SUPPORTING INFORMATION

for

Environment-specific Auxiliary Substrates Tailored for Effective Cometabolic Bioremediation of 1,4-Dioxane

Daiyong Deng, Dung Ngoc Pham, and Mengyan Li*

Department of Chemistry and Environmental Science, New Jersey Institute of Technology, Newark, NJ, USA 07102

*Address correspondence to Dr. Mengyan Li (mengyan.li@njit.edu)

Phone: +1-973-642-7095

Fax: +1-973-596-3586

S1. Materials and Methods

16S rRNA sequencing

The V3-V4 hypervariable regions of 16S rRNA were amplified using the primer set 338F (5'-ACTCCTACGGGAGGCAGCA-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3')[6] combined with barcode and adapter sequences. Briefly, the pooled-triplicate DNA samples were first amplified using a high fidelity KOD FX enzyme (KFX-101, TOYOBO Co.). The PCR products obtained from the first step were subsequently purified using the Monarch PCR & DNA Cleanup Kit (New England Biolabs). The second step of PCR was processed using Phusion DNA Polymerase (New England Biolabs). These steps were conducted following the manufacturers' instructions. All amplified amplicons were quantified using the Quant-iT 1X dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA) and pooled prior to 16S rRNA sequencing. High throughput 16S rRNA sequencing was performed by the Illumina HiSeq2500 platform for the generation of 2×250 bp paired-end reads.

An average of 170,573 raw reads was obtained per sample and further corrected using DADA2, resulting in Amplicon Sequence Variants (ASVs) ranging from 43,225 to 148,588 among 15 samples. ASVs were reported to infer true biological origins without the introduction of amplifying or sequencing bias but distinctive with as few as one nucleotide difference.[2] This approach provides a higher resolution of microbial communities than the traditional operational taxonomic unit (OTU) generated using sequencing clustering at a fixed similarity threshold (e.g. 97%). Thus, ASVs were used to perform diversity and differential abundance analysis in our study. For other downstream analyses, ASVs were clustered into operational taxonomic units (OTUs) to resolve high stochastic sequence variation at the ASV level.

Microbial community analysis

Within QIIME2**,** alpha diversity indices including observed species, Chao1 richness, Shannon, and Simpson diversity were calculated based on ASVs for each sample. The Kruskal-Wallis test was used for statistical assessment. For beta diversity, principal coordinate analysis (PCoA) was processed to visualize dissimilarities among microbial communities using the Bray Curtis distance.[7] Pairwise permutational multivariate analysis of variance (perMANOVA) analysis was performed to evaluate the significance of differences in beta diversity between each treatment using the adonis function.[8] FDR corrected *p* values according to Benjamini-Hochberg (a False Discovery Rate multiple test correction)[9] were estimated for both Kruskal-Wallis and pairwise perMANOVA tests.

Differential ranking for the identification of substrate-associated taxa

Differential ranking analysis estimates rankings of ASVs based on their log-fold change with respect to a specific factor (i.e., propane or 1-propanol vs initial).[4] The parameters were: –epochs 100000, – batch-size 5, differential-prior 0.1, –learning-rate 0.001, –min-sample-count 1000, –min-feature-count 10, –random-seed 2. This optimized model was then compared to a baseline model of 1, resulting in the Q^2 value of -0.201. Q^2 value is similar to the R^2 value computed from the classic linear regression and the negative value suggested possible overfitting related to the difference in microbial communities in different sources. The ASVs were considered as enriched due to presence of propane or 1-propanol if their log-fold changes were positive (i.e.

high-rank ASVs) and the ASVs were considered as enriched in initial conditions if their log-fold changes were negative (i.e. low-rank ASVs).

Qurro[5] was used to compute and visualize the log ratios of aforementioned ASV in propane, 1-propanol, and initial samples. It calculates natural log ratios between each selected high-rank ASV (red bar) and each selected low-rank ASV (blue bar) for each sample (i.e. propane, 1-propanol, or initial sample). This ratio-based comparison method elicits a major benefit in reducing potential false discovery rates due to unknown microbial load, which is a frequent issue when relative abundances of ASVs are employed as the results of 16S rRNA marker sequencing. Computing the logarithm of ASV ratios normalizes weights regarding their relative increase or decrease.

S2. Results and Discussion

Shifting of alpha diversity indices after bioaugmentation

With the selection pressure of the addition of DD4 and auxiliary substrates, four microbiomes (except SLU) showed decreases in richness (Chao1) and diversity (Shannon) by different degrees as compared to the initial conditions when microcosms were initially prepared (Table S1).

The decrease was modest for SGW1 and SGW2. The Chao1 richness indices were initially 794 and 558 for SGW1 and SGW2, respectively, reducing to 268~304 after the bioaugmentation with propane or 1-propanol as the substrate. Similarly, the Shannon diversity indices were initially 6.46 and 5.08 for SGW1 and SGW2, respectively, reducing to 2.90~4.99 after the bioaugmentation with propane or 1-propanol as the substrate.

The influence was greatest as observed for DGW. The Chao1 richness and Shannon diversity indices were 1045 and 6.42 before the treatment. After the bioaugmentation treatment, the Chao1 richness indices dropped to as low as 93 and 68 with propane and 1-propanol as the auxiliary substrate, respectively, suggesting the disappearance of >90% of the indigenous microbial species. Accordingly, the Shannon diversity indices were reduced to 2.19 and 2.90 with propane and 1-propanol as the auxiliary substrate, respectively. These results revealed that DD4 was able to compete with native microbes in DGW regardless of the auxiliary substrate that was applied to the microcosms.

SDT showed the highest initial Chao1 richness and Shannon diversity indices (1674 and 6.53) in the microcosms prepared with the river sediment samples. With propane as the auxiliary substrate, Chao1 richness and Shannon diversity indices dropped to 589 and 5.57, respectively. With 1-propanol as the auxiliary substrate, Chao1 richness and Shannon diversity indices dropped to 253 and 3.74, respectively. These results suggest the native microbiome was initially highly diverse and remained diverse when propane was amended, but became selected by 1-propanol.

Interestingly, increases in richness (Chao1) and diversity (Shannon) were observed in SLU where activated sludge was inoculated. The initial Chao1 richness and Shannon diversity indices were 573 and 6.38, respectively. They increased to 863 and 6.45 when propane was amended and to 1089 and 7.37 when 1-propanol was applied. This suggested the activated sludge contain a large variety of copiotrophic bacteria that were initially very low in abundance (not detected at the beginning) but grew quickly when exogenous substrates were fed to the community. These

microbes can compete with DD4 for the substrate and nutrients and hinder the effectiveness of the treatment.

Table S1. Alpha diversity of microbiomes in 5 different environmental matrices before and after the bioaugmentation with propane or 1-propanol.

Treatment	Sample	Observed ASVs	Chao1	Shannon
Initial	**I_SGW1**	789	794	6.46
	I_SGW2	550	558	5.08
	I_SDT	1657	1674	6.53
	I_SLU	539	573	6.38
	I_DGW	1028	1045	6.46
Propane	**P_SGW1**	303	304	4.99
	P_SGW2	261	279	2.90
	P_SDT	585	589	5.57
	P_SLU	815	863	6.45
	P_DGW	92	93	2.19
1-Propanol	**1P_SGW1**	259	268	4.42
	1P_SGW2	284	287	3.93
	1P_SDT	248	253	3.74
	1P_SLU	1071	1089	7.37
	1P_DGW	68	68	2.90

Table S2. Five taxa associated with DD4 bioaugmentation as revealed by the differential ranking.

OTU Number	Species Name	Percent Identity	E value	Accession Number	Description
OTU_1	*Azoarcus* sp. DD4	100%	0	CP022958.1	*Azoarcus* sp. DD4 chromosome, complete genome
OTU_2	*Ochrobactrum* sp.	100%	0	KT425063.1	*Ochrobactrum* sp. BOC2 16S ribosomal RNA gene, partial sequence
OTU_3	*Ferruginibacter alkalilentus*	97.4%	0	FJ177530.1	Sphingobacteriales bacterium HU1-GD23 16S ribosomal RNA gene, partial sequence
OTU_4	*Pseudoxanthomonas indica*	100%	0	MT269579.1	*Pseudoxanthomonas indica* strain NC24 16S ribosomal RNA gene, partial sequence
OTU_5	*Rhodococcus erythropolis*	100%	0	MT549100.1	*Rhodococcus erythropolis* strain RYA9 16S ribosomal RNA gene, partial sequence

OTU_1

GTGGGGAATTTTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGAGTGAA
GAAGGCCTTCGGGTTGTAAAGCTCTTTCGGACGGAAAGAAATCGCGCGGGCTAATA
TCCCGCGTGGATGACGGTACCGTAAGAAGAAGCACCGGCTAACTACGTGCCAGCAG
CCGCGGTAATACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGC
GCAGGCGGTTTTGTAAGACAGGTGTGAAATCCCCGGGCTTAACCTGGGAACTGCGCT
TGTGACTGCAAGGCTAGAGTACGGCAGAGGGGGGTGGAATTCCACGTGTAGCAGTG
AAATGCGTAGAGATGTGGAGGAACACCGATGGCGAAGGCAGCCCCCTGGGCCTGTA
CTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGG

OTU_2

GTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGAGTGAT
GAAGGCCCTAGGGTTGTAAAGCTCTTTCACCGGTGAAGATAATGACGGTAACCGGA
GAAGAAGCCCCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGGGCTAGC
GTTGTTCGGATTTACTGGGCGTAAAGCGCACGTAGGCGGACTTTTAAGTCAGAGGTG
AAATCCCGGGGCTCAACCCCGGAACTGCCTTTGATACTGGAAGTCTTGAGTATGGTA
GAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACAC
CAGTGGCGAAGGCGGCTCACTGGACCATTACTGACGCTGAGGTGCGAAAGCGTGGG
GAGCAAACAGG

OTU_3

GTAAGGAATATTGGTCAATGGACGAAAGTCTGAACCAGCCATGCCGCGTGAAGGAT
GAAGGTCCTCTGGATTGTAAACTTCTTTTATCTGGGACGAAAAAAGGGAATTCTTTC
TCACTTGACGGTACCAGAGGAATAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGG
TAATACGGAGGGTGCAAGCGTTATCCGGATTCACTGGGTTTAAAGGGAGCGTAGGT
GGGCAGGTAAGTCAGTGGTGAAATCTCCGGGCTTAACCCGGAAACTGCCATTGATA
CTATTTGTCTTGAATATTCTGGAGGTAAGCGGAATATGTCATGTAGCGGTGAAATGC
TTAGATATGACATAGAACACCAATTGCGAAGGCAGCTTGCTACGGAATTATTGACAC
TGAGGCTCGAAAGCGTGGGGATCAAACAGG

OTU_4

GTGGGGAATATTGGACAATGGGCGCAAGCCTGATCCAGCCATACCGCGTGGGTGAA
GAAGGCCTTCGGGTTGTAAAGCCCTTTTGTTGGGAAAGAAATCTTGCCGGTTAATAC
CTGGCGAGGATGACGGTACCCAAAGAATAAGCACCGGCTAACTTCGTGCCAGCAGC
CGCGGTAATACGAAGGGTGCAAGCGTTACTCGGAATTACTGGGCGTAAAGCGTGCG
TAGGTGGTGGTTTAAGTCTGCTGTGAAAGCCCTGGGCTCAACCTGGGAATTGCAGTG
GATACTGGGTCACTAGAGTGTGGTAGAGGGATGCGGAATTTCCGGTGTAGCAGTGA
AATGCGTAGAGATCGGAAGGAACATCCGTGGCGAAGGCGGCATCCTGGGCCAACAC
TGACACTGAGGCACGAAAGCGTGGGGAGCAAACAGG

OTU_5

GTGGGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGAT
GACGGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGACGAAGCGCAAGTGACGGTAC
CTGCAGAAGAAGCACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTG
CAAGCGTTGTCCGGAATTACTGGGCGTAAAGAGTTCGTAGGCGGTTTGTCGCGTCGT
TTGTGAAAACCAGCAGCTCAACTGCTGGCTTGCAGGCGATACGGGCAGACTTGAGT
ACTGCAGGGGAGACTGGAATTCCTGGTGTAGCGGTGAAATGCGCAGATATCAGGAG
GAACACCGGTGGCGAAGGCGGGTCTCTGGGCAGTAACTGACGCTGAGGAACGAAAG
CGTGGGTAGCGAACAGG

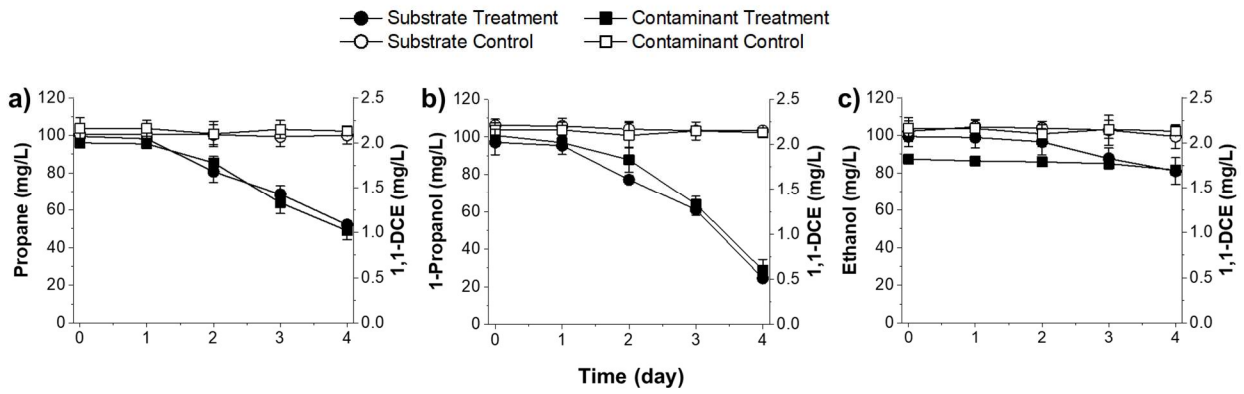

Figure S1. 1,1-DCE degradation by DD4 in NMS media with propane (a), 1-propanol (b), and ethanol (c) as primary substrates.

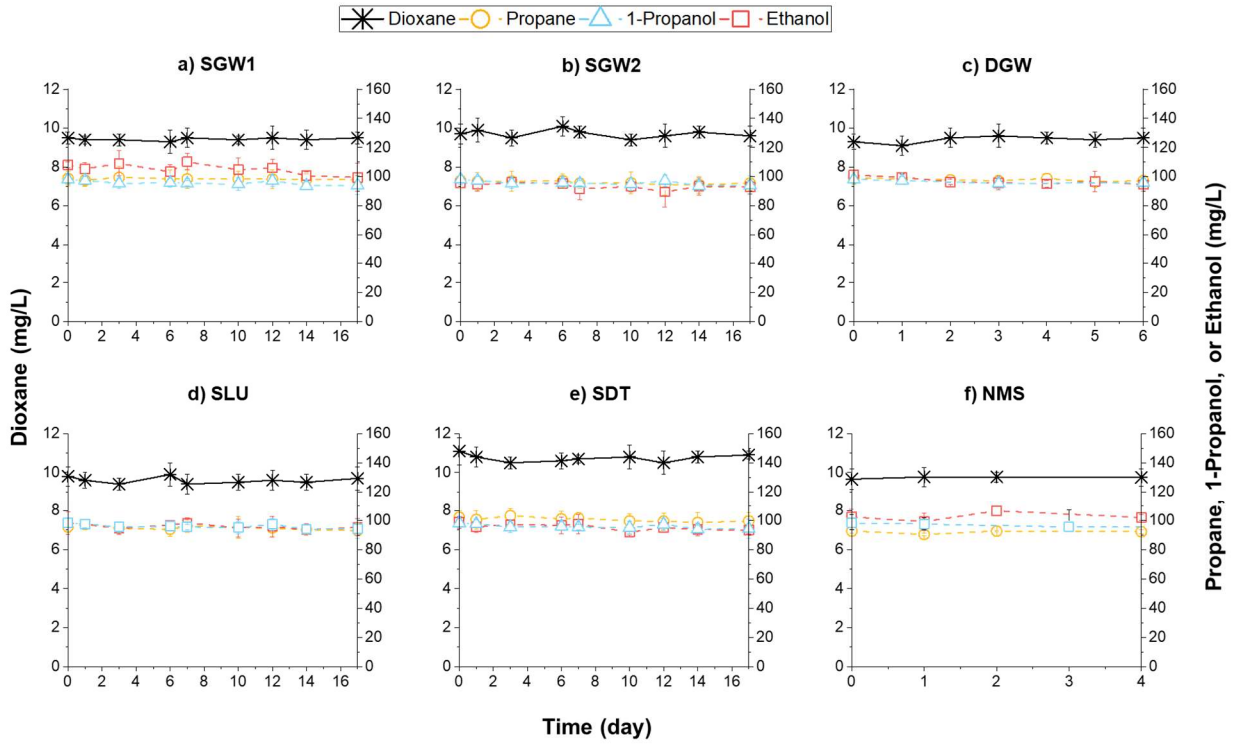

Figure S2. Concentrations of dioxane and different auxiliary substrates in control treatments inoculated with killed DD4 and biomass.

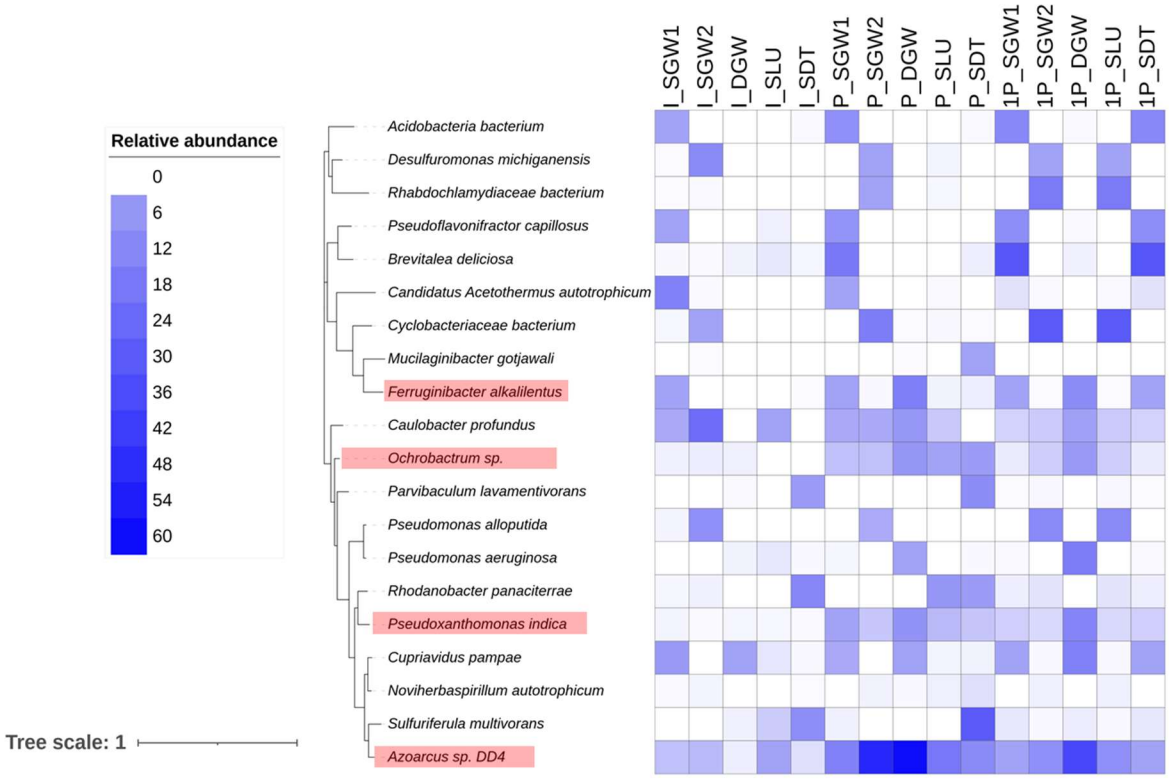

Figure S3. Phylogenic tree of the top 20 abundant OTUs based on the 500-time phylogeny bootstrap using the maximum likelihood algorithm. The heatmap shows relative abundances of OTUs in the microbial communities. I, initial; P, propane-amended; 1P, 1-propanol-amended.

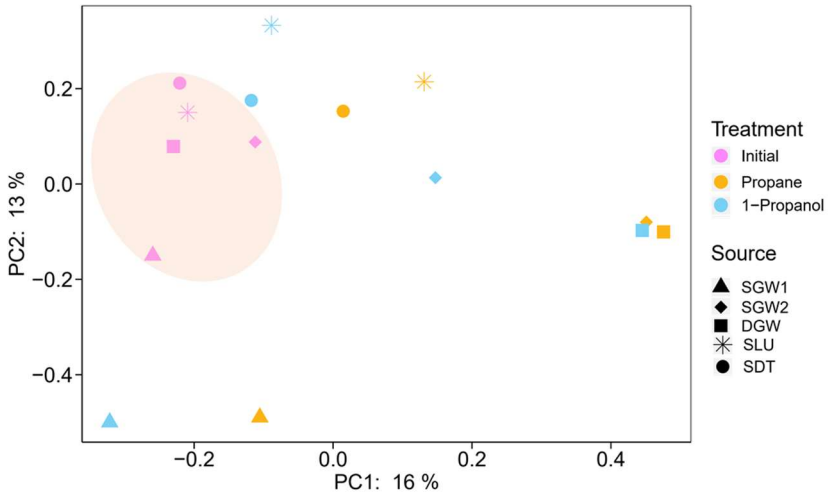

Figure S4. 2D-PCoA of the microbial community structures based on Bray Curtis distance.