activities of the crude cell extracts were assayed by measuring maleylpyruvate formation from gentisate, which could be easily measured via an increase in the absorbance at 330 nm with an UV-Vis spectrophotometer (Lack, 1959). The 700 μl reaction mixtures contained 10 mM gentisate and 500 μg crude proteins in 0.1 M potassium phosphate buffer (pH 7.4). The reactions were initiated by the addition of crude cell extracts. The relative GDO enzyme activities of wild-type CJ2 and the *nagI* knockout mutants CJ2Δ*nagI1*, CJ2Δ*nagI2* and CJ2Δ*nagI3* were calculated as percentages of the GDO activity of the wild-type strain CJ2.

Overexpression of *nagI* genes in *E. coli* and the biochemical analysis of GDO proteins. To overexpress GDO proteins in *E. coli*, three *nagI* genes were cloned into the expression vector pET21b (Novagen). Three *nagI* genes were amplified by PCR from wild-type CJ2 genomic DNA using *Pfu* DNA polymerase (Solgent) and primer pairs nagIx-F and nagIx-R, which contained *Nde*I and *Bam*HI restriction sites, respectively. The amplified PCR products were digested with *Nde*I and *Bam*HI and ligated into pET21b, resulting in the recombinant plasmids designated pET21b*nagI1*, pET21b*nagI2* and pET21b*nagI3*. *E. coli* BL21(DE3) cells were subsequently transformed with the pET21b*nagI1*, pET21b*nagI2* and pET21b*nagI3* constructs to produce GDO proteins NagI1, NagI2 and NagI3, respectively. The three *nagI* cloned plasmids were confirmed via DNA sequencing. The expression of GDO proteins and the preparation of crude cell extracts from transformed *E. coli* BL21(DE3) cells were performed according to a method described previously (Park *et al.*, 2007a).

Crude cell extracts containing the NagI2 and NagI3 proteins were applied to gel filtration and DEAE column chromatography sequentially (Amersham Biosciences). DEAE columns containing crude cell extracts were washed with binding buffer (30 mM sodium phosphate, pH 7.5), and proteins were released with elution buffer (30 mM sodium phosphate, 1 M NaCl, pH 7.5) at a flow rate of 5.0 ml min^{-1}. Fractions (2 ml) were collected and 100 mM ferrous ammonium sulfate was immediately added for an enzyme solution concentration of 0.1 mM Fe^{2+}. Native gel electrophoresis was performed in a gel containing 10 % acrylamide for the separating gel and 5 % acrylamide for the stacking gel. The subunit molecular masses were determined by SDS-PAGE and staining with Coomassie brilliant blue R-250.

The *in vitro* activities of the GDO enzymes were assayed by measuring maleylpyruvate formation from gentisate, which was assayed via an increase in absorbance at 330 nm. Activities were assayed in 1 ml reaction mixture containing 0.1 mM gentisate in 100 mM potassium phosphate buffer at pH 7.4, and the reactions were initiated by the addition of purified GDO enzymes. One enzyme unit was defined as the amount of enzyme that produced 1 μmol maleylpyruvate min^{-1} at 25 °C. The molar extinction coefficient of maleylpyruvate for the measurement of maleylpyruvate was taken as 10.8×10^3 M^{-1} cm^{-1} (Crawford *et al.*, 1975). Protein concentrations were determined by using the Bradford assay (Bradford, 1976), with BSA (Sigma) as the protein standard. A series of gentisate solutions, ranging from 2 to 1000 μM, were prepared for the determination of K_m and V_{max} values of the purified GDO enzymes. Spectrophotometric assays (A_{330}) were performed at 25 °C while maintaining a constant

Table 1. Bacterial strains, plasmids and PCR primers used in this study

Primer sequences are in 5′–3′ orientation and restriction enzyme sites (for future study) are underlined. o, Outer; F, forward; R, reverse; Km^r, kanamycin resistance; Ap^r, ampicillin resistance.

Strain, plasmid or primer	Description or sequence	Source or reference
Strains		
P. naphthalenivorans		
CJ2	Naphthalene degrader	Jeon *et al.* (2003)
CJ2Δ*nagI1*	*P. naphthalenivorans* CJ2 Δ*nagI1*	This study
CJ2Δ*nagI2*	*P. naphthalenivorans* CJ2 Δ*nagI2*	This study
CJ2Δ*nagI3*	*P. naphthalenivorans* CJ2 Δ*nagI3*	This study
CJN110	*P. naphthalenivorans* CJ2 Δ*nagR*::Km *lacZ*$^+$	Jeon *et al.* (2006)
E. coli		
S17-1/λ*pir*	Carries RK2 *tra* regulon and *pir*; host for *pir*-dependent plasmids	Kalogeraki & Winans (1997)
BL21(DE3)	F$^-$ *dcm ompT hsd*S(r_B^- m_B^-) *gal*λ BL21(DE3)	Novagen
Plasmids or vectors		
pK19*mobsacB*	Intergration vector; Km^r *ori*V_{Ec} *oriT sacB*	Schäfer *et al.* (1994)
pET21b	Expression vector, Ap^r, T7 tag, multi cloning site, His tag	Novagen
pET21b*nagI1*	*Nde*I–*Bam*HI-cut PCR fragment containing *nagI1* inserted into pET21b	This study
pET21b*nagI2*	*Nde*I–*Bam*HI-cut PCR fragment containing *nagI2* inserted into pET21b	This study
pET21b*nagI3*	*Nde*I–*Bam*HI-cut PCR fragment containing *nagI3* inserted into pET21b	This study
pK19*mobsacB*Δ*nagI1*	Carries deletion of the 2 bp fragment of *nagI1* into pK19*mobsacB*	This study
pK19*mobsacB*Δ*nagI2*	Carries deletion of the 2 bp fragment of *nagI2* into pK19*mobsacB*	This study
pK19*mobsacB*Δ*nagI3*	Carries deletion of the 2 bp fragment of *nagI3* into pK19*mobsacB*	This study
Primers		
RT1-F	CTGCTGCAAGACGAACTTGA	RT-PCR
RT1-R	GGTCCAGAATGGTCGTACTC	
RT2-F	AGTCTCCTTGACGCCATACA	
RT2-R	GGCGCGTTCAATCTCATGTT	
RT3-F	ACACCGTTCGGAAGTGACT	
RT3-R	CCATATGTTGAGGCTGATGC	
RT4-F	CATTACGCGGCTGTAGGTTA	
RT4-R	GAGCCGGCACGTATCAACAT	
RT5-F	TGACATGAACGCGGTGTTCG	
RT5-R	AACTCGATCTCACCGGCCTT	
RT6-F	AAGGCCGGTGAGATCGAGTT	
RT6-R	AGCTCGCAGTTCGCTGGAA	
RT7-F	TTGCTGGAGTTGGCAGATCG	
RT7-R	AACTCGTGGTCTGCCGAAC	
RT8-F	ACAAGGCTTACCCCATGCTG	
RT8-R	CTTGGTCTGCCCATTGACTT	
RT9-F	CAGAAGCCAATGAGTTGGGC	
RT9-R	TGTGATGATGGACGTGCTTC	
nagI1-F	AAACATATGGTGCCGCTGTGGCC	Protein expression, *Nde*I
nagI1-R	CTGGGATCCTCAGGCGCGGACTTGTA	Protein expression, *Bam*HI
nagI2-F	AAACATATGCAAGATCAAGACCGTTTTGACA	Protein expression, *Nde*I
nagI2-R	CTGGGATCCTCAGCCGCGGTTCTCGA	Protein expression, *Bam*HI
nagI3-F	AAACATATGGAAGACAACAATCGGTTTGACA	Protein expression, *Nde*I
nagI3-R	CTGGGATCCTCAGCCACGGTCCTCGTACA	Protein expression, *Bam*HI
nagI11-F	AAAAAAGCTTTTGCAAGCTGGCGAGCTGA	*nagI* knockout mutants, *Hin*dIII
nagI11-R	TGGCGCTTGGTTGAGGTGGTAGCTGGC	*nagI* knockout mutants
nagI12-F	CACCTCAACCAAGCGCCAGGCCGTCAA	
nagI12-R	CCGGAATTCCTGGTCGGCTGAGAGGTTCT	*nagI* knockout mutants, *Eco*RI
nagI21-F	AAAAAAGCTTACCGTGACGAACTGACGCAG	*nagI* knockout mutants, *Hin*dIII
nagI21-R	AGCACATCGCCAGATCACCGGCTGGTC	*nagI* knockout mutants
nagI22-F	GTGATCTGGCGATGTGCTGGATTTGCCT	

Table 1. cont.

Strain, plasmid or primer	Description or sequence	Source or reference
nagI22-R	CCGGAATTCTGTATCTGCCGAGCTGTTGC	*nagI* knockout mutants, *Eco*RI
nagI31-F	AAAAAAGCTTTCGCGACCAACTCACCCAAC	*nagI* knockout mutants, *Hin*dIII
nagI31-R	GCCCTGTCCATGTGACACTCTGGCTGG	*nagI* knockout mutants
nagI32-F	GTGTCACATGGACAGGGCTACAAGGCC	
nagI32-R	CCAIGGGTAGCGCAGCAT	*nagI* knockout mutants, *Eco*RI
rt-nagI1-F	GTATCTGGAGTGCGTGCA	Real-time RT-PCR
rt-nagI1-R	ACCGTCCACGGTGATTTGC	
rt-nagI2-F	TTCTGGCAA AGCCTATCCA	
rt-nagI2-R	AGCTGGACCGATTCGAGTTC	
rt-nagI3-F	TGCGTGAAGATGACCAGCTT	
rt-nagI3-R	GCCATCAGATAGACACGCTT	
rt-orf2-F	TAGAGAGATTCGTGAGCTGC	
rt-orf2-R	TCCACCCACCAGAGTCCAT	

enzyme concentration in 100 mM potassium phosphate buffer (pH 7.4). The initial reaction velocities were used in Lineweaver–Burk plots to determine K_m and V_{max} values.

RESULTS

Sequence analysis of the third catabolic gene cluster

Analysis of the recently completed genome of strain CJ2 (Yagi *et al.*, 2009) revealed that, in addition to the two previously described naphthalene catabolic gene clusters (*nagRAdGHAbAcAdBFCQEDJI1–orf1–tnpA* and *nagR2–XI2KL*; Jeon *et al.*, 2006), the genome has a third gene cluster related to the naphthalene catabolic pathway. The third catabolic gene cluster includes 13 open reading frames (Fig. 2a). The putative function, locus tag notation appearing at the USDOE IMG website (http://img.jgi.doe.gov), predicted product size and significant matches to the predicted gene product are summarized for each gene in Table 2.

Examination of the DNA sequence showed that the organization of the third catabolic gene cluster (containing naphthalene catabolic genes, partial *nagAa*, partial *nagH*, *nagAb* and *nagI3KL*) is similar to that of the single continuous naphthalene catabolic gene cluster in *Ralstonia* sp. U2 (Fuenmayor *et al.*, 1998) and to that of the first gene cluster of *P. naphthalenivorans* CJ2 (Jeon *et al.*, 2006). However, in the third catabolic gene cluster, the *nagR* regulator and *nagG* (salicylate-5-hydroxylase large oxygenase component) are absent, and six putative genes (*orf2*, *orf3′*, *orf3′*, *orf4′*, *orf4′* and *orf5*) unrelated to the naphthalene degradation pathway are inserted in place of the nine naphthalene catabolic genes *nagAc* to *nagJ* (Fig. 2a). Furthermore, the *nagAb*, *nagI3*, *nagK* and *nagL* genes of the third cluster are complete sequences, but *nagAa* and *nagH* are partial sequences (Table 2).

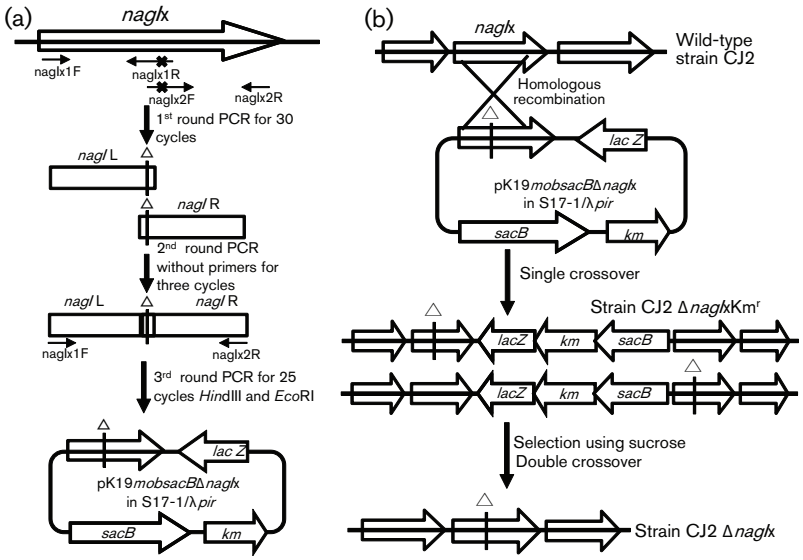

Fig. 1. Genetic manipulations used in this study. (a) Preparation of three pK19*mobsacB*Δ*naglx* constructs with 2 bp deletions in the *nagI* genes using a three-step PCR approach. (b) Construction of three *nagI* knockout mutants in *P. naphthalenivorans* CJ2 via Campbell-type homologous recombination with pK19*mobsacB*Δ*naglx*. Here, 'x' in *naglx* designates the targeted gene locus (e.g. *nagI1*, *nagI2* or *nagI3*).

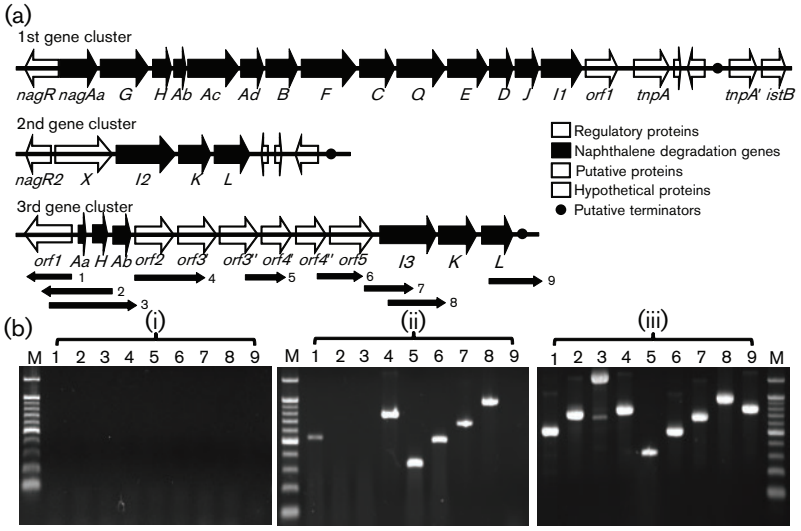

Fig. 2. Naphthalene catabolic operons in strain CJ2 and gene expression in the third gene cluster. (a) Organization of the three naphthalene catabolic gene clusters in *P. naphthalenivorans* CJ2. Arrows in the gene sequence indicate the direction of mRNA transcription in the RT-PCR assay. Numbered arrows below the third gene cluster show the location and direction of the nine primer pairs used in the RT-PCR assays. (b) RT-PCR analysis of expressed genes in the third gene cluster. The numbered lanes refer to the location of the RT-PCR fragments shown in (a). (i) PCR products from total RNA without reverse transcriptase; (ii) RT-PCR products from total RNA; (iii) PCR products from genomic DNA. M, Molecular size marker (100 bp ladder).

Examination of the DNA sequences between *orf1* and partial *nagAa* revealed that the promoter region of the third catabolic gene cluster was identical to that of the first cluster and had a high degree of identity with that of *Ralstonia* sp. U2 (Supplementary Fig. S2a, available with the online version of this paper; Jeon *et al.*, 2006; Jones *et al.*, 2003). Also, upstream of this promoter, between bases −76 and −59, is a symmetrical dyad motif $TTCAN_6TGAT$ (Supplementary Fig. S2b) that is characteristic of the LysR family, which has been identified as being important for NahR function (Jones *et al.*, 2003; Schell & Wender, 1986). Although the third catabolic gene cluster also has a transcriptional promoter matching that of *Ralstonia* sp. U2 and the first cluster of strain CJ2, the divergently transcribed protein, ORF1 (positioned analogously to *nagR*; Fig. 2) is annotated as a putative IstB-like ATP-binding protein (Table 2), not a LysR-type regulator.

Expression analyses of the third catabolic gene cluster in wild-type CJ2

For transcriptional analysis of the third gene cluster, RT-PCR was performed using purified mRNA from strain CJ2 cells grown on MSB agar with naphthalene vapour. Our primer sets covered almost all of the ORFs of the third gene cluster (Fig. 2a). Because the organization of the third gene cluster closely resembles that of *Ralstonia* sp. U2 and the first gene cluster, we presumed that expression began at *nagAa* and extended into *nagL*. The amplified products were analysed by agarose gel electrophoresis (Fig. 2b), and RT-PCR products were not observed when reverse transcriptase was omitted from the reaction mixture (Fig. 2b, i). The presence of amplified DNA fragments obtained with each primer pair suggests that contiguous genes including six putative genes unrelated to the naphthalene biodegradation pathway in the third gene cluster form a polycistronic operon that is transcribed as a single message (Fig. 2b, ii) like that of the first gene cluster (Jeon *et al.*, 2006). This indicates that the third catabolic gene cluster is transcribed from a promoter located between *orf1* and partial *nagAa*. The data in Fig. 2(b), ii, show that transcription did not extend beyond fragment 9, indicating that there is a termination sequence at the end of the third gene cluster. RT-PCR analysis also showed that the *orf1* gene is transcribed in the opposite direction to that of another promoter located between *orf1* and a partial *nagAa* like *nagR* of the first gene cluster.

Confirming the regulatory control of *nagR* on the third catabolic gene cluster

Because LysR-type regulatory genes associated with naphthalene catabolism are known to be induced by salicylate (Jones *et al.*, 2003) and the promoter of the third catabolic gene cluster was identical to that of the first gene cluster (see above and Supplementary Fig. S2), real-time RT-PCR analysis was conducted to investigate whether the transcription of the third gene cluster was controlled by the *nagR* regulator of the first gene cluster. The assay examined the influence of salicylate addition on expression of two genes within the third gene cluster (*orf2* and *nagI3*) in both wild-type and the *nagR* knockout mutant strain CJN110 (Fig. 3). The results clearly show that gene expression was induced by salicylate in wild-type strain CJ2 but not in the mutant, verifying that the *nagR* regulator controls the third catabolic gene cluster.

Real-time RT-PCR of *nagI1*, *nagI2* and *nagI3* expression

To investigate and compare the contributions of three *nagI* genes to the cellular pool of transcripts in the wild-type host, real-time RT-PCR analysis was performed. Fig. 4 shows that the transcriptional levels of the *nagI1* and *nagI3* genes were induced in cells grown on naphthalene (MSB-N) and were higher than transcription of the *nagI2* gene.

Table 2. *P. naphthalenivorans* strain CJ2 genes and gene products

Gene	Putative function	JGI locus*	Length (aa)†	Most similar gene product(s) (species, accession no.)	Percentage identity (aa)
orf1	IstB-like ATP-binding protein	2464	255 (−)	IstB-like ATP-binding protein (*Polaromonas* sp. JS666, YP_552097)IstB domain protein ATP-binding protein (*Thiomonas intermedia* K12, YP_003644160)	78 77
Partial *nagAa*	Partial ferredoxin reductase	–	89 (+)	Oxidoreductase FAD-binding subunit (*P. naphthalenivorans* CJ2, YP_982714) Ferredoxin reductase (*Burkholderia* sp. C3, ACT53245)	98 70
Partial *nagH*	Salicylate-5-hydroxylase small oxygenase component	2462	109 (+)	Salicylate-5-hydroxylase small oxygenase component (*Burkholderia* sp. C3, ACT53247) Salicylate 5-hydroxylase small oxygenase component (*Ralstonia* sp. U2, AAD12608)	86 85
nagAb	Ferredoxin	2461	104 (+)	Ferredoxin NBDF (*Comamonas* sp. JS765, AAL76201) Ferredoxin (*Ralstonia* sp. U2, AAD12609)	78 75
orf2	Extracellular ligand-binding receptor	2460	390 (+)	Extracellular ligand-binding receptor (*Polaromonas* sp. JS666, YP_547838) Extracellular ligand-binding receptor (*L. cholodnii* SP-6, YP_001789831)	70 67
orf3′	Inner-membrane translocator	2459	308 (+)	Inner-membrane translocator (*L. cholodnii* SP-6, YP_001789830) Inner-membrane translocator (*Polaromonas* sp. JS666, YP_547839)	77 74
orf3′	Inner-membrane translocator	2458	318 (+)	Putative membrane-spanning protein (*Achromobacter xylosoxidans* strain A8, YP_195867) Inner-membrane translocator (*L. cholodnii* SP-6, YP_001789829)	62 60
orf4′	ABC transporter related	2457	250 (+)	ABC transporter-related (*Polaromonas* sp. JS666, YP_547841) ABC transporter-related protein (*L. cholodnii* SP-6, YP_001789828)	72 70
orf4′	ABC transporter related	2456	237 (+)	ABC transporter-related (*Polaromonas* sp. JS666, YP_547842) ABC transporter-related protein (*L. cholodnii* SP-6, YP_001789827)	70 69
orf5	Hemerythrin HHE cation binding region	2455	196 (+)	Hemerythrin HHE cation-binding domain protein (*Thauera* sp. MZ1T, YP_002355175) Hemerythrin HHE cation-binding domain protein (*L. cholodnii* SP-6, YP_001790243)	38 37
nagI3	Gentisate 1,2-dioxygenase	2454	351 (+)	Gentisate 1,2-dioxygenase (*Ralstonia* sp. U2, AAD12619) Cupin 2 domain-containing protein (*L. cholodnii* SP-6, YP_001792689)	74 74
nagK	Fumarylpyruvate hydrolase	2453	233 (+)	Fumarylpyruvate hydrolase (*Ralstonia* sp. U2, AAD12620) Fumarylacetoacetate (FAA) hydrolase (*Delftia acidovorans* SPH-1, YP_001564718)	83 82
nagL	Maleylpyruvate isomerase	2452	212 (+)	Maleylacetoacetate isomerase (*L. cholodnii* SP-6, YP_001792691) Maleylacetoacetate isomerase (*Comamonas testosteroni* S44, ZP_07045767)	73 72
None	Putative terminator				

*The USDOE IMG JGI locus nomenclature is used, e.g. '2464' indicates Pnap_2464.
†The orientation of the coding strand is indicated in parentheses.

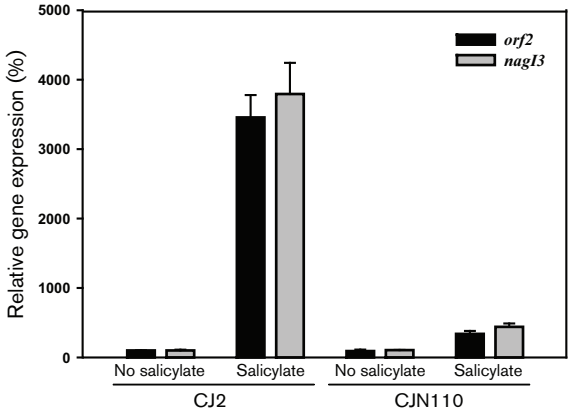

Fig. 3. Quantification of mRNA expression for *orf2* and *nagI3* in the third naphthalene catabolic gene cluster in wild-type strain CJ2 and its *nagR* regulatory mutant, strain CJN110, grown on MSB-P with and without 1 mM salicylate. The abundance of mRNA transcripts was measured by using real-time RT-PCR, and the relative expression ratios (*y*-axis) were calculated as percentages of the expression level of wild-type strain CJ2 on MSB-P agar without salicylate. Values shown are mean ± SD for three independent cultures.

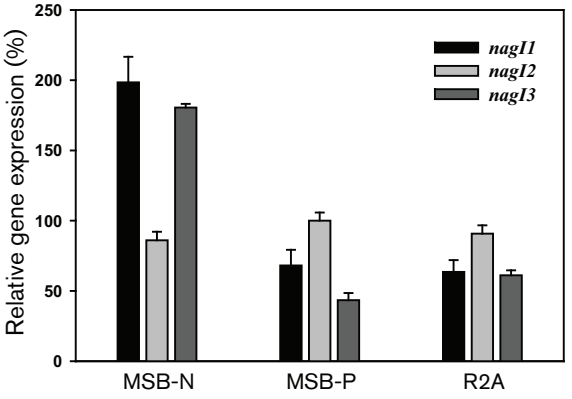

Fig. 4. Real-time RT-PCR analysis of the mRNA transcripts of the three *nagI* alleles extracted from wild-type strain CJ2 cells grown on MSB-N, MSB-P and R2A agar. The relative expression ratios (*y*-axis) were calculated as percentages of the *nagI2* gene expression level of strain CJ2 on MSB-P agar. Values shown are mean ± SD for three independent cultures.

An absence of induction of *nagI2* by naphthalene is consistent with previous results (Jeon *et al.*, 2006).

Constructing three individual *nagI* knockout mutants in strain CJ2 and phenotype of the three *nagI* alleles: analysis of growth

Though the genome of CJ2 carries three GDO-encoding genes, the relative influence of each allele has not been assessed. Three *nagI* knockout mutants of strain CJ2 (designated CJ2Δ*nagI1*, CJ2Δ*nagI2* and CJ2Δ*nagI3*) were constructed via 2 bp unmarked gene deletion mutagenesis (Fig. 1). Sequence analysis of the three deleted *nagI* genes showed that all three *nagI* knockout mutants still contained stop codons within their *nagI* genes: thus we inferred that 2 bp gene deletions did not influence protein expression downstream of the gene clusters. To further investigate the relative contributions of three *nagI* genes to the naphthalene degradation in strain CJ2, growth of the three *nagI* knockout mutants of *P. naphthalenivorans* was semiquantitatively assessed. Because many of the test strains did not grow significantly in broth culture, growth rates (wild-type and mutants) were evaluated based on the duration of incubation prior to colony formation (visible by eye) by cells spread onto R2A, MSB-P, MSB-N, MSB-S, MSB-G and MSB-H agar media. As expected, the results of the growth tests showed that none of the three *nagI* knockout mutants had impaired growth on the complex R2A medium (Table 3). However, impaired growth on aromatic substrates was dramatic. Importantly, although transcription levels of *nagI1* and *nagI3* genes were similar (Fig. 4), the *nagI3* knockout mutant, CJ2Δ*nagI3*, showed the most severe growth defect on the aromatic compounds; it is notable that the naphthalene-grown colonies (not those of the other two mutants) turned pale red, possibly due to the accumulation of gentisate. Surprisingly, the mutations influenced the ability of strains to grow on pyruvate (MSB-P), perhaps indicating that GDO might influence the global cell growth of strain CJ2. Also, even wild-type strain CJ2 did not grow on MSB-S, although salicylate is an inducer and intermediate in naphthalene biodegradation by strain CJ2. We speculate that lack of growth on salicylate (0.3 %) may have been caused by the compound's toxicity (e.g. Kamaya *et al.*, 2005) at such high concentration.

Table 3. Growth rate of wild-type strain CJ2 and three *nagI* knockout mutants on six different agar media

Values in the table indicate no. of days of incubation required to achieve a colony diameter of 2 mm. No colonies were found for any strain grown on MSB-S.

Medium	Wild-type CJ2	CJ2Δ*nagI1*	CJ2Δ*nagI2*	CJ2Δ*nagI3*
R2A	2	3	3	3
MSB-G	7	9	9	–
MSB-N	5	7	7	10*
MSB-H	5	7	7	–
MSB-P	5	7	7	9

*Colony diameter <1 mm.

GDO activity assays in wild-type CJ2 and *nagI* knockout mutants

To begin exploring the mechanistic basis for the influence of the three *nagI* alleles on the growth of strain CJ2 (Table 3), we examined the translated proteins. Because the *nagI3* knockout mutant grew poorly on MSB-N (Table 3), cells of wild-type CJ2 and the three *nagI* knockout mutants were first cultivated on R2A agar. Then the cells were further incubated on MSB-N agar for 2 days at 20 °C to achieve GDO protein induction. Crude cell extracts were prepared by disrupting resuspended cells via sonication, and their GDO enzyme activities were measured. The *nagI1* and *nagI2* knockout mutants (each retaining intact *nagI3* genes) displayed approximately 40 % reduction in GDO enzyme activity compared with that of wild-type CJ2 (data not shown). However, the *nagI3* knockout (with intact *nagI1* and *nagI2* genes) lost >95 % of its GDO capability. Clearly *nagI3* is crucial for GDO expression and dominates the activity of this enzyme for cells grown on the MSB-N medium. To our knowledge, gene regulation, protease activity and protein content were constant in all assays assessing GDO activity in all tested strains; thus, the degree of reduced enzyme activity for individual mutants was somewhat surprising. Factors governing quantitative reductions in GDO activity among the three mutants are uncertain and warrant future investigation.

Overexpression in *E. coli*, purification and characterization of GDO proteins (NagI2 and NagI3)

Each of the three *nagI* genes from *P. naphthalenivorans* CJ2 was cloned into *E. coli* BL21(DE3) under transcriptional regulation of the IPTG-inducible T7 promoter in the expression vector pET21b to produce GDO proteins. The pET21b*nagI2* and pET21b*nagI3* clones were successfully overexpressed in their native state in *E. coli*. Crude cell extracts of IPTG-induced *E. coli* BL21(DE3) harbouring the pET21b*nagI2* and pET21b*nagI3* constructs were separated using SDS-PAGE, and bands corresponding to overexpressed proteins with apparent molecular masses of the expected sizes, 39 and 38.7 kDa for NagI2 and NagI3, respectively, were observed (Supplementary Table S1 and Supplementary Fig. S3, available with the online version of this paper). However, overexpression of the *nagI1* gene failed in *E. coli*. The transformation efficiency of the pET21b*nagI1* ligation product was very low ($\sim 10^{-3}$ efficiency compared with that of the *nagI2* and *nagI3* genes), and the pET21b*nagI1* transformant always contained an altered *E. coli* rare codon [CTC (Leu)→CTA (Leu)] and did not produce a GDO protein.

Recombinant NagI2 and NagI3 proteins were purified from extracts of IPTG-induced *E. coli* BL21(DE3) cells harbouring pET21b*nagI2* and pET21b*nagI3* via gel filtration and DEAE column chromatography, respectively. The eluted fractions that showed GDO activity were separated using SDS-PAGE, and predominantly single protein bands were obtained (Supplementary Fig. S3). The predicted molecular masses of NagI2 and NagI3 were determined to be about 80 kDa using native-PAGE, which strongly suggests that these are holoenzymes (likely dimeric proteins; Supplementary Table S1). The pIs of NagI2 and NagI3 were estimated to be ~5.4, while the pI of NagI1 was estimated to be 6.8 based on a theoretical calculation by the Lasergene software. Although the pIs of NagI2 and NagI3 were similar to those of other GDO enzymes, the pI value (6.82) of NagI1 was quite different.

Because the purified GDO proteins NagI2 and NagI3 were highly unstable in phosphate buffer without ferrous ions (as reported previously; Crawford *et al.*, 1975; Feng *et al.*, 1999), 100 mM ferrous ammonium sulfate was added immediately after DEAE column purification to reach 0.1 mM Fe^{2+}, thereby stabilizing the enzymes. The optimal activity of recombinant NagI2 and NagI3 was observed at 50 and 65 °C, respectively, but they were found to retain relatively high activities (74.1 and 72.4 % of maximum activities for NagI2 and NagI3, respectively) even at 25 °C. The pH dependence of the purified GDO enzymes was investigated by assaying for enzyme activity in 0.1 M potassium phosphate buffers ranging in pH from 5.0 to 9.4; NagI2 and NagI3 enzymes exhibited maximum activities at pH 8.0 and pH 7.8, respectively.

Spectrophotometric assays were performed in 0.1 M phosphate buffer at pH 7.4 using the same amount of each enzyme to determine K_m values. The gentisate concentrations varied from 2 to 1000 μM. The purified recombinant GDO enzymes NagI2 and NagI3 displayed typical Michaelis–Menten kinetics, and 1/V versus 1/[S] double reciprocal (Lineweaver–Burk) plots of enzyme activity yielded apparent K_m values of 31 and 10 μM for the NagI2 and NagI3 enzymes, respectively (Supplementary Table S1), indicating that NagI3 has a higher affinity for gentisate than does NagI2. This high affinity, in combination with the high level of expression (Fig. 4), provides biochemical and physiological clues explaining the dominant role of the *nagI3* allele during naphthalene degradation in *P. naphthalenivorans* CJ2 (Table 3; see Discussion).

Bioinformatic evidence for *nagI* gene transfer and duplication

G+C content and oligonucleotide frequency, together with phylogenetic information and codon adaptation for functional genes, are widely accepted criteria for amassing evidence of horizontal gene transfer (e.g. Teeling *et al.*, 2004). The region of the chromosome of strain CJ2 containing the third gene cluster (as well as the first cluster) shows a major drop in G+C content compared with background trends within the overall genome (Yagi *et al.*, 2009). These observations are consistent with the hypothesis that the naphthalene catabolic genes may have been obtained from outside the genome.

An additional clue providing insight into the origin of the third naphthalene catabolic operon of strain CJ2 is evidence from gene-association analysis. The basic organization of genes unrelated to naphthalene degradation in the third gene cluster is shared in an operon present in both *Polaromonas* sp. JS666 (Mattes *et al.*, 2008) and *Leptothrix cholodnii* SP-6. These bacteria do not grow on naphthalene but do metabolize salicylate via the gentisate pathway. The gentisate operons in *Polaromonas* sp. JS666, *L. cholodnii* SP-6 and strain CJ2 feature five putative genes (*orf2*, *orf3′*, *orf3′*, *orf4′* and *orf4′*) downstream of the *nagAb* gene (Fig. 5). This commonality suggests that the three genomes share a common ancestor and/or that the cluster was transferred horizontally into the *P. naphthalenivorans* lineage, followed by gene rearrangement. The G+C contents of the five non-catabolic genes (*orf2*, *orf3′*, *orf3′*, *orf4′* and *orf4′*) common to the three gene clusters shown in Fig. 5 are far from identical [strain CJ2 (58.5 %), *Polaromonas* sp. JS666 (65.0 %) and *L. cholodnii* SP-6 (64.6 %)] and their amino acid sequence identities were low (Table 2), suggesting divergence in the ancestry since transfer. A dendrogram comparing the sequences of the three *nagI* alleles of strain CJ2 along with those of *Polaromonas* sp. JS666 and *L. cholodnii* SP-6 (Supplementary Fig. S4, available with the online version of this paper) confirms this observation: *nagI3* in strain CJ2 is more closely related to its own alleles than it is to the *nagI* alleles carried by *Polaromonas* sp. JS666 and *L. cholodnii*. Thus, if *nagI3* was obtained by strain CJ2 via horizontal transfer, the introduction likely entered early into the lineage shared by *L. cholodnii* and *Polaromonas* sp. JS666, allowing significant divergences within the ancestry to occur. Duplication within the lineage of strain CJ2 appears to be the most likely explanation for the presence of three *nagI* alleles. Clearly, the evolutionary developmental pathway is difficult to discern.

DISCUSSION

Redundancy within the genetic repertoire of bacteria has far-ranging implications for microbial biology. The gene duplication and amplification process is one of three broadly recognized mechanisms of genetic adaptation in unicellular organisms (Andersson & Hughes, 2009). Alteration in gene dosage (copy number) can be elicited in bacteria subjected to stresses ranging from nutrient limitation to antibiotic exposure (Andersson & Hughes, 2009; Gevers *et al.*, 2004). For example, carbon source starvation has led to the amplification of dioxygenase genes in bacteria carrying out catabolism of benzoate (Reams & Neidle, 2003, 2004), octane and nonanol (McBeth & Shapiro, 1984) and chlorinated organic acids (Ghosal & You, 1988; Ogawa & Miyashita, 1995; Rangnekar, 1988). Though the presence of multicopy dioxygenase genes has been noted by previous investigators (e.g. Chain *et al.* 2006; Johnson & Olsen, 1997; Maeda *et al.*, 1995), the physiological contributions of individual functional oxygenase alleles to their bacterial host have not been well explored.

In our prior study (Jeon *et al.*, 2006), we described the organization of naphthalene catabolic genes of *P. naphthalenivorans* CJ2 into two gene clusters, as shown in Fig. 2. Recent whole genome sequencing (Yagi *et al.*, 2009) revealed that there is another catabolic gene cluster related to naphthalene degradation in the CJ2 genome which is separated by approximately 8 kb on the chromosome from the first naphthalene catabolic gene cluster. In this investigation, we focused on a third naphthalene catabolic gene cluster and the role of the previously undescribed gentisate dioxygenase gene (*nagI3*) in the biology of strain CJ2.

Sequence and transcriptional analyses showed that the third cluster is controlled by *nagR* and is polycistronic, and its promoter region is identical to that of the first cluster (Figs 2 and 3, Supplementary Fig. S2). Perhaps predictably, the *nagI3* and nag*I1* transcripts were induced to similar levels in wild-type strain CJ2 cells grown on naphthalene, and these expression levels were higher than that of the *nagI2* gene (Fig. 4). Surprisingly, growth tests of knockout mutants of all three *nagI* alleles showed that, although all three *nagI* genes in *P. naphthalenivorans* CJ2 influenced naphthalene degradation, the *nagI3* gene played the most significant role (Table 3). Pursuit of an explanation of the role of *nagI3* led to two types of inquiry: GDO enzyme activity assays (examining cells from deletion mutants) and biochemical characterization of GDO enzymes. The GDO activity assays proved the dominant role of NagI3 over the other alleles in naphthalene-grown cells. Though we planned to biochemically characterize all three NagI enzymes, only *nagI2* and *nagI3* clones were successfully overexpressed. The K_m value for NagI3 (10 μM; Supplementary Table S1) was approximately three times

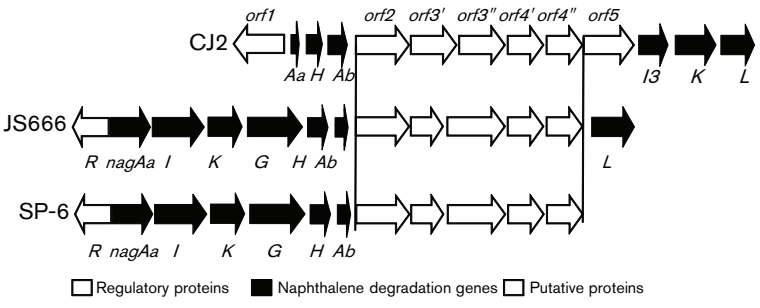

Fig. 5. Comparison of the genetic organization of the third catabolic gene cluster of strain CJ2 with corresponding gene clusters of two other bacteria, *Polaromonas* sp. JS666 and *L. cholodnii* SP-6. *R*, LysR family regulator; *orf2*, extracellular ligand-binding receptor; *orf3*, inner-membrane translocator; *orf4*, ABC transporter-related protein; *orf5*, hemerythrin HHE cation binding region.

lower than for NagI2; this may explain why *nagI3* is so influential in strain CJ2.

Our hypothesis is that the low K_m value for NagI3 leads to efficient processing of the toxic intermediary metabolite gentisate. Bacteria that metabolize aromatic hydrocarbons face the challenge of acquiring carbon and energy from compounds that are potentially toxic (Ramos *et al.*, 2002; Sikkema *et al.*, 1995). The mechanisms of toxicity are generally believed to be disruption of biological membranes (Teeling *et al.*, 2004) and the production of toxic metabolites (e.g. Park *et al.*, 2004; Vaillancourt *et al.*, 2002, 2006). Some metabolites of aromatic compounds, such as catechols and quinones, can be more toxic than the parent compounds due to an increase in solubility, production of reactive oxygen species or adduct formation with DNA and proteins (Penning *et al.*, 1999; Schweigert *et al.*, 2001). In a prior study of *P. naphthalenivorans* CJ2, Pumphrey & Madsen (2007) documented both direct inhibition of growth by naphthalene at high concentrations and the accumulation of toxic oxidation products derived from 1,2-naphthoquinone, which resulted in a complete loss of cell viability. Here we discovered a severe growth defect (Table 3) and nearly complete loss of GDO activity in a NagI3 deletion mutant. The severity of the *nagI3* deletion seems most simply explained not by loss of a redundant function (the *nagI3* deletion cells carry GDO as intact copies of *nagI1* and *nagI2*), but instead by accumulation of a toxic metabolite gentisate. This explanation is consistent with accumulation of red pigment in the CJ2Δ*nagI3* colonies (see Results) and prior reports of gentisate toxicity in bacteria (e.g. Reber, 1973). Data showing that the K_m value for NagI1 is greater than that for NagI3 would strengthen this hypothesis (but such characterization was impossible here because NagI1 expression was unsuccessful).

Yagi *et al.* (2009) presented evidence (gene synteny; Karlin signature difference; presence of conjugal transfer genes, tRNA genes, transposes, integrases, recombinases) for insertion of a 96 kb region into the chromosome of strain CJ2 that conferred naphthalene catabolism. The third naphthalene catabolic gene cluster discussed here falls within this 96 kb region. Here, gene-association analysis (Fig. 5) provided a potentially insightful clue as to the origin of the third *nagI* copy in strain CJ2: horizontal gene transfer (HGT) may have delivered a shared cluster of genes (including *nagI*) to strain CJ2, *Polaromonas* sp. JS666 and *L. cholodnii* strain SP-6. However, this hypothesis would be best supported if *nagI3* in strain CJ2 was more closely related to *nagI* in the other bacteria than to those in its own chromosome. Phylogenetic analysis of the NagI protein sequences (Supplementary Fig. S4) did not support the HGT hypothesis because NagI3 grouped more closely to its homologues within strain CJ2. The detailed evolutionary history of naphthalene and gentisate catabolism in strain CJ2 (see patterns and contrasts in gene order, operon content and regulation; Fig. 2) is probably obscured by many types of genetic change.

Bacteria in nature constantly confront stressful conditions: selection pressures such as starvation, desiccation, and non-optimal pH and temperature. There is evidence that starvation is accompanied by increased transposition of mobile elements and by other types of genetic changes (Andersson & Hughes, 2009; Kulakov *et al.*, 2005). *Polaromonas* species are frequently associated with cold and/or nutrient-poor habitats such as glacial melt waters and oligotrophic freshwater lakes (Irgens *et al.*, 1996; Liu *et al.*, 2006; Loy *et al.*, 2005). The oligotrophic heritage of strain CJ2 is consistent with its inability to grow on rich media (Jeon *et al.*, 2004). We speculate that the development of strain CJ2 may have featured relatively high rates of genetic exchange, rearrangement and duplication, leading to the three copies of the *nagI* gene in strain CJ2. Future work will aim to mechanistically explore both molecular (e.g. the active site of NagI3) and ecological (e.g. fitness traits) details of the biology of strain CJ2.

ACKNOWLEDGEMENTS

These efforts were supported by grants from the MOST/KOSEF to the 21C Frontier Microbial Genomics and Application Center Program (grant no. MG05-0104-4-0) and the Technology Development Program for Agriculture and Forestry (TDPAF) of the Ministry for Agriculture, Forestry and Fisheries. E. L. M. was supported by National Science Foundation grant no. DEB-0841999.

REFERENCES

Adams, M. A., Singh, V. K., Keller, B. O. & Jia, Z. (2006). Structural and biochemical characterization of gentisate 1,2-dioxygenase from *Escherichia coli* O157:H7. *Mol Microbiol* **61**, 1469–1484.

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

Andersson, D. I. & Hughes, D. (2009). Gene amplification and adaptive evolution in bacteria. *Annu Rev Genet* **43**, 167–195.

Bayly, R. C. & Barbour, M. G. (1984). The degradation of aromatic compounds by the meta and gentisate pathway: Biochemistry and regulation. In *Microbial Degradation of Aromatic Compounds*, pp. 253–294. Edited by D. T. Gibson. New York: Marcel Dekker.

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* **72**, 248–254.

Chain, P. S. G., Denef, V. J., Konstantinidis, K. T., Vergez, L. M., Agulló, L., Reyes, V. L., Hauser, L., Córdova, M., Gómez, L. & other authors (2006). *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. *Proc Natl Acad Sci U S A* **103**, 15280–15287.

Chapman, P. J. (1972). An outline of reaction sequences used for the bacterial degradation of phenolic compounds. In *Degradation of Synthetic Organic Molecules in the Biosphere*, pp. 17–55. Edited by P. J. Chapman & S. Dagley. Washington, DC: National Academy of Sciences.

Chen, J., Li, W., Wang, M., Zhu, G., Liu, D., Sun, F., Hao, N., Li, X., Rao, Z. & Zhang, X. C. (2008). Crystal structure and mutagenic analysis of GDOsp, a gentisate 1,2-dioxygenase from *Silicibacter pomeroyi*. *Protein Sci* **17**, 1362–1373.

Crawford, R. L., Hutton, S. W. & Chapman, P. J. (1975). Purification and properties of gentisate 1,2-dioxygenase from *Moraxella osloensis*. *J Bacteriol* **121**, 794–799.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.-F., Guindon, S., Lefort, V. & other authors (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36** (Web Server,), W465–W469.

Feng, Y., Khoo, H. E. & Poh, C. L. (1999). Purification and characterization of gentisate 1,2-dioxygenases from *Pseudomonas alcaligenes* NCIB 9867 and *Pseudomonas putida* NCIB 9869. *Appl Environ Microbiol* **65**, 946–950.

Fuenmayor, S. L., Wild, M., Boyes, A. L. & Williams, P. A. (1998). A gene cluster encoding steps in conversion of naphthalene to gentisate in *Pseudomonas* sp. strain U2. *J Bacteriol* **180**, 2522–2530.

Gevers, D., Vandepoele, K., Simillion, C. & Van de Peer, Y. (2004). Gene duplication and biased functional retention of paralogs in bacterial genomes. *Trends Microbiol* **12**, 148–154.

Ghosal, D. & You, I.-S. (1988). Gene duplication in haloaromatic degradative plasmids pJP4 and pJP2. *Can J Microbiol* **34**, 709–715.

Grund, E., Denecke, B. & Eichenlaub, R. (1992). Naphthalene degradation via salicylate and gentisate by *Rhodococcus* sp. strain B4. *Appl Environ Microbiol* **58**, 1874–1877.

Hirano, S., Morikawa, M., Takano, K., Imanaka, T. & Kanaya, S. (2007). Gentisate 1,2-dioxygenase from *Xanthobacter polyaromaticivorans* 127W. *Biosci Biotechnol Biochem* **71**, 192–199.

Hohnstock, A. M., Stuart-Keil, K. G., Kull, E. E. & Madsen, E. L. (2000). Naphthalene and donor cell density influence field conjugation of naphthalene catabolism plasmids. *Appl Environ Microbiol* **66**, 3088–3092.

Irgens, R. L., Gosink, J. J. & Staley, J. T. (1996). *Polaromonas vacuolata* gen. nov., sp. nov., a psychrophilic, marine, gas vacuolate bacterium from Antarctica. *Int J Syst Bacteriol* **46**, 822–826.

Jeon, C. O., Park, W., Padmanabhan, P., DeRito, C., Snape, J. R. & Madsen, E. L. (2003). Discovery of a bacterium, with distinctive dioxygenase, that is responsible for *in situ* biodegradation in contaminated sediment. *Proc Natl Acad Sci U S A* **100**, 13591–13596.

Jeon, C. O., Park, W., Ghiorse, W. C. & Madsen, E. L. (2004). *Polaromonas naphthalenivorans* sp. nov., a naphthalene-degrading bacterium from naphthalene-contaminated sediment. *Int J Syst Evol Microbiol* **54**, 93–97.

Jeon, C. O., Park, M., Ro, H.-S., Park, W. & Madsen, E. L. (2006). The naphthalene catabolic (*nag*) genes of *Polaromonas naphthalenivorans* CJ2: evolutionary implications for two gene clusters and novel regulatory control. *Appl Environ Microbiol* **72**, 1086–1095.

Johnson, G. R. & Olsen, R. H. (1997). Multiple pathways for toluene degradation in *Burkholderia* sp. strain JS150. *Appl Environ Microbiol* **63**, 4047–4052.

Jones, R. M., Britt-Compton, B. & Williams, P. A. (2003). The naphthalene catabolic (*nag*) genes of *Ralstonia* sp. strain U2 are an operon that is regulated by NagR, a LysR-type transcriptional regulator. *J Bacteriol* **185**, 5847–5853.

Kalogeraki, V. S. & Winans, S. C. (1997). Suicide plasmids containing promoterless reporter genes can simultaneously disrupt and create fusions to target genes of diverse bacteria. *Gene* **188**, 69–75.

Kamaya, Y., Fukaya, Y. & Suzuki, K. (2005). Acute toxicity of benzoic acids to the crustacean *Daphnia magna*. *Chemosphere* **59**, 255–261.

Kivisaar, M. (2009). Degradation of nitroaromatic compounds: a model to study evolution of metabolic pathways. *Mol Microbiol* **74**, 777–781.

Kulakov, L. A., Chen, S., Allen, C. C. R. & Larkin, M. J. (2005). Web-type evolution of *rhodococcus* gene clusters associated with utilization of naphthalene. *Appl Environ Microbiol* **71**, 1754–1764.

Lack, L. (1959). The enzymic oxidation of gentisic acid. *Biochim Biophys Acta* **34**, 117–123.

Lee, S. H., Kim, J. M., Lee, H. J. & Jeon, C. O. (2011). Screening of promoters from rhizosphere metagenomic DNA using a promoter-trap vector and flow cytometric cell sorting. *J Basic Microbiol* **51**, 52–60.

Liu, Y., Yao, T., Jiao, N., Kang, S., Zeng, Y. & Huang, S. (2006). Microbial community structure in moraine lakes and glacial meltwaters, Mount Everest. *FEMS Microbiol Lett* **265**, 98–105.

Loy, A., Beisker, W. & Meier, H. (2005). Diversity of bacteria growing in natural mineral water after bottling. *Appl Environ Microbiol* **71**, 3624–3632.

Maeda, M., Chung, S.-Y., Song, E. & Kudo, T. (1995). Multiple genes encoding 2,3-dihydroxybiphenyl 1,2-dioxygenase in the Gram-positive polychlorinated biphenyl-degrading bacterium *Rhodococcus erythropolis* TA421, isolated from a termite ecosystem. *Appl Environ Microbiol* **61**, 549–555.

Mattes, T. E., Alexander, A. K., Richardson, P. M., Munk, A. C., Han, C. S., Stothard, P. & Coleman, N. V. (2008). The genome of *Polaromonas* sp. strain JS666: insights into the evolution of a hydrocarbon- and xenobiotic-degrading bacterium, and features of relevance to biotechnology. *Appl Environ Microbiol* **74**, 6405–6416.

McBeth, D. L. & Shapiro, J. A. (1984). Reversal by DNA amplifications of an unusual mutation blocking alkane and alcohol utilization in *Pseudomonas putida*. *Mol Gen Genet* **197**, 384–391.

Ogawa, N. & Miyashita, K. (1995). Recombination of a 3-chlorobenzoate catabolic plasmid from *Alcaligenes eutrophus* NH9 mediated by direct repeat elements. *Appl Environ Microbiol* **61**, 3788–3795.

Ohmoto, T., Sakai, K., Hamada, N. & Ohe, T. (1991). Salicylic acid metabolism through a gentisate pathway by *Pseudomonas* sp. TA-2. *Agric Biol Chem* **55**, 1733–1737.

Park, W., Jeon, C. O., Cadillo, H., DeRito, C. & Madsen, E. L. (2004). Survival of naphthalene-degrading *Pseudomonas putida* NCIB 9816-4 in naphthalene-amended soils: toxicity of naphthalene and its metabolites. *Appl Microbiol Biotechnol* **64**, 429–435.

Park, M., Jeon, Y., Jang, H. H., Ro, H.-S., Park, W., Madsen, E. L. & Jeon, C. O. (2007a). Molecular and biochemical characterization of 3-hydroxybenzoate 6-hydroxylase from *Polaromonas naphthalenivorans* CJ2. *Appl Environ Microbiol* **73**, 5146–5152.

Park, M., Jeon, Y., Madsen, E. L. & Jeon, C. O. (2007b). Protection of *Polaromonas naphthalenivorans* CJ2 from naphthalene toxicity by extracellular polysaccharide capsules. *J Appl Biol Chem* **50**, 41–45.

Penning, T. M., Burczynski, M. E., Hung, C. F., McCoull, K. D., Palackal, N. T. & Tsuruda, L. S. (1999). Dihydrodiol dehydrogenases and polycyclic aromatic hydrocarbon activation: generation of reactive and redox active o-quinones. *Chem Res Toxicol* **12**, 1–18.

Pumphrey, G. M. & Madsen, E. L. (2007). Naphthalene metabolism and growth inhibition by naphthalene in *Polaromonas naphthalenivorans* strain CJ2. *Microbiology* **153**, 3730–3738.

Ramos, J. L., Duque, E., Gallegos, M.-T., Godoy, P., Ramos-Gonzalez, M. I., Rojas, A., Terán, W. & Segura, A. (2002). Mechanisms of solvent tolerance in Gram-negative bacteria. *Annu Rev Microbiol* **56**, 743–768.

Rangnekar, V. M. (1988). Variation in the ability of *Pseudomonas* sp. strain B13 cultures to utilize meta-chlorobenzoate is associated with tandem amplification and deamplification of DNA. *J Bacteriol* **170**, 1907–1912.

Reams, A. B. & Neidle, E. L. (2003). Genome plasticity in *Acinetobacter*: new degradative capabilities acquired by the spontaneous amplification of large chromosomal segments. *Mol Microbiol* **47**, 1291–1304.

Reams, A. B. & Neidle, E. L. (2004). Gene amplification involves site-specific short homology-independent illegitimate recombination in *Acinetobacter* sp. strain ADP1. *J Mol Biol* **338**, 643–656.

Reber, H. (1973). Comparative studies with two pseudomonads on the sequential degradation of aromatic substances metabolized via different pathways. *Arch Mikrobiol* **89**, 305–315.

Ritalahti, K. M., Amos, B. K., Sung, Y., Wu, Q., Koenigsberg, S. S. & Löffler, F. E. (2006). Quantitative PCR targeting 16S rRNA and reductive dehalogenase genes simultaneously monitors multiple *Dehalococcoides* strains. *Appl Environ Microbiol* **72**, 2765–2774.

Sambrook, J. & Janssen, K. J. (2001). *Molecular cloning: a laboratory manual*, 3rd edn. New York: Cold Spring Harbor Laboratory.

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G. & Pühler, A. (1994). Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145**, 69–73.

Schell, M. A. & Wender, P. E. (1986). Identification of the *nahR* gene product and nucleotide sequences required for its activation of the *sal* operon. *J Bacteriol* **166**, 9–14.

Schweigert, N., Zehnder, A. J. & Eggen, R. I. (2001). Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals. *Environ Microbiol* **3**, 81–91.

Sikkema, J., de Bont, J. A. & Poolman, B. (1995). Mechanisms of membrane toxicity of hydrocarbons. *Microbiol Rev* **59**, 201–222.

Stanier, R. Y., Palleroni, N. J. & Doudoroff, M. (1966). The aerobic pseudomonads: a taxonomic study. *J Gen Microbiol* **43**, 159–271.

Teeling, H., Meyerdierks, A., Bauer, M., Amann, R. & Glöckner, F. O. (2004). Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environ Microbiol* **6**, 938–947.

Vaillancourt, F. H., Labbe, G. M., Drouin, N. M., Fortin, P. D. & Eltis, L. D. (2002). The mechanism-based inactivation of 2,3-dihydroxybiphenyl 1,2-dioxygenase by catecholic substrates. *J Biol Chem* **277**, 2019–2027.

Vaillancourt, F. H., Bolin, J. T. & Eltis, L. D. (2006). The ins and outs of ring-cleaving dioxygenases. *Crit Rev Biochem Mol Biol* **41**, 241–267.

Yagi, J. M., Sims, D., Brettin, T., Bruce, D. & Madsen, E. L. (2009). The genome of *Polaromonas naphthalenivorans* strain CJ2, isolated from coal tar-contaminated sediment, reveals physiological and metabolic versatility and evolution through extensive horizontal gene transfer. *Environ Microbiol* **11**, 2253–2270.

Edited by: A. S. Ball