Supporting Information

1,4-Dioxane Biodegradation by *Mycobacterium dioxanotrophicus* PH-06 Is Associated with a Group-6 Soluble Di-iron Monooxygenase

Ya He[a], Jacques Mathieu[a], Yu Yang[a], Pingfeng Yu[a], Marcio L. B. da Silva[a], Pedro J. J. Alvarez[a#]

[a] Department of Civil and Environmental Engineering, Rice University, Houston, TX, 77005, USA

Corresponding author: Pedro J. J. Alvarez

Email: alvarez@rice.edu; Phone: +1-7133485903; Fax: +1-7133485268

Table of Contents

Material and Methods

Total RNA extraction

PH-06 cells in exponential growth phase (after half of the added substrates were consumed) were used to extract total RNA for RNA sequencing (RNA-seq) and reverse transcription quantitative polymerase chain reaction (RT-qPCR). RNAprotect Bacteria Reagent (Qiagen, Hilden, Germany) was added to each culture prior to extraction according to the manufacturer's instructions, and cultures were then incubated at room temperature (25 ℃) for 5 min to stabilize the RNA before further processing. The stabilized bacterial cultures were centrifuged for 10 min at 5000 × *g*. The supernatant was decanted and the pellets were immersed and lysed in an appropriate volume of Buffer RLT (RNase Mini Kit, Qiagen, Hilden, Germany) with 2-mercaptoethanol (Sigma-Aldrich, St. Louis, MO, USA). After incubation for 5 min at room temperature, the lysed biomass in Buffer RLT was transferred into MicroRNA Bead Tubes (UltraClean® Microbial RNA Isolation Kit, MO BIO, Carlsbad, CA, USA) containing 0.25 g of beads and vortexed at maximum speed for 5 min using a Vortex-Genie® 2 (MO BIO, Carlsbad, CA, USA) to further homogenize the cell lysate. Then the lysis solution was centrifuged at maximum speed for 30 s and the supernatant was transferred into a new tube and the total RNA was purified using an RNase Mini Kit (Qiagen, Hilden, Germany). On-column DNA digestion was performed during the purification using an RNase-Free DNase Set (Qiagen, Hilden, Germany) to get high quality of RNA.

Quantification of gene expression

The $2^{-\Delta\Delta C_T}$ method was used to quantify differential gene expression, and the results were analyzed with SDS 1.3.1.[1] The calculation equation is:

$$\Delta\Delta C_{T,\ \text{Target gene}} = (C_{T,\ \text{Target gene}} - C_{T,\ \text{Housekeeping gene}})_{\ \text{Treatment}} - (C_{T,\ \text{Target gene}} - C_{T,\ \text{Housekeeping gene}})_{\ \text{Control}}$$

(C_T is cycle threshold, treatment is dioxane or THF, and control is glucose or acetate). This value is used to calculate the expression fold change of the target gene, as follows:

$$\text{Expression fold change} = 2^{-\Delta\Delta C_T}$$

Here, a housekeeping gene (16S rRNA of PH-06) was used as an internal standard for RT-qPCR.

Design and evaluation of the prmA primer/probe set

Amino acid sequences of alpha hydroxylase subunits from different groups of SDIMOs were aligned using Clustal X 2.0[2] and the phylogenetic tree was then visualized using TreeGraph2.[3] Partial DNA sequence (from 200 to 600) of the *prmA* gene of PH-06 was used as the input sequence in Primer Quest (Integrated DNA Technologies, Coraville, IA) to generate a series of possible primer/probe sets that satisfied the design criteria for Taqman assays. After manual comparison and adjustment, the final primer/probe set was chosen. The amplicon was 103 bp in length.

The sequences (5'-3') of 16S rRNA primer/probe set are: Forward primer (CGGTGAATACGTTCYCGG), Reverse primer (GGWTACCTTGTTACGACTT), and Probe (FAM-CTTGTACACACCGCCCGTC-BHQ1).[4]

The genome size of PH-06 is 8.08 Mb[5] and DNA molecule weight is 9.12×10^{11} bp /ng of DNA [i.e., (6.022×10^{14} Da/ng of DNA)/(660 Da/bp)].The gene copy numbers were calculated using the following equation:

$$\frac{gene\ copies}{\mu L} = \left(\frac{ng\ of\ DNA}{\mu L}\right) \times \left(\frac{9.12\times10^{11} bp}{ng\ of\ DNA}\right) \times \left(\frac{genome}{8.08\times10^{6} bp}\right) \times \left(\frac{gene\ copies}{genome}\right)$$

The overall detection limit (DL) of *prmA* biomarker (in copy numbers/ml of bacterial culture or aquifer materials) was determined by the following equation:

$$Overall\ DL = \frac{qPCR\ instrument\ DL}{DNA\ recovery} \times Proportion\ of\ DNA\ used\ in\ qPCR\ reaction$$

qPCR instrument DL is identified as the minimum detectable copy number when seven sequential analyses are successful.

Dioxane degradation rates

The dioxane degradation experiments with four consortia (Table S3) were conducted in 250-mL bottles amended with 100-mL AMS medium with an initial dioxane concentration of 500 mg/L, while shaking on a Digital Linear Shaker (Scilogex, Rocky Hill, CT, USA) at 150 rpm (T=30 ℃, pH=7). For the dioxane impacted aquifer, it was a natural attenuation experiment (no nutrients were added). The aquifer is from a dioxane impacted site in Seattle, WA. The dioxane concentration was about 300 μg/L. The natural attenuation experiment was conducted at 15 ℃ to mimic the actual average groundwater temperature at this site.

Results and Discussion

Biomarker testing results of the aquifer

No *prmA* gene was detected in the dioxane impacted aquifer. The detection limit is 6000 – 7000 gene copies/ml aquifer, which is higher than the culture detection limit (3000-4000 gene copies/ml culture) because the DNA recovery from aquifer is lower than that from cultures. The abundance of 16S rRNA is 7.89 (log) gene copies per ml aquifer.

Supporting Figures

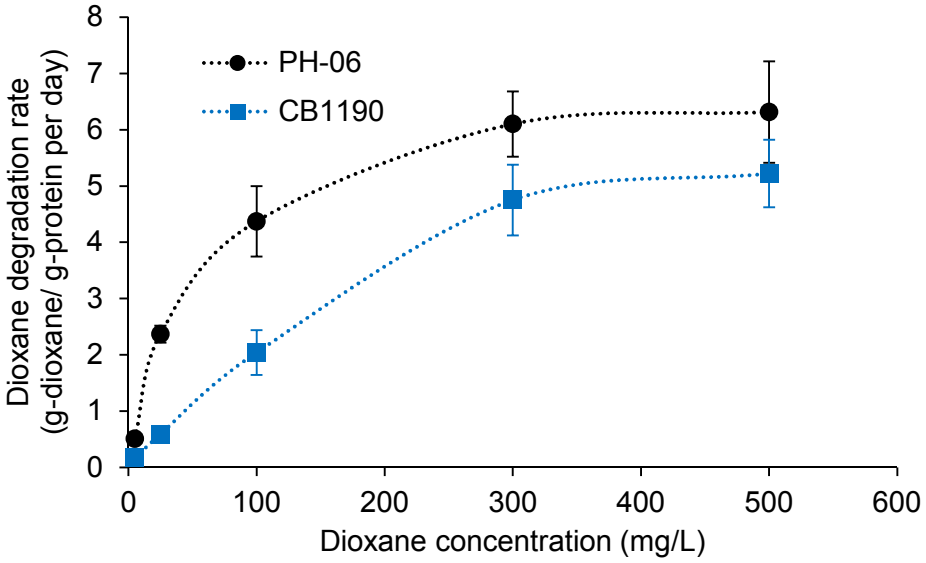

Figure S1. Monod plot for the dioxane degradation by *Mycobacterium dioxanotrophicus* PH-06 (Ks = 78±10 mg/L, Y = 0.16 g protein/g dioxane) and *Pseudonocardia dioxanivorans* CB1190 (Ks = 145±17 mg/L, Y = 0.11 g protein/ g dioxane). Our data for CB1190 are consistent with the previously reported values (Ks = 160±44 mg/L, Y = 0.09 g protein/ g dioxane).[6]

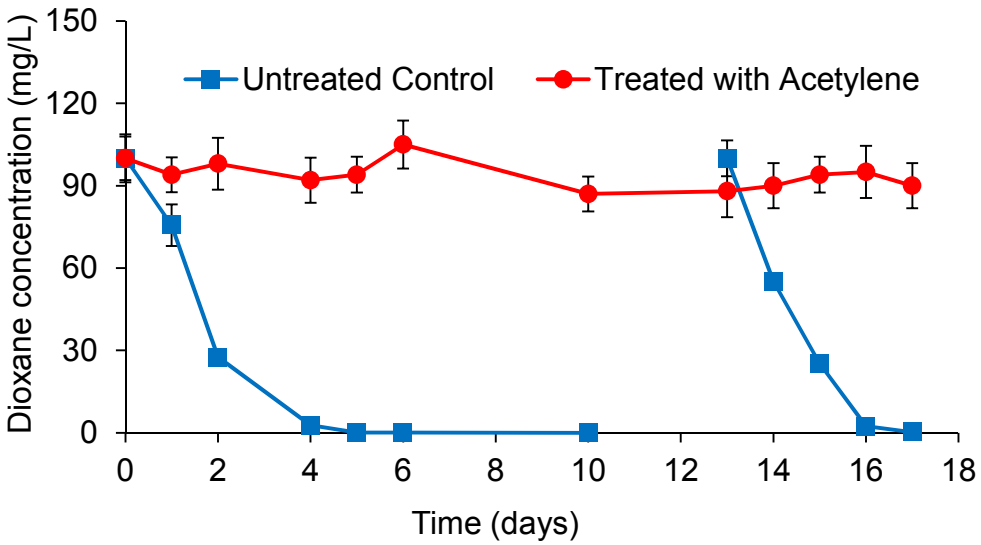

Figure S2. Inhibition by acetylene on biodegradation of dioxane by *Mycobacterium dioxanotrophicus* PH-06. The degradation capability was significantly inhibited by acetylene (8% in headspace) and was not recovered after the removal of acetylene. The initial dioxane and biomass concentration (measured as protein density) were 100 mg/L and 20 mg/L, respectively. On Day 13, acetylene was removed and 100 mg/L of dioxane was added into the control. Error bars represent ± one standard deviation from the mean of triplicate measurements.

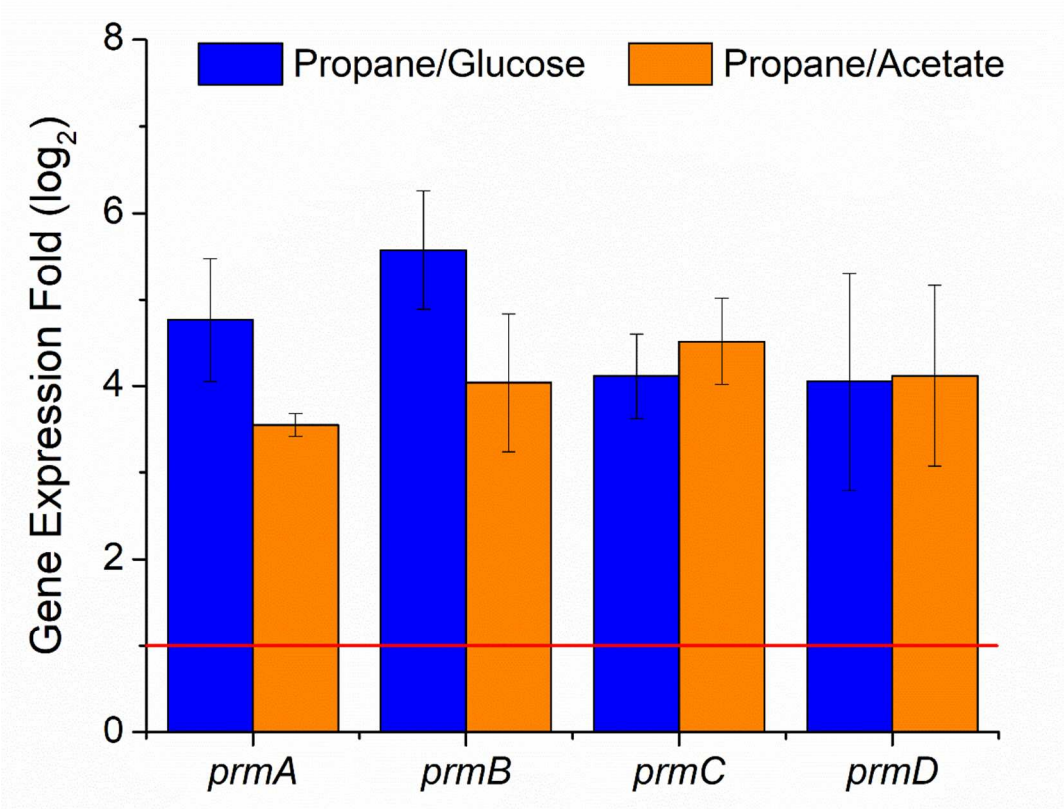

Figure S3. Gene expression fold of *prmABCD* in PH-06 pure cultures grown on propane relative to glucose or acetate. 16S rRNA gene of PH-06 was used as the housekeeping gene for error control among samples. The red horizontal lines indicated upregulation criteria (two-fold gene expression change).

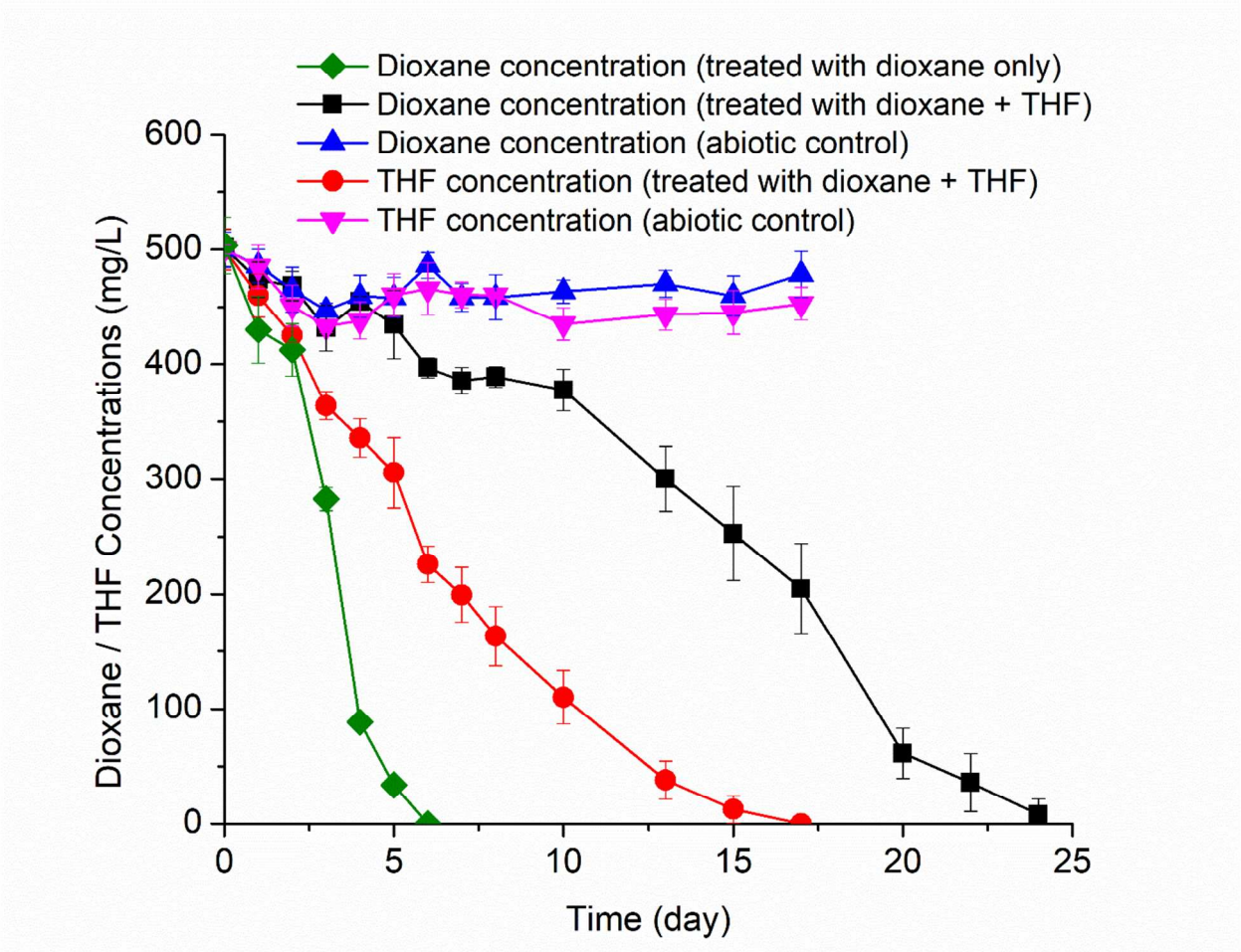

Figure S4. Inhibition of dioxane degradation by THF. Dioxane was degraded much faster when present alone than with THF. When dioxane and THF were added concurrently, the degradation of dioxane was inhibited initially, but after THF was consumed (80% degraded to 100 mg/L), dioxane was more rapidly degraded.

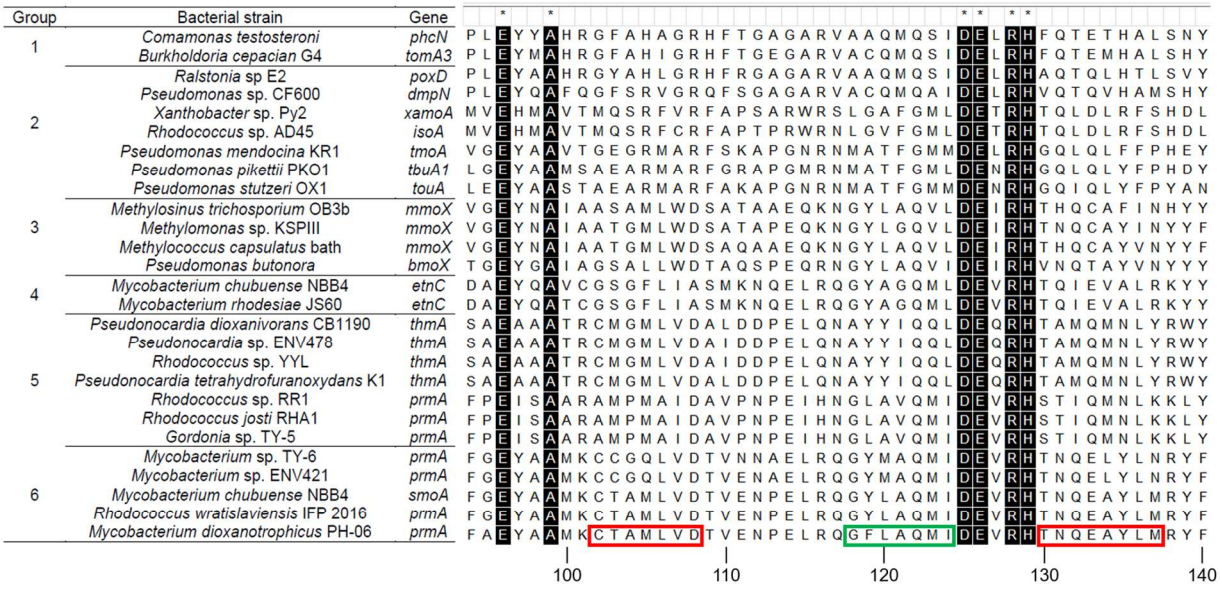

Figure S5. Alignment of the amino acid sequences of the alpha hydroxylase subunits for different groups of SDIMO. The amino acid residues selected for primers and probe design are marked in red and green boxes, respectively. The highlighted columns in black were 100% identity, including the di-iron center (i.e., DE*RH). Numbers below the sequences correspond to the numbering for PH-06.

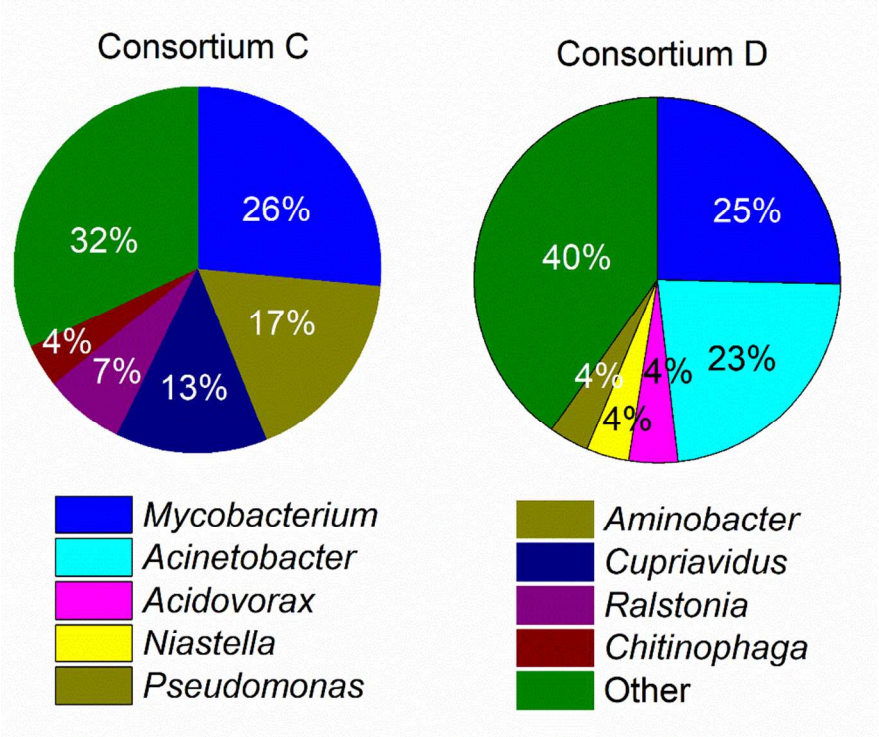

Figure S6. Composition of the microbial communities in Consortium C and D. The sum of all bacteria genera representing less than 3.5% of total community is indicated as "Other". The composition of the microbial communities in Consortium A and B can be found in a paper[7] published on the journal *Microbial Biotechnology* (DOI: 10.1111/1751-7915.12850).

Supporting Tables

Table S1. Specificity and coverage tests for the designed *prmA* primer/probe set.

Bacterial Strain	Dioxane degrader[a]	Gene name	Encoding enzyme[b]	SDIMO group	*prmA* detection[c]
Mycobacterium dioxanotrophicus PH-06	+	*prm*	Dioxane MO	6	+
Rhodococcus wratislaviensis IFP 2016	-	*prm*	Propane MO	6	-
Mycobacterium sp. ENV421	+	*prm*	Propane MO	6	-
Mycobacterium chubuense NBB4	-	*smo*	Propane MO	6	-
Pseudonocardia tetrahydrofuranoxydans K1	+	*thm*	Tetrahydrofuran MO	5	-
Pseudonocardia dioxanivorans CB1190	+	*thm*	Tetrahydrofuran MO	5	-
Rhodococcus jostii RHA1	+	*prm*	Propane MO	5	-
Mycobacterium chubuense NBB4	-	*etn*	Ethene MO	4	-
Mycobacterium chubuense NBB4	-	*smo*	Soluble methane MO	3	-
Pseudomonas putida CF600	-	*dmp*	Phenol HD	2	-
Pseudonocardia dioxanivorans CB1190	+	*phm*	Phenol HD	1	-
Mycobacterium vanbaalenii PYR-1	-	*nid*	PAH DO	-	-
Pseudomonas putida F1	-	*tod*	Toluene DO	-	-
Escherichia coli K-12	-	-	-	-	-
Bacteriophage λ	-	-	-	-	-

[a] Plus and minus signs indicate whether a specific strain can degrade dioxane or not, respectively, including metabolism and cometabolism.
[b]Abbreviations: MO, monooxygenase; HD, hydroxylase; DO, dioxygenase; PAH, polycyclic aromatic hydrocarbon.
[c] Plus and minus signs indicate positive and negative detection, respectively.

Table S2. Primers used in RT-qPCR to determine the relative gene expression levels of *prmABCD* under two different treatments.

Targeted Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')
prmA	TTT CCA GGG CTC GAC TTA	TCC AAC CGC TCG ATG TA
prmB	GCAGCGTCCTTACATCAA	CGCCAACGCTTCGTAATA
prmC	CCAGGCGACCGTGCATTA	GTCTCGAGCAGGAAGAACGTG
prmD	GGC GTC AAC CTG CAA TA	TCC TCG GAA TCG CAT AGA
16S rRNA	CAG GAC GCT GGT AGA GAT A	CGG GAC TTA ACC CAA CAT C

Table S3. Properties of the enriched consortia and dioxane impacted aquifer.

Name	Enriched source	Dioxane degrading rate (g dioxane/g protein/day)
Consortium A	Uncontaminated garden soil	7.1
Consortium B	Uncontaminated garden soil	5.7
Consortium C	Dioxane impacted site	3.02
Consortium D	Wastewater sludge	3.8
Dioxane impacted aquifer	-	0

Table S4. Monooxygenase genes upregulated during growth on dioxane relative to glucose shown by RNA-seq[a].

Gene ID[b]	Reference[c]	Similarity[d]	Gene product[e]	Replicon[f]	Expression fold change
BTO20_RS38210	**WP_005572960**	**94%**	**hydroxylase alpha chain (*prmA*)**	**P3**	**6.5**
BTO20_RS38215	**WP_005572958**	**90%**	**hydroxylase beta chain (*prmB*)**	**P3**	**8.1**
BTO20_RS38220	**WP_014805364**	**93%**	**coupling protein (*prmC*)**	**P3**	**9.7**
BTO20_RS38225	**WP_014805363**	**90%**	**reductase protein (*prmD*)**	**P3**	**10.7**
BTO20_RS38395	WP_014211353	90%	ammonia monooxygenase subunit C	P3	5.5
BTO20_RS38400	WP_014211354	100%	ammonia monooxygenase subunit A	P3	7.5
BTO20_RS38405	WP_014211355	94%	ammonia monooxygenase subunit B	P3	13.3
BTO20_RS22865	KUH99082	84%	cyclohexanone monooxygenase	C	13.9
BTO20_RS00055	WP_083164222	88%	antibiotic biosynthesis monooxygenase	C	8.9
BTO20_RS27410	WP_083164672	96%	4-hydroxyacetophenone monooxygenase	C	8.6
BTO20_RS11965	WP_083159483	94%	FAD-binding monooxygenase	C	6.8
BTO20_RS26085	WP_062831686	97%	alkane 1-monooxygenase	C	5.8
BTO20_RS22540	WP_095533860	98%	alkanesulfonate monooxygenase	C	4.7
BTO20_RS08645	WP_083163302	99%	nitronate monooxygenase	C	4.2
BTO20_RS31515	WP_062827933	99%	cytochrome P450	C	4.2
BTO20_RS24595	WP_083165874	99%	FMNH2-dependent monooxygenase	C	4.0
BTO20_RS24130	WP_091256594	89%	phenol 2-monooxygenase	C	4.0
BTO20_RS25380	WP_003208288	99%	dimethylaniline monooxygenase	C	3.3
BTO20_RS16950	WP_018599938	100%	phenylalanine 4-monooxygenase	C	3.3

[a] RNA sequence data were submitted to the GenBank database under accession number SRP116048.
[b] Gene ID: locus_tag on GenBank.
[c] Reference: accession number of the reference protein used to ascribe the gene product.
[d] Similarity: similarity with the reference obtained by NCBI BLAST.
[e] Gene product: The upregulated four subunits of *prmABCD* were highlighted in bold.
[f] Replicon: C, Chromosome; P3, Plasmid_3.

Table S5. Properties of the designed *prmA* primer/probe set.

Probe/Primer	Name	Sequence (5'-3')	Size	G+C (%)	T_m
Forward primer	*prmA*_307F	ACTGCGATGCTGGTTGAC	18	55.6	62
Reverse primer	*prmA*_410R	TCAGGTACGCCTCCTGATT	19	52.6	62
Taqman Probe	*prmA*_355P	/56-FAM/TTCCTCGCG/ZEN/ CAGATGATCGACG/3IABkFQ/	22	59	68

Table S6. Detection limit of various catabolic genes.

Biomarker	Targeted gene	Detection limit (reported)	Detection limit (converted)[a]	Reference
prmA	propane monooxygenase	3 - 4 $\times 10^3$ copies/ml	3 - 4 $\times 10^3$ copies/ml	This study
dxmB	dioxane monooxygenase	2 $\times$ 10^6 copies/ L	2 $\times$ 10^3 copies/ ml	8
poxD	phenol-2-monooxygenase	2 $\times$ 10^6 copies/ L	2 $\times$ 10^3 copies/ ml	8
thmA	dioxane monooxygenase	7 - 8 $\times 10^3$ copies/g of soil	2.4 - 2.7 $\times$ 10^4 copies/ ml	9
pmoA	particulate methane monooxygenase	1.9 $\times$ 10^4 copies/g of soil	6.5 $\times$ 10^4 copies/ml	10
PAH-RHDα	PAH-ring hydroxylating dioxygenase	2 $\times$ 10^3 copies/g of soil	6.8 $\times$ 10^3 copies/ml	11
nahAc	naphthalene dioxygenase	5 $\times$ 10^3 copies/g of soil	1.7 $\times$ 10^4 copies/ml	12

[a] Gene copies per g-soil were converted to copies per mL of pore volume by multiplying the former by the soil or aquifer material bulk density and dividing by its porosity. Assuming a bulk density of 1.7 g/ ml[13] and a porosity of 0.5,[14] the conversion factor is 3.4 g of soil/ ml of pore water.

References

1. Livak, K. J.; Schmittgen, T. D., Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **2001,** *25*, 402-408.
2. Larkin, M. A.; Blackshields, G.; Brown, N.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R., Clustal W and Clustal X version 2.0. *Bioinformatics* **2007,** *23*, 2947-2948.
3. Stöver, B. C.; Müller, K. F., TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC bioinformatics* **2010,** *11*, 7.
4. Suzuki, M. T.; Taylor, L. T.; DeLong, E. F., Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5′-nuclease assays. *Appl. Environ. Microbiol.* **2000,** *66*, 4605-4614.
5. He, Y.; Wei, K.; Si, K.; Mathieu, J.; Li, M.; Alvarez, P. J., Whole-genome sequence of the 1,4-dioxane-degrading bacterium *Mycobacterium dioxanotrophicus* PH-06. *Genome Announc.* **2017,** *5*, e00625-17.
6. Mahendra, S.; Alvarez-Cohen, L., Kinetics of 1,4-dioxane biodegradation by monooxygenase-expressing bacteria. *Environ. Sci. Technol.* **2006,** *40*, 5435-5442.
7. He, Y.; Mathieu, J.; Silva, M. L. B. d.; Li, M.; Alvarez, P. J., 1,4-Dioxane degrading consortia can be enriched from uncontaminated soils: Prevalence of *Mycobacterium* and soluble di-iron monooxygenase genes. *Microb. Biotechnol.* **2017**. (Accepted) DOI: 10.1111/1751-7915.12850
8. Gedalanga, P. B.; Pornwongthong, P.; Mora, R.; Chiang, S.-Y. D.; Baldwin, B.; Ogles, D.; Mahendra, S., Identification of biomarker genes to predict biodegradation of 1, 4-dioxane. *Appl. Environ. Microbiol.* **2014,** *80*, 3209-3218.

9. Li, M.; Mathieu, J.; Liu, Y.; Van Orden, E. T.; Yang, Y.; Fiorenza, S.; Alvarez, P. J., The abundance of tetrahydrofuran/dioxane monooxygenase genes (*thmA/dxmA*) and 1,4-dioxane degradation activity are significantly correlated at various impacted aquifers. *Environ. Sci. Technol. Lett.* **2013,** *1*, 122-127.

10. Kolb, S.; Knief, C.; Stubner, S.; Conrad, R., Quantitative detection of methanotrophs in soil by novel *pmoA*-targeted real-time PCR assays. *Appl. Environ. Microbiol.* **2003,** *69*, 2423-2429.

11. Cébron, A.; Norini, M.-P.; Beguiristain, T.; Leyval, C., Real-time PCR quantification of PAH-ring hydroxylating dioxygenase (PAH-RHD$_\alpha$) genes from Gram positive and Gram negative bacteria in soil and sediment samples. *J Microbiol. Methods* **2008,** *73*, 148-159.

12. Park, J.-W.; Crowley, D. E., Dynamic changes in *nahAc* gene copy numbers during degradation of naphthalene in PAH-contaminated soils. *Appl. Microbiol. Biotechnol.* **2006,** *72*, 1322-1329.

13. Twum, E. K.; Nii-Annang, S., Impact of soil compaction on bulk density and root biomass of *Quercus petraea* L. at Reclaimed Post-Lignite Mining Site in Lusatia, Germany. *Appl. Environ. Soil Sci.* **2015,** *2015*.

14. Nimmo, J., Porosity and pore size distribution. *Encyclopedia of Soils in the Environment* **2004,** *3*, 295-303.