

Supporting information

A benzene-degrading nitrate-reducing microbial consortium displays aerobic and anaerobic benzene degradation pathways

Siavash Atashgahi,^{1*#} Bastian Hornung,^{2#} Marcelle J van der Waals,^{1,3} Ulisses Nunes da Rocha,^{4,5} Floor Hugenholtz,¹ Bart Nijse,² Douwe Molenaar,⁴ Rob van Spanning,⁴ Alfons J.M. Stams,^{1,6} Jan Gerritse³ & Hauke Smidt¹

¹ Wageningen University & Research, Laboratory of Microbiology, Stippeneng 4, 6708 WE Wageningen, the Netherlands

² Wageningen University & Research, Laboratory of Systems and Synthetic Biology, Stippeneng 4, 6708 WE Wageningen, the Netherlands

³ Deltares, Subsurface and Groundwater Systems, Daltonlaan 600, 3584 BK Utrecht, the Netherlands

⁴ Vrije Universiteit Amsterdam, Department of Molecular Cell Biology, De Boelelaan 1108, 1081 HZ Amsterdam, the Netherlands

⁵ Department of Environmental Microbiology, Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany

⁶ Centre of Biological Engineering, University of Minho, Braga, Portugal

* For correspondence: E-mail: siavash.atashgahi@wur.nl, Tel: +31 317 484683

Equal contribution

Text mining

All domains derived from InterproScan were matched against the Brenda database with the following procedure. The text mining algorithm included lower casing all characters, removal of non-alphanumeric characters (colons, commas, brackets, apostrophes, dashes, terminal points), removal of partial and generic terms (type, terminal, subunit, domain, enzyme, like, hypothetical, conserved, operon, active site, probably, central, 51 kd, respiratory chain, c terminal, n terminal), manual rejection of too generic final result terms (kinase, cytochrome, protein, methyltransferase) and reduction of certain terms (deletion of PEP/pyruvate binding, removal of "prokaryotic" in "prokaryotic cytidylate kinase"; "family" in "cytidilate kinase family"; "phosphorylating" in "glyceraldehyde phosphate dehydrogenase phosphorylating"; "iron containing" in "iron containing alcohol dehydrogenase"; "zinc containing" in "zinc containing alcohol dehydrogenase"; "manganese containing" in "manganese containing catalase"; "20 kd" in "nadh ubiquinone oxidoreductase 20 kd", replacement of "carboxyltransferase" with "carboxylase" in "pyruvate carboxyltransferase"). Furthermore all terms containing only one character were also removed in case the remaining name contained more than two words.

Figure S1. Proposed anaerobic biodegradation of benzene via methylation to toluene mediated by an unknown presumed methylase (shown by a question mark) followed by fumarate addition to form benzylsuccinate that is subsequently metabolized to benzoyl-CoA. The genes encoding different enzymes of this pathway are: *bssABC*, benzylsuccinate synthase; *bbsEF*, succinyl-CoA:(R)-benzylsuccinate CoA-transferase; *bbsG* (R)-benzylsuccinyl-CoA dehydrogenase; *bbsH*, phenylitaconyl-CoA hydratase; *bbsCD*, 2-[hydroxy(phenyl)methyl]-succinyl-CoA dehydrogenase; *bbsAB*, benzoylsuccinyl-CoA thiolase.

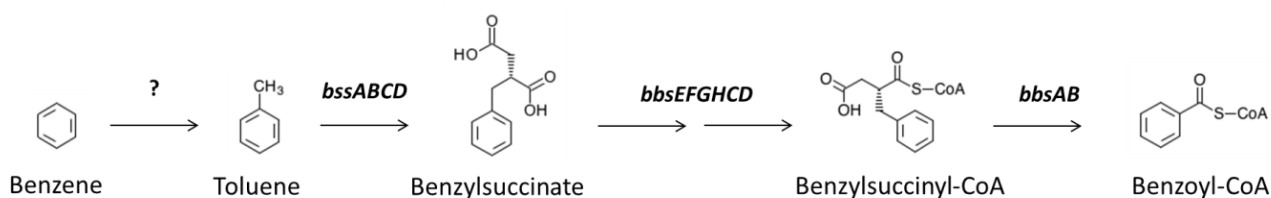


Figure S2. Putative benzene degradation gene cluster of the iron-reducing culture BF ¹ (A) and the nitrate-reducing culture used in this study (B) where Contig-100 is the default IDBA_UD output for a kmer-run of 100, the following number is the contig number (in this case 0) and the last number is the gene number on that contig (shown above each corresponding gene).

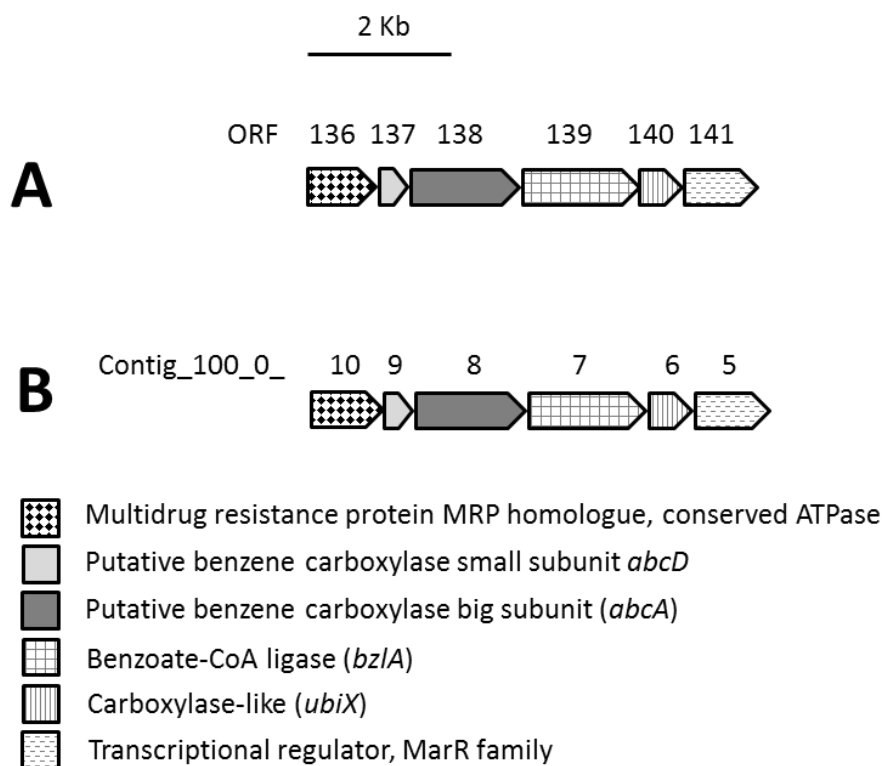


Figure S3. Oxygen concentration in the continuous culture. The black line with circles indicates the oxygen concentration (%) in the culture liquid, measured with an oxygen electrode. The red diamonds indicate oxygen in the headspace gas, measured with a GC-TCD. A spike of 0.5 mM nitrate to the continuous culture is indicated by arrow 1. After 2.5 hours of constant oxygen levels, a spike of 1 mM formate (arrow 2) was added, after which the oxygen concentration increased to 2.2% within 1.5 hours.

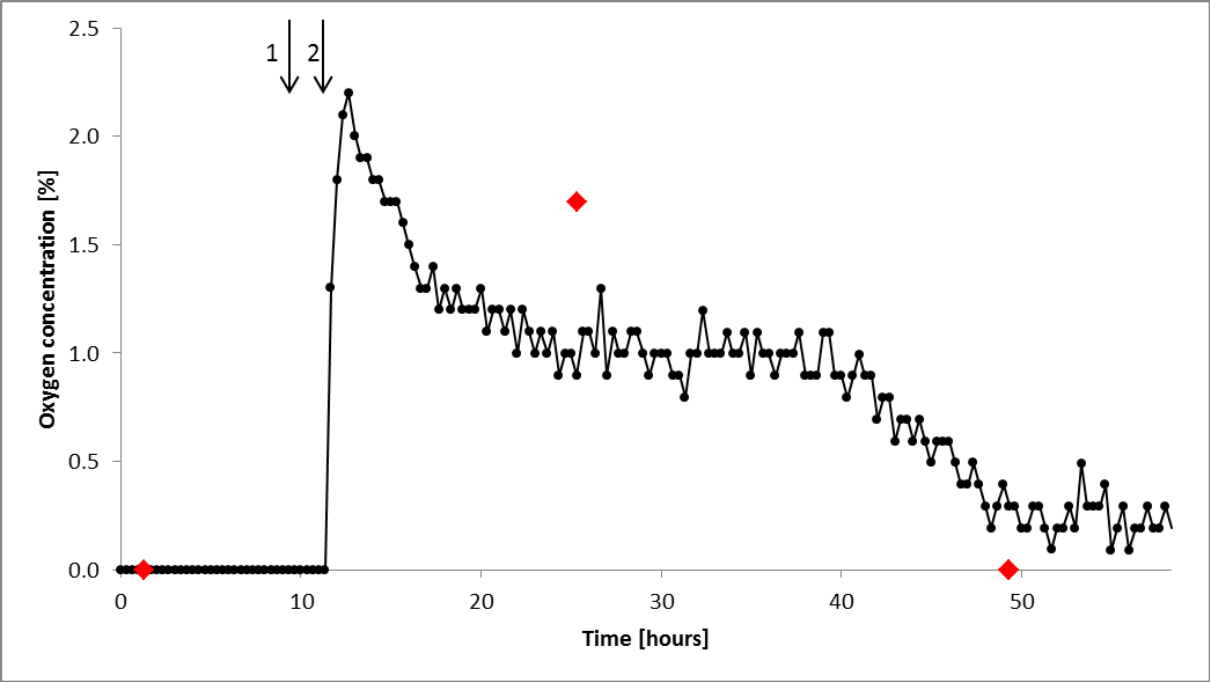


Figure S4. Relative abundances (%) of gene transcripts identified in biofilm samples involved in nitrate metabolism. The bar showing the number of relative abundance was log scaled and 0 values were removed.

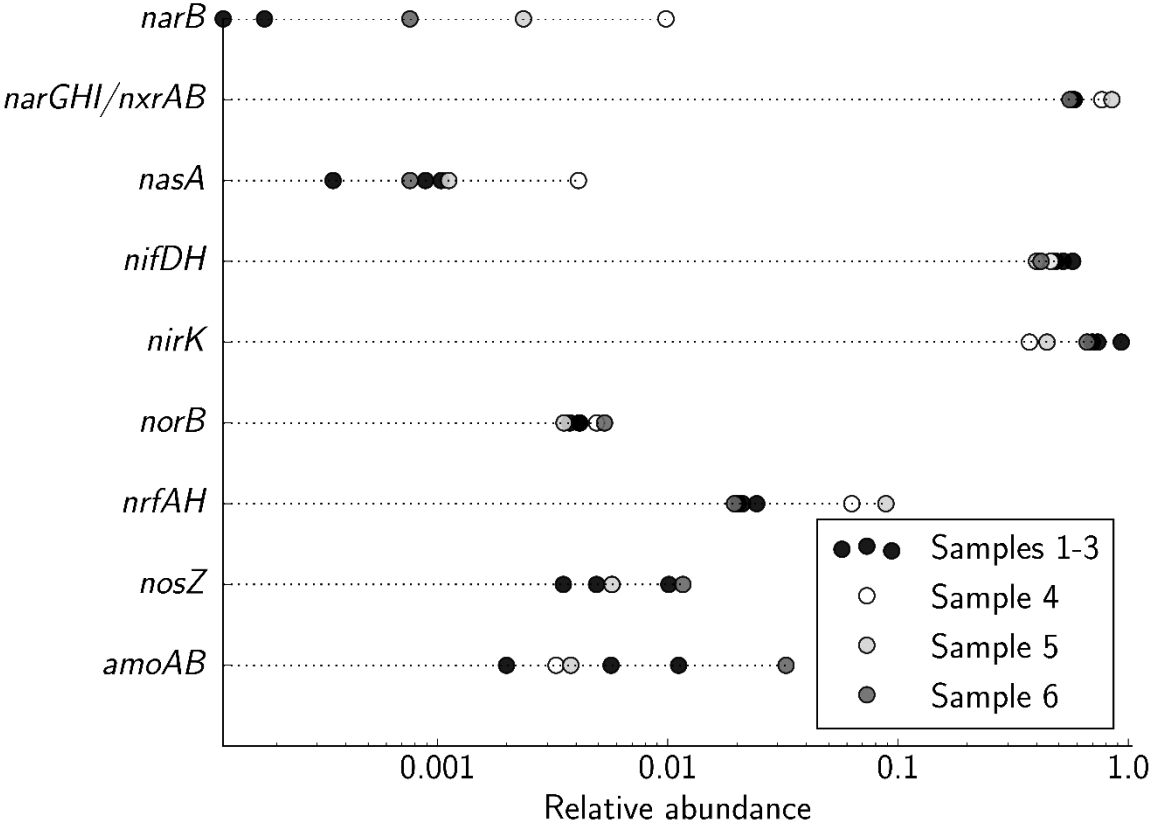


Table S1. Details of samples used in this study. The total amount of biofilm protein in the culture vessel was assessed by extrapolating the amount of protein released from defined surface areas to the total area of biofilm covered surfaces in the system. Protein content in the liquid (15 mg), white (233 mg) and brown biofilm (2485 mg) was used to calculate the protein content percentage in the liquid and biofilm in the samples. Protein content was analyzed as described ².

#	Biofilm morphologies	Sampling date	Samples volume (ml)	Scraped off biofilm area (cm ²)		Protein content (%)	
				White biofilm	Brown biofilm	Liquid	Biofilm
Sample 1	White	31-10-2014	50	6.3	0	89	11
Sample 2	White	31-10-2014	50	6.3	0	89	11
Sample 3	White	31-10-2014	50	6.3	0	89	11
Sample 4	White	3-11-2014	50	6.3	0	89	11
Sample 5	Brown	3-11-2014	50	22.3	4	30	70
Sample 6	Brown	3-11-2014	50	22.3	8	0	100

Table S2. Sequence information summary for the bioreactor biofilm and effluent samples during growth on benzene and nitrate.

Samples	Read number	Non-rRNA read (%)	Reads that passed PRINSEQ Lite quality filtering (%)	Read mean length (bp)	Percentage of reads passing all quality control	Mapping rate of reads passing quality control to the assembly
Sample 1	46403410	46282920 (99.7)	977319 (2.1)	97	2.1	54.6
Sample 2	6768045	6746142 (99.7)	367039 (5.4)	93.22	5.4	53.7
Sample 3	12225282	11938239 (97.7)	3580969 (30)	97.66	30.0	68.3
Sample 4	1169416	1089029 (93.1)	565017 (51.9)	126.37	51.9	78.3
Sample 5	6015250	5694797 (94.7)	2267743 (39.8)	127.87	39.8	84.7
Sample 6	11080970	10896851 (98.3)	1175823 (10.8)	126.33	10.8	81.9

Table S4. Summary of transcribed genes involved in nitrogen metabolism. The first column lists the transcribed genes followed by the locus tag of each gene. The third column shows the taxonomy of the locus tag, based on megablast/blastn hits of the whole contig against the NCBI NT database. The fourth column is the relative contribution of this locus tag to this function (e.g. if two genes with equal expression were assigned to one function, both would have 50% contribution to that function). The last four columns show the function of the most similar protein as identified by blastp (based on the locus' protein sequence) in the Uniprot database, followed by the accession number of the hit, the identity on protein level and the taxonomy of this entry, respectively.

Gene(s)	Locus tag	Taxon of closest match	Contribution to function (%)	Best blast hit ^a	Accession number of the blast hit	Identity (%)	Taxonomy of the best blast hit
<i>narB</i>	Contig-100_6458_1 ^b	<i>Candidatus</i> <i>Kuenenia stuttgartiensis</i>	60	Similar to FeS molybdopterin oxidorecutcase proteins	Q1PVH1	94	<i>Candidatus</i> <i>Kuenenia stuttgartiensis</i>
<i>narG/nxrA</i>	Contig-100_26_5	unclassified_Bacteria	37	Uncharacterized protein	A0A0B5GVF0	43	<i>Haloarcula</i> sp. CBA1115
<i>narG/nxrA</i>	Contig-100_18_5	<i>Candidatus</i> <i>Kuenenia stuttgartiensis</i>	11	Similar to nitrate reductase subunit NarG	Q1PZD8	99	<i>Candidatus</i> <i>Kuenenia stuttgartiensis</i>
<i>narG/nxrA</i>	Contig-100_7_6	Unclassified bacteria	8	Uncharacterized protein	A0A0M2U6Y8	71	<i>Clostridiales</i> bacterium PH28_bin88
<i>narH/nxrB</i>	Contig-100_26_4	Unclassified bacteria	12	Uncharacterized protein	T0MN75	53	Candidate division <i>Zixibacteria</i> bacterium RBG-1
<i>narH/nxrB</i>	Contig-100_18_2	<i>Candidatus</i> <i>Kuenenia stuttgartiensis</i>	6	Strongly similar to nitrate reductase (NarH)	Q1PZD5	100	<i>Candidatus</i> <i>Kuenenia stuttgartiensis</i>
<i>narH/nxrB</i>	Contig-100_7_7	Unclassified bacteria	7	Chemotaxis protein CheY	A0A0M2U7X6	75	<i>Clostridiales</i> bacterium PH28_bin88
<i>narI</i>	Contig-100_51_2	Unclassified <i>Peptococcaceae</i>	7	Putative iron-sulfur-binding reductase	F7NDM4	64	<i>Acetonema longum</i> DSM 6540
<i>nasA</i>	Contig-100_1580_1	N/A ^c	100	Uncharacterized protein	X0UVK5	50	Marine sediment metagenome
<i>nifD</i>	Contig-100_75_2	<i>Desulfitobacterium hafniense</i>	4	Oxidoreductase, nitrogenase component 1	G9XNJ1	74	<i>Desulfitobacterium hafniense</i> DP7
<i>nifH</i>	Contig-100_0_18	Unclassified <i>Clostridia</i>	78	Putative multidrug resistance protein	D8WWN7	99	<i>Clostridia</i> bacterium enrichment culture clone BF
<i>nifH</i>	Contig-100_75_1	<i>Desulfitobacterium hafniense</i>	15	Putative nitrogenase iron protein	G9XNJ2	77	<i>Desulfitobacterium hafniense</i> DP7
<i>nirK</i>	Contig-100_276_2	<i>Desulfomonile tiedjei</i>	91	Uncharacterized protein	A0A0M2U6Z1	68	<i>Clostridiales</i> bacterium PH28_bin88
<i>norB</i>	Contig-100_4457_1	<i>Pontibacter</i> sp. BAB1700	58	Putative nitric oxide reductase	F8S994	74	Uncultured bacterium
<i>norB</i>	Contig-100_865_2	<i>Formosa</i> sp. AK20	26	Nitric oxide reductase	A0A0F2NRP9	96	<i>Flavobacteriales</i> bacterium BRH_c54
<i>nrfA</i>	Contig-100_77_1	<i>Candidatus</i> <i>Kuenenia stuttgartiensis</i>	39	Hypothetical (Triheme) protein	Q1Q0T9	99	<i>Candidatus</i> <i>Kuenenia stuttgartiensis</i>
<i>nrfA</i>	Contig-100_124_2	<i>Candidatus</i> <i>Kuenenia stuttgartiensis</i>	0	Cytochrome c family protein	Q1Q5N8	97	<i>Candidatus</i> <i>Kuenenia stuttgartiensis</i>
<i>nrfH</i>	Contig-100_3851_1	<i>Thermincola potens</i>	3	NapC/NirT cytochrome c domain-containing protein	A0A0L6W6D3	67	<i>Thermincola ferriacetica</i>
<i>nosZ</i>	contig-100_430_2	uncultured bacterium	100	Nitrous oxide reductase protein NosZ	I6ZZA8	74	<i>Melioribacter roseus</i> (strain JCM 17771 / P3M-2)
<i>amoA/pmoA</i>	Contig-100_1144_2	Unclassified <i>Nitrosomonas</i>	70	AmoA	E3UMY3	91	Uncultured bacterium

<i>amoB/pmoB</i>	Contig-100_8938_1	<i>Nitrosomonas eutropha</i>	18	Ammonia monooxygenase	A0A0F7KEV5	86	<i>Nitrosomonas communis</i>
------------------	-------------------	------------------------------	----	-----------------------	------------	----	------------------------------

^a Based on uniprot May 11, 2016

^b Contig-100 is the default IDBA_UD output for a kmer-run of 100, the following number is the contig number and last number is the gene number on that contig

^c Not assigned. No taxonomy could be assigned, because either no relevant blast hit was found, or relevant blast hits were too diverse

Reference

- 1 Abu Laban, N., Selesi, D., Rattei, T., Tischler, P. & Meckenstock, R. U. Identification of enzymes involved in anaerobic benzene degradation by a strictly anaerobic iron-reducing enrichment culture. *Environ. Microbiol.* **12**, 2783-2796 (2010).
- 2 van der Waals, M. J. *et al.* Benzene degradation in a denitrifying biofilm reactor: activity and microbial community composition. *Appl. Microbiol. Biotechnol.* **101**, 5175–5188 (2017).