TAXONOMIC AND FUNCTIONAL CHARACTERIZATION OF MICROBIAL COMMUNITIES LINKED TO CHLORINATED SOLVENT, 1,4-DIOXANE AND RDX BIODEGRADATION IN GROUNDWATER AND SOIL MICROCOSMS

By

Hongyu Dang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Environmental Engineering—Doctor of Philosophy

2021

ABSTRACT

TAXONOMIC AND FUNCTIONAL CHARACTERIZATION OF MICROBIAL COMMUNITIES LINKED TO CHLORINATED SOLVENT, 1,4-DIOXANE AND RDX BIODEGRADATION IN GROUNDWATER AND SOIL MICROCOSMS

By

Hongyu Dang

Bioremediation is becoming increasing popular for the remediation of sites contaminated with a range of different contaminants. Molecular methods such as 16S rRNA gene amplicon sequencing, shotgun sequencing, and high throughput quantitative PCR offer much potential for examining the microorganisms and functional genes associated with contaminant biodegradation, which can provide critical additional lines of evidence for effective site remediation.

In this work, the first project examined the taxonomic and functional biomarkers associated with chlorinated solvent and 1,4-dioxane biodegradation in groundwater from five contaminated sites. Each site had previously been bioaugmented with the commercially available dechlorinating mixed culture, SDC-9. The results highlighted the occurrence of numerous genera previously linked to chlorinated solvent biodegradation. The functional gene analysis indicated two reductive dehalogenase genes (*vcrA* and *tceA*) from *Dehalococcoides mccartyi* were abundant. Additionally, aerobic and anaerobic biomarkers for the biodegradation of various chlorinated compounds were observed across all sites. The approach used (shotgun sequencing) is advantageous over many other methods because an unlimited number of functional genes can be examined and a more complete picture of the functional abilities of microbial communities can be depicted.

Another research project evaluated the functional genes and species associated with RDX biodegradation at a RDX contaminated Navy site where biostimulation had

been adopted. For this, DNA samples extracted from groundwater samples pre- and post-biostimulation were subject to shotgun sequencing and high throughput qPCR. DNA sequences from thirty-one RDX biodegraders were detected, with the most abundant species being *Variovorax* sp. JS1663. Further, nine RDX biodegrading species significantly ($p<0.05$) increased in abundance following biostimulation. Both the sequencing data and qPCR indicated *xenA* and *xenB* exhibited the highest relative abundance among the six functional genes examined. Four genes, *diaA, nsfI*, *xenA* and *pnrB,* exhibited higher relative abundance values in some wells following biostimulation. The study provides a comprehensive approach for assessing biomarkers during RDX bioremediation and provides evidence that biostimulation generated a positive impact on a set of key species and genes.

A third study examined laboratory microcosms to determine the phylotypes and functional genes associated with the biodegradation of *cis*-dichloroethene (cDCE) and 1,4-dioxane. The impact of amending with lactate on cDCE and 1,4-dioxane biodegradation was also investigated. Stable isotope probing (SIP) was then used to determine which phylotypes were actively involved in biodegradation. The most enriched phylotypes for ^{13}C assimilation from 1,4-dioxane included *Rhodopseudomonas* and *Rhodanobacter*. The dominant enriched phylotypes for ^{13}C assimilation from cDCE included *Bacteriovorax*, *Pseudomonas* and *Sphingomonas*. The functional genes associated with the degradation of these contaminants was predicted using PICRUSt2. The results suggest aerobic concurrent biodegradation of cDCE and 1,4-dioxane should be considered for chlorinated solvent site remediation.

Overall, the data generated and approaches utilized in all three projects have the potential to be incorporated into diagnostic molecular tools for assessing biodegradation potential and for evaluating bioremediation performance at contaminated sites.

To my Mom and Dad,
for encouraging me to go on every adventure,
nurturing me to cherish the value of knowledge and everything else.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to Dr. Alison Cupples, who continually supported and guided me in the past four years. Without her help and exceptional scientific guidance, the work in this thesis would not have been possible. I would like to thank my committee members, Dr. Syed Hashsham, Dr. Irene Xagoraraki and Dr. Brian Teppen, for their valuable research advice.

I would like to thank the current and past members in our laboratory and Dr. Hashsham's laboratory, particularly, Yogendra Kanitkar, Robert Stedtfeld and Maggie Williams, for their willingness to teach me many methods during the early stages of my doctoral program. I am also grateful to the research efforts of Jenny Collier and Vidhya Ramalingam, as their work built foundations for my research. Also, thanks Dr. Benli Chai for introducing me to bioinformatic tools, which had lasting effects across this thesis. In addition, thanks to my friends, Jianzhou He and Xun Qian for enhancing my life and providing constructive advice in my research during my time at MSU.

Lastly, I would like to thank my family for their endless love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

PCE	Tetrachloroethene
TCE	Trichloroethene
DCE	Dichloroethene
VC	Vinyl Chloride
RDX	Hexahydro-1,3,5-trinitro-1,3,5-triazine
qPCR	Quantitative Polymerase Chain Reaction
RDase	Reductive Dehalogenase
SIP	Stable Isotope Probing
MG-RAST	Metagenomics Rapid Annotation using Subsystem Technology
DIAMOND	Double Index Alignment of Next-Generation Sequencing Data
KBS LTER	Kellogg Biological Station Long-Term Ecological Research
STAMP	Statistical Analysis of Taxonomic and Functional Profiles
KEGG	Kyoto Encyclopedia of Genes and Genomes
OTU	Operational Taxonomic Unit

CHAPTER 1 Introduction

1. Chlorinated Solvent, RDX and 1,4-Dioxane Contamination

Chlorinated Solvent Contamination. The chlorinated solvents tetrachloroethene (PCE) and trichloroethene (TCE) and their metabolites, *cis*-dichloroethene (cDCE) and vinyl chloride (VC), are persistent groundwater contaminants, requiring remediation because of their risks to human health (1). From the list of co-contaminants found in soil and groundwater, the chlorinated solvents and their metabolites are particularly prevalent (found at > 3000 active superfund sites (2), Table 1) and problematic due to their tendency to form large, dissolved-phase plumes and their recalcitrant nature (3).

RDX Contamination. RDX (Hexahydro-1,3,5-trinitro-1,3,5-triazine), also known as Royal Demolition Explosive, is a synthetic product and commonly used explosive (4). It is classified as a possible human carcinogen and its exposure can lead to irritation, nausea, kidney damage and other health impacts (4). RDX has caused soil, groundwater and sediment contamination because of the denotation of ordnance, firing of munitions on military training ranges, and the manufacturing and transport of munitions. Currently, there are 39 active sites with RDX on the National Priority List (2) (Table 1.1).

1,4-Dioxane Contamination. 1,4-Dioxane, a probable human carcinogen and common chlorinated solvent stabilizer, has been found at numerous contaminated sites across the U.S. (4, 5). In an examination from 49 remediation installations at U.S. Air Force sites, 1,4-dioxane was detected in 781 groundwater wells, with the percentage of 1,4-dioxane with TCE in all 1,4-dioxane detection-positive wells being 64.4% (6). In an evaluation of >2000 sites in California,

the chlorinated solvents were found in 94% of the sites with detections of 1,4-dioxane (7). There are 69 active super fund sites contaminated with 1,4-dioxane (2) (Table 1).

Table 1. 1. Number of active superfund sites with above contaminants as of 04/06/2021 (2)

Contaminants		Number of active sites
Chloroethene	PCE	898
	TCE	1007
	DCE (*cis*- and *trans*-DCE)	815
	VC	677
RDX		39
1,4-dioxane		68

2. Biodegradation of Chlorinated Solvents, 1,4-Dioxane and RDX

In the past decade, biostimulation (e.g. the addition of electron donors) has become increasingly popular for many organic contaminants. For the chlorinated solvents, in many cases, bioaugmentation is practiced, involving the addition of both electron donor and mixed microbial communities. Bioaugmentation is starting to become more recognized as a potential remediation method for 1,4-dioxane and RDX, but the applications are currently more limited compared to the chlorinated solvents.

Biodegradation of Chlorinated Solvents. *Dehalococcoides mccartyi*, is a key microorganism for the complete transformation of PCE to the non-hazardous end product, ethene (8, 9). *D. mccartyi* strains reduce chlorinated compounds obtaining energy from the reduction process (10-12). Examples of commercially available mixed cultures containing *D. mccartyi* for chlorinated solvent remediation include SDC-9 (from APTIM) and KB-1 (from SiREM) (13). It was estimated that several hundred sites in the U.S. have been subject to bioaugmentation with cultures containing *D. mccartyi* (14). Following bioaugmentation, remediation professionals commonly monitor *D. mccartyi* populations, typically targeting reductive dehalogenase (RDase) genes such as *pceA*, *tceA*, *vcrA* and *bvcA* (15-17) using quantitative PCR (qPCR) on nucleic acids extracted from groundwater (18-20). Other genera expressing RDases include

Dehalogenimonas (21), *Desulfitobacterium* (22), *Dehalobacter* (23), *Geobacter* (24), *Sulfurospirillum* (25) and *Anaeromyxobacter* (26).

In addition to anaerobic dehalogenation, chloroethenes can also be biodegraded aerobically. The genes encoding for the enzymes associated with aerobic VC degradation include *etnC* (alkene monooxygenase) and *etnE* (epoxyalkane: CoM transferase) (27-31). Also, the gene encoding for cytochrome P450 from *Polaromonas sp.* strain JS666 initializes the biodegradation of *cis*-1,2-dichoroethene (32). Other genes associated with co-metabolism of chlorinated ethenes include the α subunits of soluble and particulate methane monooxygenases (*mmoX* and *pmoA*) (33-35).

Biodegradation of RDX. RDX biodegradation is initiated by a number of bacteria and their associated enzymes under aerobic or anaerobic conditions. Under aerobic conditions, *nfsI* (present in both *Morganella morganii* strain B2 and *Enterobacter cloacae* strain 96-3) encodes a type I nitroreductase, which is responsible for the nitroreduction of RDX (36). Another RDX degrading functional gene, *pnrB*, was associated with *Pseudomonas* sp. and *Stenotrophomonas maltophilia* (37). Microorganisms within the genera *Rhodococcus, Gordonia, Williamsia* and *Microbacterium* have the well-studied *xplA* gene, which has been linked with nitro group removal and ring cleavage (38-41). Under anaerobic conditions, RDX transformation through nitro group denitration was initiated by the enzyme encoded by *diaA* from *Clostridium kluyveri* (42, 43). Finally, *xenA* and *xenB*, from the genus *Pseudomonas*, encode enzymes for the transformation of RDX to methylenedinitramine (44). These genes have previously been monitored at contaminated sites as evidence for RDX degradation, often using qPCR (45-47).

Biodegradation of 1,4-Dioxane. Many bacteria and enyzymes have been associated with the metabolic or co-metabolic biodegradation of 1,4-dioxane under aerobic conditions (48, 49),

with limited information available for anaerobic biodegradation. *Pseudonocardia dioxanivorans* strain CB1190 is a well studied 1,4-dioxane degrader able to use 1,4-dioxane as the carbon source via dioxane monooxygenase (50, 51). Other enyzmes involved in 1,4-dioxane biodegradation included toluene monooxygenase, propane monooxygenase, tetrahydrofuran monooxygenase and methane monooxygenase (49).

At contaminated sites, 1,4-dioxane is often present with the chlorinated solvents, which can impact aerobic 1,4-dioxane biodegradation. For example, for *P. dioxanivorans* CB1190, 1,1-DCE and cDCE had a greater effect on 1,4-dioxane biodegradation compared to TCE, while the effect of 1,1,1-trichloroethane (1,1-TCA) was negligible (52).

3. High Throughput Sequencing for Monitoring Biodegradation

Current approaches to detect biodegraders (targeting 16S rRNA genes or functional genes) for many contaminants in groundwater have typically focused on quantitative polymerase chain reaction or qPCR (16, 19, 45, 47, 53-56). Although qPCR has a high level of sensitivity, it can only target a limited number of genes (unless high throughput qPCR is used).

For example, during the natural attenuation of chlorinated solvents (57-59), following biostimulation (60) and batch 1,4-dioxane biodegradation (61), 16S rRNA gene amplicon sequencing has been used to monitor the dynamics of microbial communities. However, functions from the microbial communities can only be predicted (62, 63) rather than directly detected. High throughput sequencing offers an additional valuable tool for monitoring biomarkers environmental samples because a limitless number of biomarkers can be investigated. In contrast to 16S rRNA gene amplicon sequencing, shotgun sequencing captures random pieces of DNA, thus can sequence both the taxonomic and the potential functional characteristics of microbial communities (64).

4. Dissertation Outline and Objectives

The following summarizes the key objectives of each of the following chapters. Chapter 2 has been published (Dang, H., Kanitkar, Y. H., Stedtfeld, R. D., Hatzinger, P. B., Hashsham, S. A. and A. M. Cupples. 2018. Abundance of chlorinated solvent and 1,4-dioxane degrading microorganisms at five chlorinated solvent contaminated sites determined via shotgun sequencing. Environmental Science and Technology. 52 (23): 13914–13924), whereas Chapters 3 and 4 are currently being prepared for submission to peer reviewed journals.

Chapter 2. The project examined nucleic acids extracted from SDC-9 and groundwater from five chloroethene contaminated sites, previously bioaugmented with SDC-9. The overall objective was to develop the methodology for both taxonomic and functional analysis for chlorinated solvent contaminated sites. The specific objectives were to 1) determine the relative abundance of genera associated with chloroethene biodegradation; 2) investigate the relative abundance of reductive dehalogenases and other functional biomarkers involved in the biodegradation of chlorinated contaminants and 1,4-dioxane and 3) correlate the abundance of all biomarkers across individual wells.

Chapter 3. The project examined nucleic acids extracted from groundwater at an RDX contaminated military site using shotgun sequencing and high throughput qPCR. The specific objectives were to 1) determine the relative abundance of each functional gene, 2) ascertain the taxonomy of the microorganisms associated with each functional gene, 3) investigate changes in gene abundance following biostimulation and 4) ascertain if previously identified RDX degraders were present at the site and if their abundance changed following biostimulation.

Chapter 4. The study examined the concurrent biodegradation of cDCE and 1,4-dioxane in laboratory soil microcosms. The specific objectives were to 1) examine removal rates of the

co-contaminants cDCE and 1,4-dioxane by two soil microcosms, with and without lactate addition, 2) identify the microorganisms responsible for the uptake of ^{13}C from cDCE and 1,4-dioxane and 3) determine the functional genes present and correlate their presence to specific bacteria.

Chapter 5. This chapter generalizes the key findings and provides suggestions for future research.

REFERENCES

REFERENCES

1. Mattes TE, Alexander AK, Coleman NV. 2010. Aerobic biodegradation of the chloroethenes: pathways, enzymes, ecology, and evolution. FEMS Microbiol Rev 34:445-75.

2. EPA. 2021. Search Superfund Site Information. https://cumulis.epa.gov/supercpad/cursites/srchsites.cfm.

3. Yang Y, McCarty PL. 2000. Biologically Enhanced Dissolution of Tetrachloroethene DNAPL. Environmental Science & Technology 34:2979-2984.

4. EPA. 2017. Technical Fact Sheet – Hexahydro-1,3,5-trinitro1,3,5-triazine (RDX). https://www.epa.gov/sites/production/files/2017-10/documents/ffrro_ecfactsheet_rdx_9-15-17_508.pdf.

5. Zenker MJ, Borden RC, Barlaz MA. 2003. Occurrence and Treatment of 1,4-Dioxane in Aqueous Environments. Environmental Engineering Science 20:423-432.

6. Anderson RH, Anderson JK, Bower PA. 2012. Co-occurrence of 1,4-dioxane with trichloroethylene in chlorinated solvent groundwater plumes at US Air Force installations: Fact or fiction. Integrated Environmental Assessment and Management 8:731-737.

7. Adamson DT, Mahendra S, Walker KL, Rauch SR, Sengupta S, Newell CJ. 2014. A Multisite Survey To Identify the Scale of the 1,4-Dioxane Problem at Contaminated Groundwater Sites. Environmental Science & Technology Letters 1:254-258.

8. Muller JA, Rosner BM, von Abendroth G, Meshulam-Simon G, McCarty PL, Spormann AM. 2004. Molecular identification of the catabolic vinyl chloride reductase from *Dehalococcoides* sp strain VS and its environmental distribution. Applied and Environmental Microbiology 70:4880-4888.

9. He JZ, Ritalahti KM, Aiello MR, Loffler FE. 2003. Complete detoxification of vinyl chloride by an anaerobic enrichment culture and identification of the reductively dechlorinating population as a *Dehalococcoides* species. Applied and Environmental Microbiology 69:996-1003.

10. Maymo-Gatell X, Anguish T, Zinder SH. 1999. Reductive dechlorination of chlorinated ethenes and 1,2-dichloroethane by "*Dehalococcoides ethenogenes*" 195. Applied and Environmental Microbiology 65:3108-3113.

11. He J, Sung Y, Krajmalnik-Brown R, Ritalahti KM, Loffler FE. 2005. Isolation and characterization of *Dehalococcoides* sp strain FL2, a trichloroethene (TCE)- and 1,2-dichloroethene-respiring anaerobe. Environmental Microbiology 7:1442-1450.

12. Krajmalnik-Brown R, Holscher T, Thomson IN, Saunders FM, Ritalahti KM, Loffler FE. 2004. Genetic identification of a putative vinyl chloride reductase in *Dehalococcoides* sp strain BAV1. Applied and Environmental Microbiology 70:6347-6351.

13. Steffan RJ, Vainberg S. 2013. Production and handling of *Dehalococcoides* bioaugmentation cultures, p 89-115. *In* Stroo HF, Leeson A, Ward CH (ed), Bioaugmentation for Groundwater Remediation. Springer, New York.

14. Lyon DY, Vogel TM. 2013. Bioaugmentation for groundwater remediation: an overview, p 1-38. *In* Stroo HF, Leeson A, Ward CH (ed), Bioaugmentation for groundwater remediation. Springer, New York.

15. Kanitkar YH, Stedtfeld RD, Hatzinger PB, Hashsham SA, Cupples AM. 2017. Development and application of a rapid, user-friendly, and inexpensive method to detect *Dehalococcoides* sp reductive dehalogenase genes from groundwater. Applied Microbiology and Biotechnology 101:4827-4835.

16. Perez-de-Mora A, Zila A, McMaster ML, Edwards EA. 2014. Bioremediation of chlorinated ethenes in fractured bedrock and associated changes in dechlorinating and nondechlorinating microbial populations. Environmental Science & Technology 48:5770-5779.

17. Petrovskis EA, Amber WR, Walker CB. 2013. Microbial monitoring during bioaugmentation with *Dehalococcoides*, p 171-197. *In* Stroo HF, Leeson A, Ward CH (ed), Bioaugmentation for groundwater remediation. Springer, New York.

18. Hatt JK, Loffler FE. 2012. Quantitative real-time PCR (qPCR) detection chemistries affect enumeration of the *Dehalococcoides* 16S rRNA gene in groundwater. Journal of Microbiological Methods 88:263-270.

19. Lee PKH, Macbeth TW, Sorenson KS, Deeb RA, Alvarez-Cohen L. 2008. Quantifying genes and transcripts to assess the in situ physiology of "*Dehalococcoides*" spp. in a trichloroethene-contaminated groundwater site. Applied and Environmental Microbiology 74:2728-2739.

20. Liang Y, Liu XK, Singletary MA, Wang K, Mattes TE. 2017. Relationships between the abundance and expression of functional genes from vinyl chloride (VC)-degrading bacteria and geochemical parameters at VC-contaminated sites. Environmental Science & Technology 51:12164-12174.

21. Yang Y, Higgins SA, Yan J, Simsir B, Chourey K, Iyer R, Hettich RL, Baldwin B, Ogles DM, Loffler FE. 2017. Grape pomace compost harbors organohalide-respiring *Dehalogenimonas* species with novel reductive dehalogenase genes. ISME J 11:2767-2780.

22. Futagami T, Fukaki Y, Fujihara H, Takegawa K, Goto M, Furukawa K. 2013. Evaluation of the inhibitory effects of chloroform on ortho-chlorophenol- and chloroethene-dechlorinating *Desulfitobacterium* strains. AMB Express 3:30.

23. Maillard J, Schumacher W, Vazquez F, Regeard C, Hagen WR, Holliger C. 2003. Characterization of the corrinoid iron-sulfur protein tetrachloroethene reductive dehalogenase of *Dehalobacter restrictus*. Appl Environ Microbiol 69:4628-38.

24. Sung Y, Fletcher KE, Ritalahti KM, Apkarian RP, Ramos-Hernandez N, Sanford RA, Mesbah NM, Loffler FE. 2006. *Geobacter lovleyi* sp. nov. strain SZ, a novel metal-reducing and tetrachloroethene-dechlorinating bacterium. Appl Environ Microbiol 72:2775-82.

25. Buttet GF, Holliger C, Maillard J. 2013. Functional genotyping of *Sulfurospirillum* spp. in mixed cultures allowed the identification of a new tetrachloroethene reductive dehalogenase. Appl Environ Microbiol 79:6941-7.

26. He Q, Sanford RA. 2002. Induction characteristics of reductive dehalogenation in the ortho-halophenol-respiring bacterium, *Anaeromyxobacter dehalogenans*. Biodegradation 13:307-16.

27. Hartmans S, Debont JAM. 1992. Aerobic vinyl chloride metabolism in *Mycobacterium aurum* L1. Applied and Environmental Microbiology 58:1220-1226.

28. Danko AS, Saski CA, TomkinS JP, Freedman DL. 2006. Involvement of coenzyme M during aerobic biodegradation of vinyl chloride and ethene by *Pseudomonas putida* strain AJ and *Ochrobactrum* sp strain TD. Applied and Environmental Microbiology 72:3756-3758.

29. Coleman NV, Spain JC. 2003. Epoxyalkane: Coenzyme M transferase in the ethene and vinyl chloride biodegradation pathways of *Mycobacterium* strain JS60. Journal of Bacteriology 185:5536-5545.

30. Mattes TE, Coleman NV, Spain JC, Gossett JM. 2005. Physiological and molecular genetic analyses of vinyl chloride and ethene biodegradation in *Nocardioides* sp strain JS614. Archives of Microbiology 183:95-106.

31. Coleman NV, Spain JC. 2003. Distribution of the coenzyme m pathway of epoxide metabolism among ethene- and vinyl chloride-degrading *Mycobacterium* strains. Applied and Environmental Microbiology 69:6041-6046.

32. Nishino SF, Shin KA, Gossett JM, Spain JC. 2013. Cytochrome P450 Initiates Degradation of cis-Dichloroethene by *Polaromonas* sp Strain JS666. Applied and Environmental Microbiology 79:2263-2272.

33. Lee SW, Keeney DR, Lim DH, Dispirito AA, Semrau JD. 2006. Mixed pollutant degradation by *Methylosinus trichosporium* OB3b expressing either soluble or particulate methane monooxygenase: Can the tortoise beat the hare? Applied and Environmental Microbiology 72:7503-7509.

34. Yoon S, Im J, Bandow N, DiSpirito AA, Semrau JD. 2011. Constitutive expression of pMMO by *Methylocystis* strain SB2 when grown on multi-carbon substrates: implications for biodegradation of chlorinated ethenes. Environmental Microbiology Reports 3:182-188.

35. Chang HL, Alvarez-Cohen L. 1996. Biodegradation of individual and multiple chlorinated aliphatic hydrocarbons by methane-oxidizing cultures. Applied and Environmental Microbiology 62:3371-3377.

36. Kitts CL, Green CE, Otley RA, Alvarez MA, Unkefer PJ. 2000. Type I nitroreductases in soil *Enterobacteria* reduce TNT (2,4,6,-trinitrotoluene) and RDX (hexahydro-1,3,5-trinitro-1,3,5-triazine). Can J Microbiol 46:278-82.

37. Caballero A, Lazaro JJ, Ramos JL, Esteve-Nunez A. 2005. PnrA, a new nitroreductase-family enzyme in the TNT-degrading strain *Pseudomonas putida* JLR11. Environ Microbiol 7:1211-9.

38. Indest KJ, Crocker FH, Athow R. 2007. A TaqMan polymerase chain reaction method for monitoring RDX-degrading bacteria based on the *xplA* functional gene. J Microbiol Methods 68:267-74.

39. Bernstein A, Adar E, Nejidat A, Ronen Z. 2011. Isolation and characterization of RDX-degrading *Rhodococcus* species from a contaminated aquifer. Biodegradation 22:997-1005.

40. Seth-Smith HM, Rosser SJ, Basran A, Travis ER, Dabbs ER, Nicklin S, Bruce NC. 2002. Cloning, sequencing, and characterization of the hexahydro-1,3,5-Trinitro-1,3,5-triazine degradation gene cluster from *Rhodococcus rhodochrous*. Appl Environ Microbiol 68:4764-71.

41. Andeer PF, Stahl DA, Bruce NC, Strand SE. 2009. Lateral transfer of genes for hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) degradation. Appl Environ Microbiol 75:3258-62.

42. Bhushan B, Halasz A, Spain JC, Hawari J. 2002. Diaphorase catalyzed biotransformation of RDX via N-denitration mechanism. Biochem Biophys Res Commun 296:779-84.

43. Chakraborty S, Sakka M, Kimura T, Sakka K. 2008. Cloning and expression of a *Clostridium kluyveri* gene responsible for diaphorase activity. Biosci Biotechnol Biochem 72:735-41.

44. Fuller ME, McClay K, Hawari J, Paquet L, Malone TE, Fox BG, Steffan RJ. 2009. Transformation of RDX and other energetic compounds by xenobiotic reductases *XenA* and *XenB*. Appl Microbiol Biotechnol 84:535-44.

45. Fuller ME, McClay K, Higham M, Hatzinger PB, Steffan RJ. 2010. Hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) Bioremediation in Groundwater: Are Known RDX-Degrading Bacteria the Dominant Players? Bioremediation Journal 14:121-134.

46. Michalsen MM, King AS, Rule RA, Fuller ME, Hatzinger PB, Condee CW, Crocker FH, Indest KJ, Jung CM, Istok JD. 2016. Evaluation of Biostimulation and Bioaugmentation To Stimulate Hexahydro-1,3,5-trinitro-1,3,5,-triazine Degradation in an Aerobic Groundwater Aquifer. Environ Sci Technol 50:7625-32.

47. Michalsen MM, King AS, Istok JD, Crocker FH, Fuller ME, Kucharzyk KH, Gander MJ. 2020. Spatially-distinct redox conditions and degradation rates following field-scale bioaugmentation for RDX-contaminated groundwater remediation. Journal of Hazardous Materials 387:121529.

48. Mahendra S, Alvarez-Cohen L. 2006. Kinetics of 1,4-dioxane biodegradation by monooxygenase-expressing bacteria. Environ Sci Technol 40:5435-42.

49. He Y, Mathieu J, Yang Y, Yu PF, da Silva MLB, Alvarez PJJ. 2017. 1,4-Dioxane biodegradation by *Mycobacterium dioxanotrophicus* PH-06 is associated with a group-6 soluble di-iron monooxygenase. Environmental Science & Technology Letters 4:494-499.

50. Grostern A, Sales CM, Zhuang W-Q, Erbilgin O, Alvarez-Cohen L. 2012. Glyoxylate Metabolism Is a Key Feature of the Metabolic Degradation of 1,4-Dioxane by *Pseudonocardia dioxanivorans* Strain CB1190. Applied and Environmental Microbiology 78:3298.

51. Gedalanga PB, Pornwongthong P, Mora R, Chiang SY, Baldwin B, Ogles D, Mahendra S. 2014. Identification of biomarker genes to predict biodegradation of 1,4-dioxane. Appl Environ Microbiol 80:3209-18.

52. Zhang S, Gedalanga PB, Mahendra S. 2016. Biodegradation Kinetics of 1,4-Dioxane in Chlorinated Solvent Mixtures. Environ Sci Technol 50:9599-607.

53. Crocker FH, Indest KJ, Jung CM, Hancock DE, Fuller ME, Hatzinger PB, Vainberg S, Istok JD, Wilson E, Michalsen MM. 2015. Evaluation of microbial transport during aerobic bioaugmentation of an RDX-contaminated aquifer. Biodegradation 26:443-51.

54. Li M, Mathieu J, Liu Y, Van Orden ET, Yang Y, Fiorenza S, Alvarez PJJ. 2014. The Abundance of Tetrahydrofuran/Dioxane Monooxygenase Genes (*thmA*/*dxmA*) and 1,4-Dioxane Degradation Activity Are Significantly Correlated at Various Impacted Aquifers. Environmental Science & Technology Letters 1:122-127.

55. He Y, Mathieu J, da Silva MLB, Li M, Alvarez PJJ. 2018. 1,4-Dioxane-degrading consortia can be enriched from uncontaminated soils: prevalence of *Mycobacterium* and soluble di-iron monooxygenase genes. Microbial Biotechnology 11:189-198.

56. Ritalahti KM, Amos BK, Sung Y, Wu Q, Koenigsberg SS, Loffler FE. 2006. Quantitative PCR targeting 16S rRNA and reductive dehalogenase genes simultaneously monitors multiple *Dehalococcoides* strains. Appl Environ Microbiol 72:2765-74.

57. Kotik M, Davidova A, Voriskova J, Baldrian P. 2013. Bacterial communities in tetrachloroethene-polluted groundwaters: A case study. Science of the Total Environment 454:517-527.

58. Nemecek J, Dolinova I, Machackova J, Spanek R, Sevcu A, Lederer T, Cernik M. 2017. Stratification of chlorinated ethenes natural attenuation in an alluvial aquifer assessed by hydrochemical and biomolecular tools. Chemosphere 184:1157-1167.

59. Simsir B, Yan J, Im J, Graves D, Loffler FE. 2017. Natural attenuation in streambed sediment receiving chlorinated solvents from underlying fracture networks. Environmental Science & Technology 51:4821-4830.

60. Atashgahi S, Lu Y, Zheng Y, Saccenti E, Suarez-Diez M, Ramiro-Garcia J, Eisenmann H, Elsner M, Stams AJM, Springael D, Dejonghe W, Smidt H. 2017. Geochemical and microbial community determinants of reductive dechlorination at a site biostimulated with glycerol. Environmental Microbiology 19:968-981.

61. Miao Y, Johnson NW, Gedalanga PB, Adamson D, Newell C, Mahendra S. 2019. Response and recovery of microbial communities subjected to oxidative and biological treatments of 1,4-dioxane and co-contaminants. Water Research 149:74-85.

62. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI. 2020. PICRUSt2 for prediction of metagenome functions. Nat Biotechnol 38:685-688.

63. Aßhauer KP, Wemheuer B, Daniel R, Meinicke P. 2015. Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. Bioinformatics 31:2882-2884.

64. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. 2017. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol 35:833-844.

CHAPTER 2 Abundance of Chlorinated Solvent and 1,4-Dioxane Degrading Microorganisms at Five Chlorinated Solvent Contaminated Sites Determined via Shotgun Sequencing

This chapter is a modified version of a published work in Environmental Science & Technology: Hongyu Dang, Yogendra H. Kanitkar, Robert D. Stedtfeld, Paul B. Hatzinger, Syed A. Hashsham, and Alison M. Cupple Abundance of Chlorinated Solvent and 1,4-Dioxane Degrading Microorganisms at Five Chlorinated Solvent Contaminated Sites Determined via Shotgun Sequencing *Environmental Science & Technology* 2018 *52* (23), 13914-13924.

Abstract

Shotgun sequencing was used for the quantification of taxonomic and functional biomarkers associated with chlorinated solvent bioremediation in twenty groundwater samples (five sites), following bioaugmentation with SDC-9. The analysis determined the abundance of 1) genera associated with chlorinated solvent degradation, 2) reductive dehalogenase (RDases) genes, 3) genes associated with 1,4-dioxane removal, 4) genes associated with aerobic chlorinated solvent degradation and 5) *D. mccartyi* genes associated with hydrogen and corrinoid metabolism. The taxonomic analysis revealed numerous genera previously linked to chlorinated solvent degradation, including *Dehalococcoides, Desulfitobacterium* and *Dehalogenimonas*. The functional gene analysis indicated *vcrA* and *tceA* from *D. mccartyi* were the RDases with the highest relative abundance. Reads aligning with both aerobic and anaerobic biomarkers were observed across all sites. Aerobic solvent degradation genes, *etnC* or *etnE,* were detected in at least one sample from each site, as were *pmoA* and *mmoX*. The most abundant 1,4-dioxane biomarker detected was *Methylosinus trichosporium OB3b mmoX*. Reads aligning to *thmA* or *Pseudonocardia* were not found. The work illustrates the importance of shotgun sequencing to provide a more complete picture of the functional abilities of microbial communities. The

approach is advantageous over current methods because an unlimited number of functional genes can be quantified.

1. Introduction

The chlorinated solvents tetrachloroethene (PCE) and trichloroethene (TCE) and their metabolites, dichloroethene (DCE) and vinyl chloride (VC), are persistent groundwater contaminants, requiring remediation because of their risks to human health. Remediation efforts have involved biostimulation, through the addition of carbon sources, or bioaugmentation, which involves the injection of mixed microbial cultures containing *Dehalococcoides mccartyi* (1). *D. mccartyi* is a key microorganism for the complete transformation of these chemicals to the non-hazardous end product, ethene (2, 3). *D. mccartyi* strains reduce chlorinated compounds obtaining energy from the reduction process (4-6). Examples of commercially available mixed cultures containing *D. mccartyi* for chlorinated solvent remediation include SDC-9 (from APTIM, formerly CB&I, also marketed under several different names) and KB-1 (from SiREM) (1). It was estimated that several hundred sites in the US have been subject to bioaugmentation with cultures containing *D. mccartyi* (7). With the expansion of this remedial practice over the last decade, the number of sites in the US now numbers well over 2,300, and bioaugmentation has been performed in at least 11 other countries (*P Hatzinger, pers comm*). Following bioaugmentation, remediation professionals commonly monitor *D. mccartyi* populations, typically targeting reductive dehalogenase (RDase) genes such as *vcrA* (8-10) using quantitative PCR (qPCR) on nucleic acids extracted from groundwater (11-13).

While qPCR has been successful for documenting the occurrence and dechlorinating activity of *D. mccartyi* (9, 12, 14, 15) most laboratories only have the instrumentation (bench-top real-time thermal cycler) to target a small number of functional genes. Next generation

sequencing (NGS) is now becoming the tool of choice for environmental samples. For example, 16S rRNA gene amplicon NGS (16S rRNA-NGS) has been used to monitor microbial communities during chlorinated solvent natural attenuation (16-18), following biostimulation (9, 19) (20-22), during zerovalent iron-based (22, 23) and thermal-based (24, 25) chlorinated solvent remediation.

In contrast to 16S rRNA-NGS, shotgun (or whole genome) sequencing offers the opportunity to investigate both the taxonomic and the potential functional characteristics of microbial communities. However, only a limited number of researchers have adopted this approach for describing chlorinated solvent groundwater microbial communities. Notably, these studies have primarily focused on taxonomic data, without specifically addressing RDases or other functional genes related to chlorinated solvent degradation (26, 27). Others have examined dehalogenating genes in forest soils using shotgun sequencing (28). To our knowledge, the current work represents the first study to target contaminant degrading functional genes in groundwater from chlorinated solvent contaminated sites using shotgun sequencing.

The samples included groundwater (from twenty injection or monitoring wells, post bioaugmentation with SDC-9) from five contaminated sites as well as the bioaugmentation culture, SDC-9. Although other researchers have used NGS to study *D. mccartyi* containing enrichment cultures (e.g. KB-1, D2, ANAS) (29, 30), limited data is available on SDC-9. The overall objective was to develop the methodology to quantify chlorinated solvent and 1,4-dioxane degrading microorganisms in contaminated site groundwater using both taxonomic and functional analyses. We propose that this approach (or a derivative) will ultimately be the method of choice for predicting biodegradation potential at contaminated sites.

2. Methods

2.1 DNA Extraction from Groundwater and SDC-9

Groundwater samples from injection and monitoring wells were collected at five different chlorinated solvent sites (San Antonio TX, Tulsa OK, Edison NJ, Quantico VA, and Indian Head MD) through traditional low-flow sampling (31). Only one of the five locations (Tulsa, OK) was known to be contaminated with 1,4-dioxane. The water was pumped into sterile amber bottles (1L), which were placed on ice and then shipped overnight to Michigan State University. All sites were previously bioaugmented with the commercially available reductive dechlorinating culture, SDC-9 (32, 33). Details concerning groundwater sampling times and site characteristics have been summarized (Supplementary Table 2.1). Additional site information, when available, has also been provided (e.g. plume maps, plot layouts, concentration data over time) for each site (Supplementary Figures 2.1-15). DNA was extracted (collection on a filter, bead-beating and chemical lysis) from groundwater and mixed culture (SDC-9) samples using the PowerWater DNA isolation kit (Mo Bio Laboratories, a Qiagen Company) and previously described methods (8, 34).

2.2 Sequencing and Taxonomic Analysis

DNA extracts from twenty groundwater samples and the culture SDC-9 were submitted for library generation and sequencing to the Research Technology Support Facility Genomics Core at Michigan State University (MSU). Details on the preparation of libraries, the sequencing platform (Illumina HiSeq 4000) and the taxonomic analysis (Meta Genome Rapid Annotation using Subsystem Technology (MG-RAST) (35) are provided in the Supplementary Section (Supplementary Table 2.2).

2.3 Reference Sequences Collection, Functional Gene Analysis, qPCR

Two approaches were employed to analyze the functional gene data. First, protein sequences associated with RDases for published genomes were collected from the National Center for Biotechnology Information (NCBI). The microorganisms and genome information utilized in this analysis has been summarized (Supplementary Tables 2.3 and 2.4). Secondly, to enable a wider number of sequences to be examined, protein sequences were collected from additional sources e.g. Functional Gene Pipeline and Repository (FunGene) (36), NCBI BLAST. DIAMOND (double index alignment of next-generation sequencing data) (37) was used as the alignment tool for all functional genes. A stringent screening criteria approach (minimum sequence identity of 90% and alignment length of 49 amino acids) was adopted because of the similarity in many of the *D. mccartyi* genes (e.g. hydrogenases and corrinoid metabolism genes) between different strains. Detailed information on the collection of these sequences and the DIAMOND analysis has been provided (Supplementary Section). Quantitative PCR was performed to enumerate *vcrA* gene copies in each DNA extract using methods previously developed (34, 38) (see Supplementary Section).

3. Results

3.1 Sequencing and Taxonomic Analysis

From the twenty groundwater DNA extracts, the majority (seventeen) generated between ~4 and ~6 million sequences each, post quality control. Three samples (PMW2, MWAW1, IW7) produced lower sequence counts (157,000, 471,513 and 1,547,247). The average sequence length varied from 226 to 241 bp (standard deviations from 34 to 41 bp) (Supplementary Table 2.2). The rarefaction curves plateaued indicating the analysis had captured the majority of the diversity within the samples (Supplementary Figure 2.16).

Sequencing analysis of SDC-9 indicated the genera *Dehalococcoides* (31% of all sequences) and *Methanocorpusculum* (10%) were major components of the culture (Supplementary Figure 2.17). Other important microorganisms included those within the phyla *Bacteroidetes* (23%, primarily the genera *Parabacteroides* and *Bacteroides*), *Firmicutes* (19%, primarily *Desulfitobacterium, Desulfotomaculum, Clostridium* and *Bacillus*) and *Proteobacteria* (9%). For the groundwater, between two and five samples were studied for each of the five sites, with *Proteobacteria* and *Archaea* being dominant in many samples (Supplementary Figures 2.18-22)

3.2 Occurrence of Chlorinated Solvent Degrading Microorganisms in SDC-9 and In Situ

The sequencing data for each site was examined to determine the relative abundance of genera previously associated with chlorinated solvent degradation (Figure 2.1). It is important to note that this analysis is only at the genus level and therefore, except for *Dehalococcoides,* may overestimate the abundance of possible degrading microorganisms. *Dehalobacter* and *Desulfomonile* were not detected in any of the culture or groundwater samples by MG-RAST and are not included in Figure 2.1.

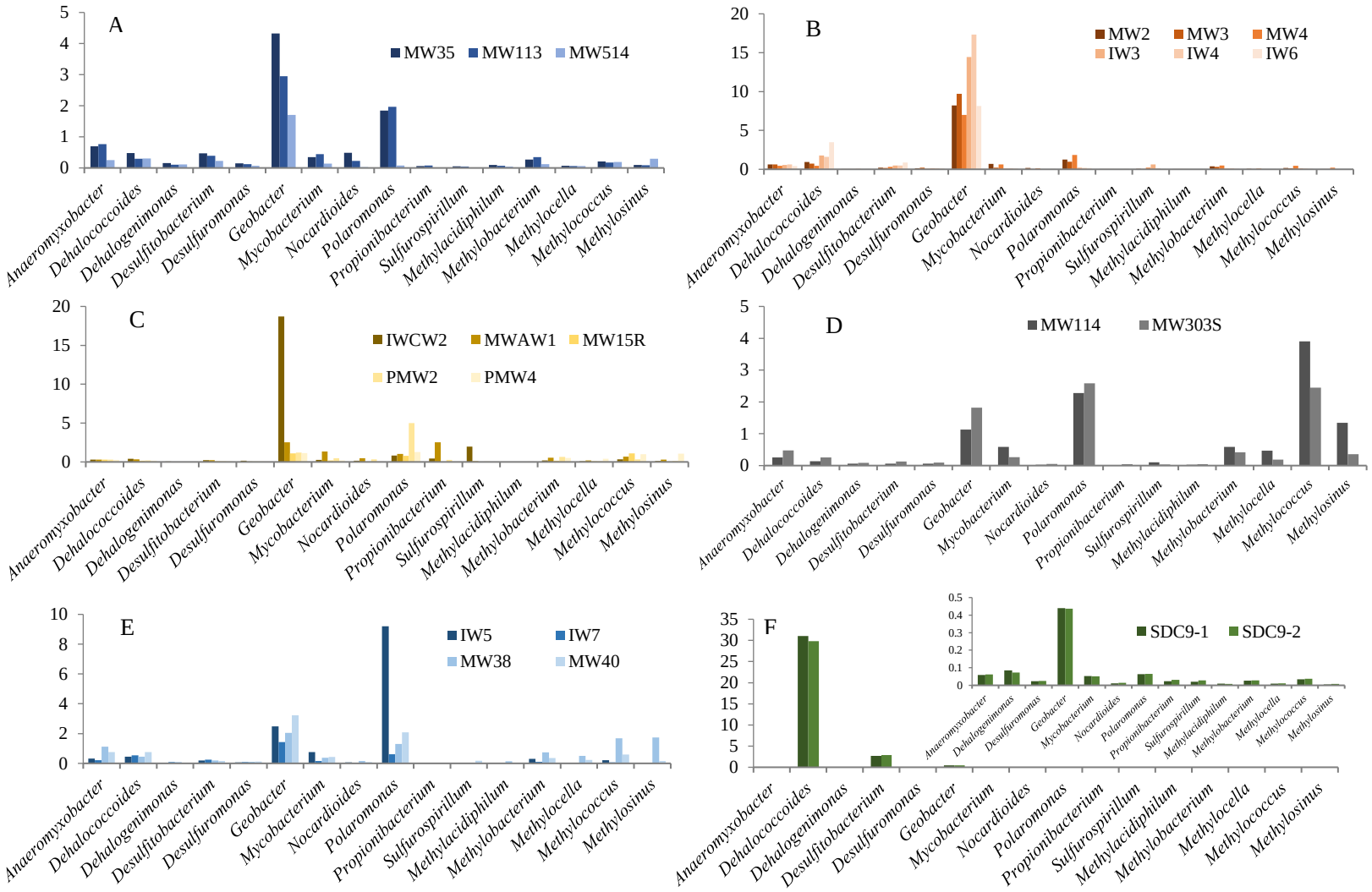


Figure 2. 1. Relative abundance (%, as determined using MG-RAST) of methanotrophs and genera associated with chlorinated solvent biodegradation in groundwater from San Antonio (A), Tulsa (B), Quantico (C), Edison (D), Indian Head (E) and SDC-9 (F). The genus *Dehalococcoides* was present in all groundwater samples ranging from 0.1 – 3.5%. Note,"MW" in name refers to a monitoring well and "IW" in name refers to an injection well. The insert in F does not include *Dehalococcoides* or *Desulfitobacterium* to enable a y-axis with a different scale.

The relative abundance of methanotrophs in the groundwater samples was also investigated (Figure 2.1). Methanotrophs are important because of their ability to use particulate and soluble methane monoxygenases (pMMO and sMMO) to cometabolically oxidize several chlorinated solvents (39-41).

Dehalococcoides, the key dechlorinating genera in SDC-9 (31% in SDC-9), was detected in every sample at every site (averages for each site ranging from 0.2 to 1.4%). The sites had been bioaugmented with SDC-9 from ~ 6.5 months (Quantico) to more than 6 years (Edison) prior to groundwater sample collection (Supplementary, Table 1). The abundance of *Dehalococcoides* was greater in the injection wells (IW3, IW4, IW5, IW, CW2) compared to the monitoring wells (Figure 2.1B, C). *Dehalococcoides* relative abundance levels (0.14-0.26%) were lowest at the Edison site (Figure 2.1D) which had the longest time between bioaugmentation and sample collection (76 months). The lower *Dehalococcoides* levels at the Quantico site (0.15-0.19%, Figure 2.1C) are puzzling, since it had the shortest time between bioaugmentation and sampling (6.5 months), and may be related to the electron donor utilized (hydrogen compared to a fermentable substrate). At the Tulsa site, *Dehalococcoides* relative abundance levels were on the higher side (monitoring wells, 0.44 -0.96%, Figure 2.1B), perhaps as a result of higher TCE concentrations at the time of sampling (Supplementary Figure 2.12). *Dehalococcoides* abundance levels were also higher at the Indian Head site (0.40-0.75%), possibly related to a shorter time between bioaugmentation and sampling (9 months).

Desulfitobacterium was detected at all five sites, although the relative abundance (average ranging from 0.1 to 0.4%) was typically less than that of *Dehalococcoides*. Except for *Dehalococcoides*, *Desulfitobacterium* was present at a higher relative abundance in SDC-9 (2.7%) compared to other dechlorinating microorganisms (<0.4%). At three sites, *Geobacter* was

the most abundant genus in this group (Figure 2.1A, B and C) and at two sites, it was either the second or third most abundant (Figure 2.1D and E).

The five methanotrophs examined were present only at low levels in SDC-9 (averages ranging from 0.006-0.035%). In the groundwater samples, *Methylococcus* or *Methylosinus* were typically the most abundant, followed by *Methylobacterium* and *Methylocella*.

3.3 Functional Gene Analysis

The groundwater sequencing data were aligned to characterized RDases from *D. mccartyi* and three other genera (*Dehalogenimonas*, *Dehalobacter* and *Desulfitobacterium*) (Figure 2.2). Not surprisingly, RDases from *D. mccartyi* were the most abundant (Figure 2.2A). Samples from Tulsa illustrated some of the highest values for *tceA* and *vcrA*, again a pattern perhaps caused by the higher chlorinated ethene concentrations at this site (Supplementary Table 2.1, Supplementary Figure 2.12). Following Tulsa, the wells at Indian Head contained the second most abundant reads aligning to RDases from *D. mccartyi*. These results agree with the MG-RAST analysis, which illustrated the highest relative abundance of *Dehalococcoides* at Indian Head and Tulsa (Figure 2.1B and E).

The abundance of RDases from *Dehalogenimonas*, *Dehalobacter* and *Desulfitobacterium* were found in lower numbers and the results varied between sites (Figure 2.2B, C, D). The majority of reads aligning with *cerA* and *tdrA* from *Dehalogenimonas* were from Tulsa (MW2, MW3, MW4, IW3, IW4, IW6), followed by Indian Head (IW5, IW7, MW38, MW40) and Edison (MW303S) (Figure 2.2B). The average relative abundance values for *Dehalogenimonas* from the MG-RAST analysis indicated the highest values for San Antonio, Edison and Indian Head (Figure 2.1A, D, E). Reads aligning to RDases from *Dehalobacter* and *Desulfitobacterium* were less abundant but were found in at least one well from three of the five sites (except San

Antonio and Edison) (Figure 2.2C, D). Although *Desulfitobacterium* was detected with the MG-RAST analysis, *Dehalobacter* was not.

Additional differences between the MG-RAST and the functional gene data sets included the presence of the genera *Anaeromyxobacter* and *Sulfurospirillum* with MG-RAST, but the absence of functional genes (associated with the removal of chlorinated chemicals) from these microorganisms. Also, *Geobacter* and *Polaromonas* were present at all sites, however, reads aligning to *pceA* of *Geobacter lovleyi* and cytochrome P450 of *Polaromonas JS666* were observed from only one sample each (MW40 and MW4, respectively, data not shown). These findings emphasize the importance of functional gene analysis to clearly define *in situ* potential biodegradation capabilities.

The majority of the RDases found in SDC-9 were from *D. mccartyi,* with *tceA* and *vcrA* being the most abundant (~two orders of magnitude higher than the RDases from other species) (inserts in Figure 2.2). RDases from *Dehalogenimonas, Dehalobacter, Desulfitobacterium* were also present in SDC-9.

Reads aligning to the genes associated with the aerobic degradation of 1,4-dioxane (42) were also investigated (Figure 2.3). From the twelve genes examined, only six were identified in the groundwater samples (Figure 2.3A). These genes were detected in at least one sample from all five sites, despite the fact that only one of the sites (Tulsa) was known to be contaminated with 1,4-dioxane. Surprisingly, no genes associated with *Pseudonocardia* were detected. The MG-RAST taxonomic data were examined for the presence of the genera associated with these genes

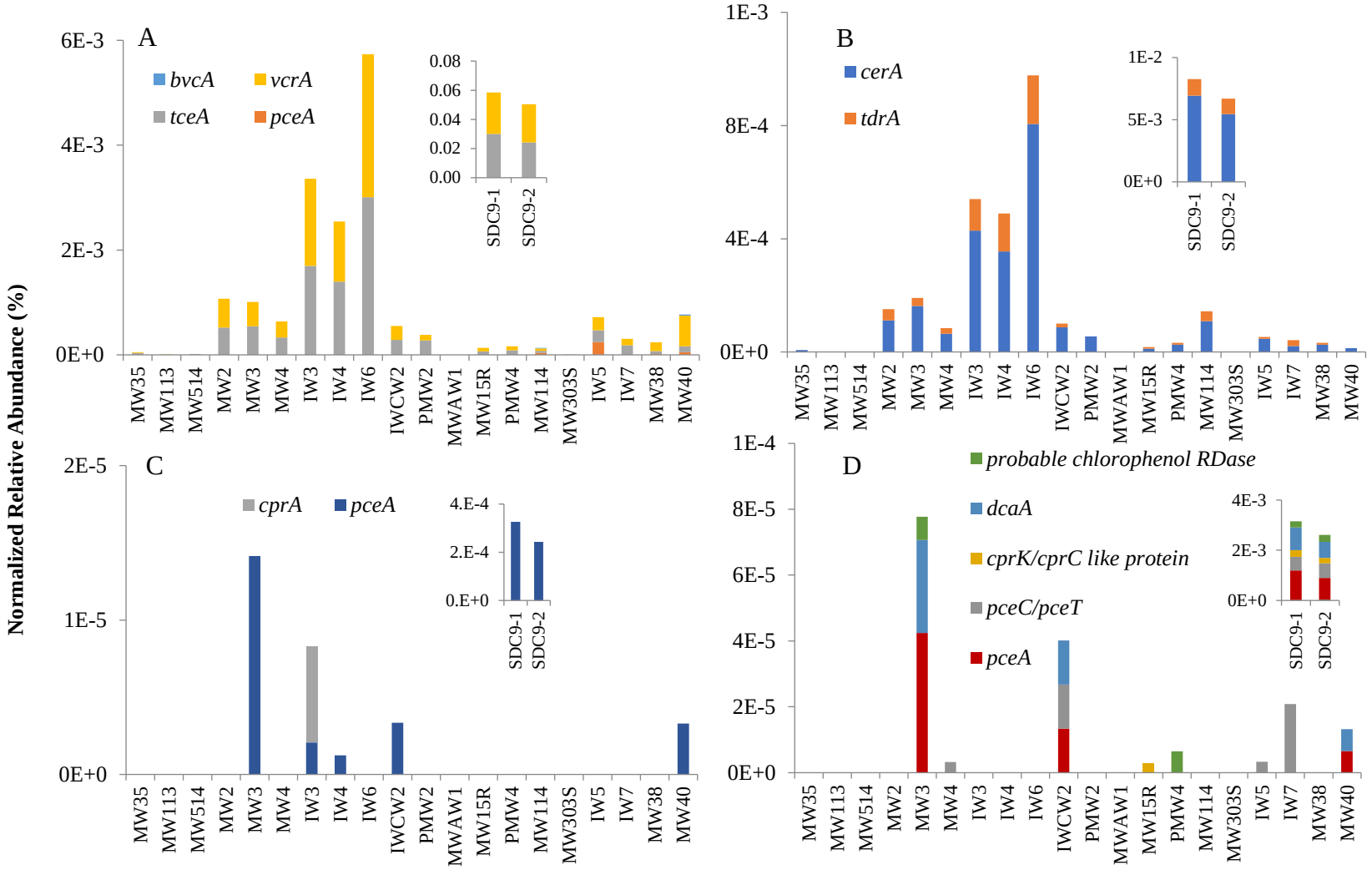


Figure 2. 2. Normalized relative abundance (%, as determined by DIAMOND) of genes associated with reductive dechlorination in *Dehalococcoides mccartyi* (A), *Dehalogenimonas* spp. (B), *Dehalobacter* spp. (C) and *Desulfitobacterium* spp. (D) in SDC-9 (inserts) and in groundwater from the five chlorinated solvent sites. The highest abundance values are from *tceA* and *vcrA* from *Dehalococcoides*, followed by *cerA* from *Dehalogenimonas*.

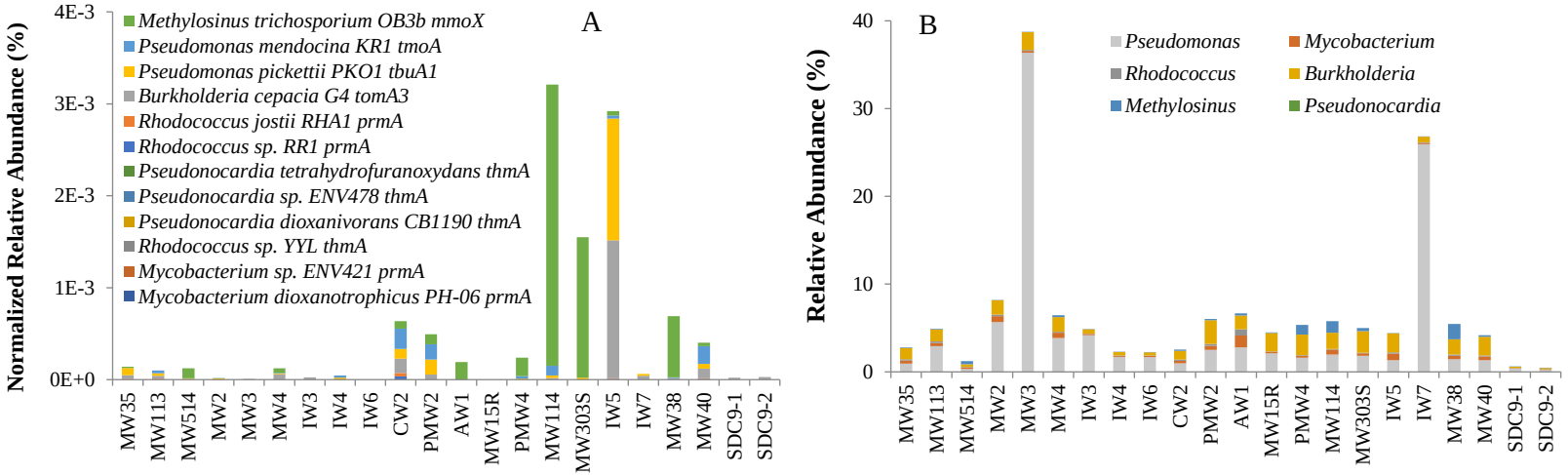


Figure 2. 3. Normalized relative abundance (%, determined with DIAMOND) of genes (A) and relative abundance (%, determined with MG-RAST) of genera (B) previously associated with 1,4-dioxane degradation in all groundwater samples and in SDC-9. The relative abundance of *Pseudonocardia* was zero in all groundwater samples and in SDC-9. *Methylosinus trichosporium* OB3b *mmoX* was the dominant 1,4-dioxane degrading gene in the majority of the groundwater samples.

(Figure 2.3B). From this group, *Pseudomonas* was the most dominant genus, followed by *Burkholderia, Mycobacterium, Methylosinus* and *Rhodococcus*. Similar to the functional gene data, the genus *Pseudonocardia* was not detected in any groundwater sample.

The shotgun data sets were also queried against reference databases that contained both RDases from complete genomes as well as those from uncultured microorganisms (Figure 2.4A). The results were consistent with those found using sequences from complete genomes only (Figure 2.2A). Reads aligning with the genes associated with the aerobic degradation of the chlorinated ethenes (*pmoA, mmoX, etnC, etnE)* (40, 41, 43) were detected in the groundwater samples from a number of samples from Edison and Indian Head (Figure 2.4B, C). Additionally, *etnC* and *etnE* were also found at high levels in the monitoring wells from the Tulsa site, again perhaps as a result of higher cVOC concentrations at the time of sampling. Notably, the highest normalized relative abundance values for *etnC* and *etnE* were two orders of magnitude lower than *vcrA* or *tceA*.

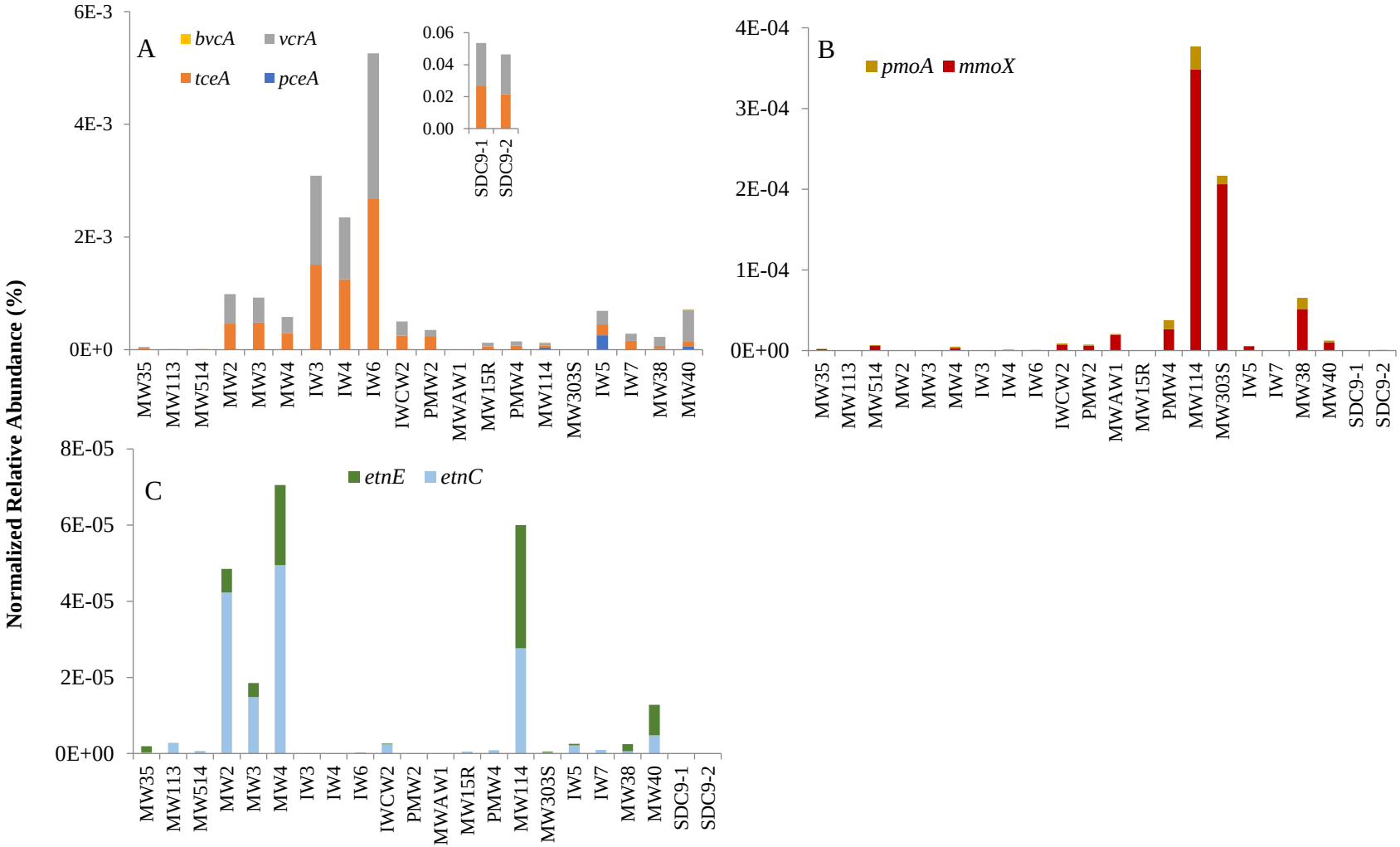


Figure 2. 4. Normalized relative abundance (%, determined with DIAMOND) of genes associated with the chlorinated solvent reductive dechlorination (A) and the aerobic degradation of the chlorinated solvents (B, C) in SDC-9 (insert for A) and in groundwater from the five chlorinated solvent sites. The aerobic genes occurred at lower levels compared to the anaerobic genes. Note, the analysis approach differed from the approach used to generated Figure 2.2, in that all sequences from the databases were compared to each dataset.

The DIAMOND analysis included alignments to a gene encoding for a formate dehydrogenase-like protein (*fdhA*), hydrogenase genes (*hup, vhc, hym* and *ech*) and corrinoid metabolism genes (*btu, cbi* and *cob*) from *D. mccartyi*. In previous research, the formate dehydrogenase-like protein was found to be highly expressed and ubiquitous in *D. mccartyi*, representing a specific indictor for activity (44). Hydrogenases are thought to oxidize H_2, the electron donor for *D. mccartyi* (45). Corrinoid metabolism genes are relevant for up-taking and transforming of cobamides and cobinamide, which are critical for *D. mccartyi* RDases (45). Samples containing the most abundant reads of *fdhA* were from Tulsa following by samples from Indian Head (Supplementary Figure 2.23). The abundance patterns for the hydrogenase and corrinoid metabolism genes across samples were similar to those for *vcrA, tceA* and *fdhA* (Supplementary Figures 2.24 and 2.25). The *fdhA* abundance patterns across samples were similar to those observed for *tceA* and *vcrA* (Spearman's rank correlation coefficients 0.939 and 0.89 for *fdhA* vs. *vcrA* and *fdhA* vs. *tceA*, respectively, p values both < 0.0001, Supplementary Figure 2.26), indicating this gene acts as an effective biomarker for *D. mccartyi*.

To investigate the accuracy of the shotgun sequencing data quantification method, the relative abundance of *vcrA* determined via shotgun sequencing was compared to *vcrA* gene copies determined via qPCR (Supplementary Figure 2.27). In general, the abundance of *vcrA* determined using shotgun sequencing correlated well (Spearman's rank correlation coefficient 0.808, p value < 0.0001) with the qPCR data (3.9×10^4 to 7.0×10^9 *vcr* gene copies per L).

Principal component analyses were completed for the functional genes (Figure 2.5A) and genera (Figure 2.5B) associated with chlorinated solvent and 1,4-dioxane biodegradation. The genes *tdrA, vcrA* and *tceA* were positively correlated to *fhdA* as well as the hydrogenase and corrinoid metabolism genes, consistent with their similar abundance distribution in the wells.

These genes correlated with injective wells from the Tulsa site, which would be expected considering the high relative abundance of *Dehalococcoides* in these samples. Genes relevant to aerobic chlorinated ethene degradation correlated with *mmoX* (from *M. trichosporium* OB3b) suggesting the genetic potential for degradation of these co-contaminants occurs at the same site. In this case, the genes correlated with MW114 from the Edison site. The remaining genes associated with 1,4-dioxane degradation correlated together (bottom left quadrant) perhaps indicating multiple functional genes will contribute to 1,4-dioxane degradation at the same site. RDases (*pceA*) from *Desulfitobacterium* and *Dehalobacter* also correlated together, along with MW3 (from Tulsa) which was previously found to contain RDases from these genera. For the taxonomic principal component analysis (Figure 2.5B), the anaerobic genera *Dehalococcoides, Desulfitobacterium and Desulfuromonas* correlated together along with the injection wells from the Tulsa site. For the methanotrophs, *Methylococcus* and *Methylobacterium* illustrated a positive correlation to each other and to the wells from several sites e.g. MW114, MW303, PMW4. The genera PCA is less meaningful because it is unknown if the majority of these microorganisms are truly associated with contaminant degradation.

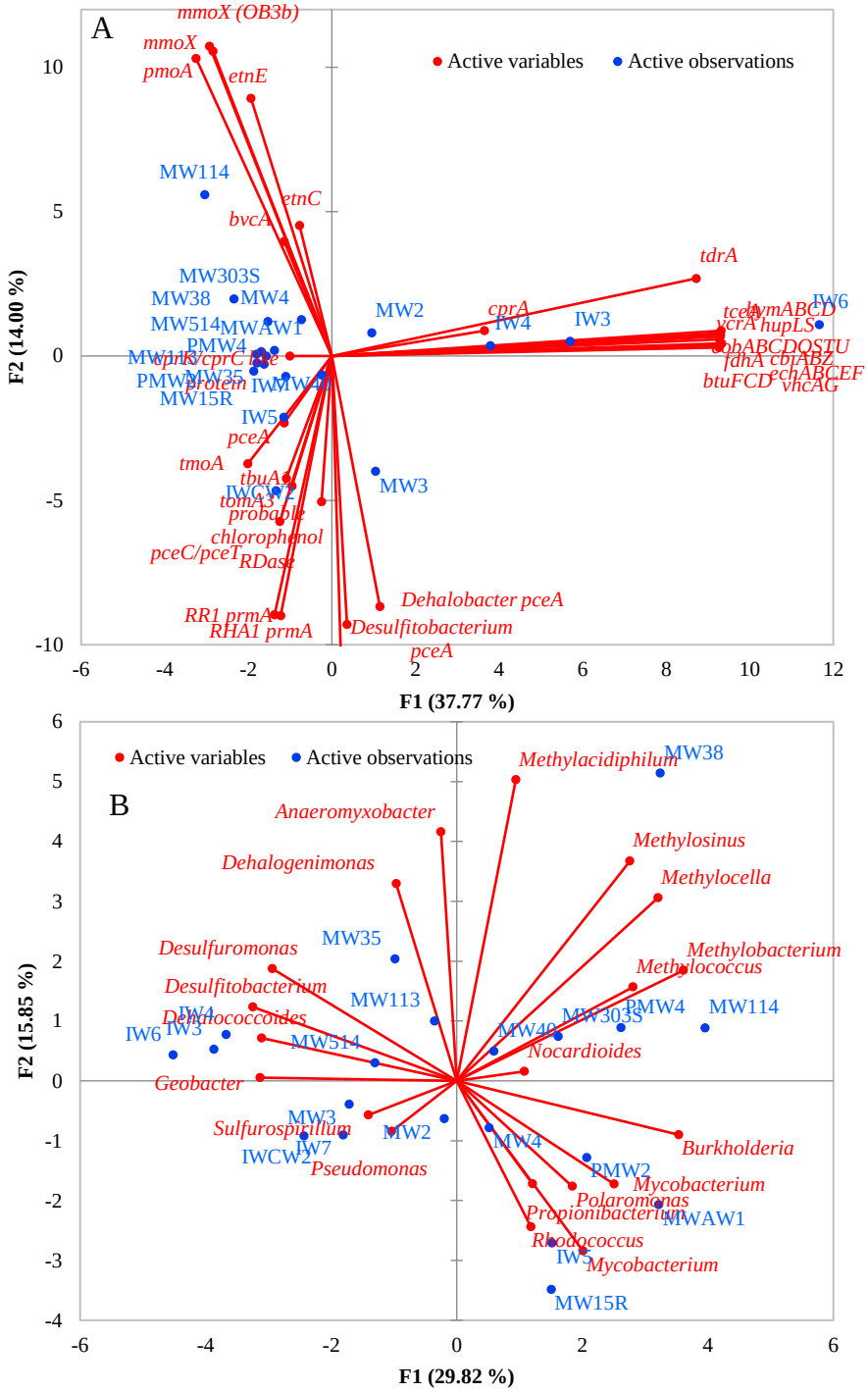


Figure 2. 5. Principle component analyses of functional genes (A) and genera (B) associated with chlorinated solvent and 1,4-dioxane biodegradation in all groundwater samples.

4. Discussion

Biostimulation and bioaugmentation are becoming increasingly popular approaches for the remediation of groundwater contaminated with PCE, TCE and their daughter products. However, limited research has focused on groundwater microbial communities post bioaugmentation. This work is important because of the requirement for *Dehalococcoides* to co-exist with other "supporting" microorganisms and to survive over time. Further, it is also valuable to determine if other chlorinated solvent degrading microorganisms are present, and the extent to which these organisms persist following bioaugmentation with exogenous strains.

Not surprisingly, the genus *Dehalococcoides* was a major component of SDC-9. This was also reported for another common bioaugmentation culture, KB-1 (29). Of additional interest is the presence (4% relative abundance) of *Desulfitobacterium* in SDC-9, as this genus has also been associated with dechlorination (46-50). Similarly, others have reported *Desulfitobacterium* type RDase genes in *Dehalococcoides* enriched cultures (29). Other genera linked to chlorinated solvent degradation were also detected in SDC-9 (as discussed above); however, their relative abundance in the culture was low compared to *Dehalococcoides* or *Desulfitobacterium.*

As in other *Dehalococcoides* enrichment cultures, SDC-9 contained methanogens (*Methanocorpusculum*), acetogens (*Clostridiaceae*) and *Geobacter* (9, 29). *Geobacter* has previously been associated with interspecies corrinoid transfer with *Dehalococcoides* (51). In addition, *Geobacter* has also been associated with dechlorination (52, 53). The genera *Thermosinus* and *Selenomonas* within the family *Veillonellaceae* were detected in SDC-9 at low levels (3% and 0.4%, respectively). *Veillonellaceae* were previously found to be important corrinoid supplying microorganisms to *Dehalococcoides* in another

enrichment culture (54). SDC-9 contained *Desulfovibrio* (2.5%), which, in previous research, was linked to more robust dechlorination rates and growth when grown in co-culture with *Dehalococcoides* (55). It was reported that *Desulfovibrio* can support *Dehalococcoides* by providing acetate, hydrogen and corrinoid cofactors (55).

Following *Dehalococcoides* and *Methanocorpusculum*, the third and fourth most abundant genera in SDC-9 were *Bacteroides* (5.4%) and *Parabacteroides* (10%) (within *Bacteroidetes*). Members of the *Bacteroidetes* phylum have also been reported as important bacteria in other dechlorinating mixed cultures (29) and in contaminated groundwater (9).

For the groundwater samples, *Geobacter* was more abundant at all sites compared to the SDC-9 culture and may therefore be important in playing a supportive role for *Dehalococcoides* at contaminated sites. In contrast, *Thermosinus* and *Selenomonas* were not detected in any groundwater samples. Other potentially supportive microorganisms, including *Desulfovibrio, Bacteroides* and *Parabacteroides*, were detected in the groundwater at all sites (ranging from 0.1- 4.2%) and therefore may also play a supportive role for *Dehalococcoides in situ*.

Similar to many previous studies examining microbial communities at chlorinated solvent sites (undergoing some kind of bioremediation), the genera *Dehalococcoides, Dehalogenimonas* and *Geobacter* were found in groundwater from all five sites (9, 19-27). The current study identified *Desulfitobacterium* and *Anaeromyxobacter* in the majority of samples and these genera have also been frequently detected at contaminated sites (17, 20-22, 24). In contrast, fewer previous studies have reported the presence of *Polaromonas* and *Nocardioides* (16, 17, 25). Previous researchers have also detected methanotrophs *in situ* (16, 18, 24). It was surprising that *Dehalobacter* was absent in the MG-RAST data, as this genus has been commonly reported in groundwater from chlorinated solvent contaminated sites (17, 19, 21, 22, 24). However, *cprA*

and *pceA* from *Dehalobacter* were found in the functional gene analysis, suggesting this genus could be present, but at levels undetectable by the MG-RAST analysis.

Although taxonomic data is important for characterizing microbial communities *in situ*, it is well recognized that certain limitations are associated with such data. A key limitation concerns an inability to classify to the species level when short sequences are analyzed. This issue is particularly relevant to bioremediation applications, as it impacts an identification of a known degrader, e.g. *Polaromonas* JS666 (56) or *Geobacter lovleyi* (57), over others in the same genus that are not capable of contaminant degradation. In the current study, relying on taxonomic data alone would have been misleading, because although the genera *Polaromonas* and *Geobacter* were present, the functional genes were largely absent (*P450* from *Polaromonas* JS666 and *pceA* from *Geobacter* were detected only once). Another related limitation concerns the inability of taxonomic data to provide in-depth information on function. This concern is important when considering *D. mccartyi*, as strains with similar 16S rRNA gene sequences may contain different RDases. Clearly, to generate a full picture of the functional abilities of microorganisms to degrade contaminants *in situ*, both taxonomic and functional analyses are needed.

The taxonomic and functional analysis detected both aerobic and anaerobic biomarkers across the five sites. For example, both *vcrA* and *etnC* were found in MW2, MW3, MW4 from the Tulsa site (although the values for *vcrA* were higher). This trend has previously been noted for groundwater from other chlorinated solvent sites (13, 16, 58). The genes *etnC* or *etnE* were detected in at least one groundwater sample from each site, with the normalized relative abundance values covering a wide range. Similarly, *pmoA* and *mmoX* were detected in at least one groundwater sample from each site and

were particularly abundant at the Edison site. Given the occurrence of these genes in the current study, future research directions should include a consideration of both aerobic and anaerobic genes when accounting for chlorinated solvent removal rates.

To our knowledge, this study represents the first analysis of the genes associated with 1,4-dioxane degradation in groundwater using shotgun sequencing. Here, from the twelve sequences investigated, the most abundant number of reads (collectively, in all groundwater samples) aligned to *Methylosinus trichosporium OB3b mmoX,* followed by *Burkholderia cepacia G4 tomA3* and *Pseudomonas pickettii PKO1 tbuA1.* Notably, although *mmoX* from *M. trichosporium OB3b* has been associated with 1,4-dioxane degradation at high concentrations (59), at low, environmental relevant concentrations, no removal was observed (60). Three others (*Pseudomonas mendocina KR1 tmoA, Rhodococcus jostii RHA1 prmA, Rhodococcus sp. RR1 prmA*) were detected at lower levels in at least one well from each site. In some cases, remarkably, the normalized relative abundance values were in the same range as those for *vcrA* and *tceA,* even though 1,4-dioxane was not previously reported at 4 of the 5 sites, and reducing conditions (i.e., negative oxidation-reduction potential; nORP) generally prevailed. Previously, others have observed *thmA* in samples from 1,4-dioxane contaminated sites using qPCR primers designed to *thmA* from *Pseudonocardia* (61-63). However, reads aligning to *thmA* were not found in the current study. The taxonomic data confirmed this finding, as the genus *Pseudonocardia* was absent from the MG-RAST results. Reads aligning to *Mycobacterium* 1,4-dioxane degrading gene sequences (*prmA*) were also not detected in the current study, even though the taxonomic MG-RAST data indicated this genus was present. This discrepancy again illustrates the importance of functional gene data to corroborate taxonomic data and assumptions about function. Further, the current work illustrates the importance of shotgun sequencing to

provide a more complete picture of the potential of *in situ* microbial communities to degrade 1,4-dioxane compared to qPCR, which typically only targets a small number of genes.

Previous research indicated that transcripts of the proteins Fdh and Hup may be better indicators of cell respiration compared to RDases (64, 65). In fact, it was concluded that HupL transcripts were the most robust activity biomarker across multiple *D. mccartyi* strains (66). Given importance of Hup, the relative abundance of *fdhA* and other genes encoding for hydrogenases from *D. mccartyi* were investigated in the groundwater samples. Building on the approach developed in the current study, future research could include shotgun sequencing of transcripts to obtain an improved indicator of *D. mccartyi* cell respiration. These gene targets, as well as those involved in corrinoid metabolism, could be used as additional biomarkers for *D. mccartyi*.

To examine the quantitative robustness of the data generated, the normalized relative abundance values for *vcrA* were compared to those obtained via TaqMan qPCR. The correlation indicated the methods produced similar values across a range of concentrations for the five sites. Two important future research directions for using shotgun sequencing for bioremediation applications will be 1) to determine detection limits and 2) to generate more in depth comparisons to values determined with qPCR.

In summary, methods were developed to determine the abundance of genes associated with chlorinated solvent and 1,4-dioxane biodegradation in groundwater samples from multiple samples from multiple contaminated sites. The use of shotgun sequencing enabled a larger selection of genes to be targeted compared to traditional qPCR. In fact, the number of functional genes that can be analyzed is limitless. The

method also does not require primer design or primer assay verification for each target (as is the case for qPCR). The most labor-intensive part of the approach involved the collection of reference fasta files for the DIAMOND alignment (following this, all remaining steps were not time consuming). The sequencing price is perhaps the largest limitation to the method. In the current study, for 22 samples, the cost was approximately $210 per sample. However, it is likely that sequencing costs will drop as the technology evolves, making the approach more attractive. The data indicated the presence of both aerobic and anaerobic biomarkers for chlorinated solvent degradation. Not surprisingly, the taxonomic data alone was insufficient to determine the functional abilities of these communities. The relative abundance of hydrogenases and corrinoid metabolism genes suggest these may be appropriate additional biomarkers for *D. mccartyi.* The approach developed will enable researchers to investigate the abundance of any contaminant degrading gene in any sample, greatly expanding the analytical toolbox for natural attenuation, biostimulation or bioaugmentation.

Acknowledgements

Thanks to Simon Vainberg, Sheryl Streger, Robert E. Mayer, Michael Martinez and David Lippincott from APTIM Federal Services for providing groundwater samples. Thanks to James Cole and Benli Chai from RDP (MSU). Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government. Views, opinions, and/or findings contained in this report are those of the authors and should not be construed as an official Department of Defense position or decision unless designated by other official documentation.

APPENDIX

APPENDIX

Text A1. Background on Functional Genes

The abundance of the genes associated with reductive dechlorination were examined from the genera *Dehalococcoides, Dehalogenimonas, Desulfitobacterium, Dehalobacter, Geobacter, Sulfurospirillum* and *Anaeromyxobacter*. The genes encoding for enzymes associated with aerobic VC degradation were also targeted (*etnC*/alkene monooxygenase and *etnE*/epoxyalkane:CoM transferase) (67-71). Also, the gene encoding for cytochrome P450 from *Polaromonas* JS666 was investigated, as this initializes the degradation of *cis*-1,2-dichoroethene (72). The genes encoding for the α subunits of soluble and particulate methane monooxygenases (*mmoX* and *pmoA*) were examined due to their role in chlorinated ethene degradation (41, 73, 74). The genes encoding the enzymes associated with 1,4-dioxane biodegradation (as summarized in (42)) were also investigated. This chemical is a probable human carcinogen (75) and is frequently detected at sites where the chlorinated solvents are present (76-78). Finally, the genes encoding for enzymes associated with hydrogen metabolism (*fdhA, hupSL, vhcAG, hymABCD* and *echABCEF*) and corrinoid metabolism (*btuFCD, cobA, cobB, cobC, cobD, cobQ, cobS, cobT, cobU, cbiA, cbiB* and *cbiZ*) in *D. mccartyi* were also quantified (45).

Methods

Text A2. Collection of Sequences for Functional Genes

A primary source for RDases was the Functional Gene Pipeline and Repository (FunGene) (36) website using the link ‘vcrA_ver2’. For this, the contents (e.g. score, protein and nucleotide accession numbers, microorganism name, length of the protein) of the topmost 3000 sequences from ‘vcrA_ver2’ were exported into excel. Then, accession number lists of RDases for each

species were created by setting a filter in Excel. Each accession number was checked for accuracy using complete genome information (NCBI) (Supplementary Table 2.3). The protein sequence fasta files were also downloaded from FunGene (again by selecting the same 3000 sequences of 'vcrA_ver2'). The accession number RDase lists were used to collect RDases reference protein sequences from the protein sequence fasta files (from FunGene) using a tool (Readseq.jar) developed by Ribosomal Database Project (https://github.com/rdpstaff/RDPTools). The overall process produced individual RDase protein sequence files for each microorganism.

A number of microorganisms (including *Dehalococcoides mccartyi JNA, SG1, Sulfurospirillum strains, Anaeromyxobacter dehalogenans 2CP-C, Geobacter lovleyi SZ, Dehalobacter E1 and FTH1, Desulfitobacterium sp. PCE1, Desulfitobacterium hafniense TCP-A and PCP-1* and *Polaromonas sp. JS666*) did not have functional gene data on FunGene, therefore their reference sequences were collected manually by downloading the fasta files from NCBI complete genomes (Supplementary Table 2.3).

The *vcrA* reference sequences (39 *vcrA* sequences, protein fasta files) were collected from the link 'vcrA' in FunGene by selecting those sequences with a score higher than 900 (Hidden Markov Model score alignment by FunGene). The *tceA* and *bvcA* reference sequences were collected using both NCBI and FunGene. Protein sequences (AAN85588, AAT64888) previously used for designing *tceA* and *bvcA* primers (38) were first used to collect sequences from the NCBI database using BLAST (79). Sequences with a maximum score higher than 900 from the BLAST search were collected (31 *tceA* sequences, 13 *bvcA* sequences). In FunGene, using the 'Probe Match Search' function, primers for *tceA* (TceA1270F, TceA1336R and TceA1294Probe) and *bvcA* (Bvc925F, Bvc1017R and Bvc977Probe) (38) were used to search the first 3000 sequences of 'vcrA_ver2', producing 38 *tceA* sequences and 11 *bvcA* sequences.

The *tceA* and *bvcA* sequences from the two sources were compared. The *tceA* sequences from NCBI (except ADV18463) were all present in the *tceA* sequences obtained from FunGene. Therefore, the final *tceA* reference list consisted of the sequences from FunGene along with ADV18463. A similar approach was used for generating the *bvcA* reference list.

Sequences for *pceA* (5 sequences) and *fdrA* (30 sequences) from *Dehalococcoides mccartyi* were collected by downloading the fasta files from the NCBI complete genomes. The sequence for the putative VC RDase (*cerA*) from *Dehalogenimonas* (80) was kindly provided by Dr. Frank Loeffler (Locus Tag JP09_004725, Protein ID PPD58423.1).

Reference sequences for *etnC* and *etnE* (31 *etnC* sequences, 95 *etnE* sequences) were collected from FunGene using scores higher than 700 and 500, respectively. Additionally, primers for *etnC* (RTC_F (*etnC*) and RTC_R (*etnC*)) and *etnE* (RTC_F (*etnE*) and RTC_R (*etnE*)) (43) were used with the 'Probe Match Search' function in FunGene to search for sequences in all pages of 'etnC' and 'etnE', resulting in 9 *etnC* sequences, 31 *etnE* sequences. Reference sequences for *etnC* and *etnE* (20 *etnC* sequences, 53 *etnE* sequences) were also collected from UniProt. The final *etnC* and *etnE* reference sequences were generated by combining all data sets discussed above.

mmoX and *pmoA* reference sequences (21 *mmoX* sequences, 30 *pmoA* sequences) were first collected using 'mmoX' and 'pmoA' links in FunGene (sequences with a score higher than 980 and 500, respectively). Additionally, all other sequences annotated as 'mmoX' or 'soluble methane monooxygenase' in all pages of *mmoX* in FunGene were also collected. For this, information, such as score, protein and nucleotide accession number, name of the microorganism, length of the protein, was imported to excel. Then, a filter in excel was set for the name of the gene to create an accession number list for *mmoX*. The accession number list was

used to collect reference protein sequences from protein sequences downloaded from FunGene using Readseq.jar (generating 580 sequences). Sequences annotated as '*pmoA*' or 'particulate methane monooxygenase' were also collected using methods similar to those described from *mmoX* (generating 8327 sequences).

A list of functional genes (12 sequences) associated with 1,4-dioxane metabolism or cometabolism was obtained from a recent publication (42). The protein sequences of these genes were then collected from NCBI. The functional genes associated with hydrogen and corrinoid metabolism in *D. mccartyi* were also examined. Reference sequences for all hydrogenase (*hupLS*, *vhcAG*, *echABCEF* and *hymABCD*) and corrinoid (*btuFCD*, *cbiABZ* and *cobABCDQSTU*) metabolism genes were collected by using NCBI BLAST search. Additional information on the collection of sequences associated with hydrogen and corrinoid metabolism is provided in the supplementary section.

Sequences ACZ61293.1 and ACZ61294.1 from *D. mccartyi* VS were used for starting the BLAST search for *hupL* and *hupS*, separately. Then *hupL* and *hupS* reference sequences (13 *hupL* sequences, 6 *hupS* sequences) were collected with an identity > 95% and >94%, respectively. All identity values were selected because of the large identity decrease after the last selected reference sequences. *vhcA* and *vhcG* reference sequences (9 *vhcA* sequences, 8 *vhcG* sequences) were collected with an identity > 90%. The sequences used for the BLAST search were ACZ61705.1 and ACZ61704.1 from *D. mccartyi* VS. *hymA1* and *hymA2* reference sequences (4 *hymA1* sequences, 5 *hymA2* sequences) were collected with an identity > 98% and >96%, respectively. The sequences used for starting the BLAST search were ACZ61326.1 and ACZ61777.1 from *D. mccartyi* VS. *hymB1* and *hymB2* (3 *hymB1* sequences, 19 *hymB2* sequences) were collected with an identity > 98% and > 97%, respectively. The sequences used

for starting the BLAST search were ACZ61327.1 and ACZ61778.1 from *D. mccartyi* VS. *hymC1* and *hymC2* (15 *hymC1* sequences, 15 *hymC2* sequences) were collected with an identity >96% and >87%, respectively. The sequences used to start the BLAST search were ACZ61328.1 and ACZ61779.1 from *D. mccartyi* VS. *hymD1* (11 *hymD1* sequences) was collected with an identity > 89%. The sequence used to start the BLAST search was ACZ61329.1 from *D. mccartyi* VS.

Additional *hymABC* genes were found in *D. mccartyi* 195 and following the similar nomenclature for the genes, they were named *hymA3, A4, B3* and *C3*. *hymA3* and *hymA4* (9 *hymA3* sequences, 13 *hymA4* sequences) were collected with identities > 98% and > 93%, respectively. The sequences for the BLAST search were AAW39863.1 and AAW40249.1. *hymB3* (18 sequences) was collected with an identity > 94%. The sequence used for starting the BLAST search was AAW39862.1 *hymC3* (13 sequences) was collected with an identity > 90%. The sequence used for starting the BLAST search was AAW39861.1

The sequences used for starting BLAST search for *echABCEF* were from *D. mccartyi* CBDB1 with accession number of CAI82985.1, CAI82986.1, CAI82987.1, CAI82992.1 and CAI82993.1. *echABCEF* reference sequences (23 *echA* sequences, 16 *echB* sequences, 7 *echC* sequences, 10 *echE* sequences, 11 *echF* sequences) were collected with an identity > 92%, 94%, 94%, 96% and 84%, respectively. The sequences used for starting BLAST search for *btuFCD* were from *D. mccartyi* DCMB5 with accession number of AGG06280.1, AGG06281.1 and AGG06282.1. *btuFCD* reference sequences (17 *btuF* sequences, 14 *btuC* sequences, 6 *btuD* sequences) were collected with an identity > 89%, 93% and 93%, respectively.

The sequences used for starting BLAST search for *cbi* were from *D. mccartyi* VS. *cbiA* and *cbiB* reference sequences (16 *cbiA* sequences, 13 *cbiB* sequences) were collected both with

an identity > 90%. The sequences used for starting the BLAST search were ACZ61308.1 and ACZ61741.1.

There were four *cbiZ* sequences from *D. mccartyi* VS (hereafter named *cbiZ1234*). The accession number of the sequences of *cbiZ1234* for the BLAST search were ACZ61242.1, ACZ61249.1, ACZ61740.1 and ACZ62455.1. *cbiZ1234* reference sequences (16 *cbiZ1* sequences, 11 *cbiZ2* sequences, 10 *cbiZ3* sequences, 46 *cbiZ4* sequences) were collected with an identity > 72%, 97%, 92% and 87%, respectively.

The majority of the sequences used for starting BLAST search for *cob* were from *D. mccartyi* VS, with one from *cobC* from *D. mccartyi* CBDB1. The accession number of sequences used for starting BLAST search for *cobA123* were with of AAW40449.1, AAW39561.1 and AAW39547.1. cobA123 reference sequences (8 *cobA1* sequences, 14 *cobA2* sequences, 14 *cobA3* sequences) were collected with an identity > 91%, 92% and 94%, respectively.

The accession number of sequences used for starting BLAST search for *cobBCQ* were with of AAW40541.1, CAI82815.1 and AAW39791.1. *cobBCQ* reference sequences (16 *cobB* sequences, 10 *cobC* sequences, 20 *cobQ* sequences) were collected with an identity > 89%, 90% and 92%, respectively. *cobD1* and *cobD4* reference sequences (11 *cobD1* sequences, 17 *cobD4* sequences) were collected with an identity > 84% and 80%, respectively. The sequences used for starting the BLAST search were AAW40448.1 and AAW39562.1. *cobS1*, *cobT1* and *cobU1* reference sequences (14 *cobS1* sequences, 12 *cobT1* sequences, 14 *cobU1* sequences) were collected with an identity > 90%, 94% and 90%, respectively. The sequences used for starting the BLAST search were AAW40093.1, AAW40094.1 and AAW40091.1.

The BLAST search of *cobD2* and *cobD3* generated the same results as *cbiB*. Also, the BLAST results of *cobS2*, *cobT2* and *cobU2* were the same as those of *cobS1*, *cobT1* and *cobU1*, respectively. Therefore, the results of *cobD2*, *cobD3 cobS2*, *cobT2* and *cobU2* were not included in the analysis.

Text A3. Library Preparation, Sequencing, MG-RAST and DIAMOND analysis

The Takara ThruPLEX low input DNA library preparation kit was used to generate libraries based on manufacturer's recommendations. Completed libraries were subject to quality control and quantification using a combination of Qubit dsDNA HS and Caliper LabChipGX HS DNA assays. All libraries were pooled in equimolar amounts to a maximum usable volume based on quantification obtained using the Kapa Biosystems Illumina Library Quantification qPCR kit. This pool was loaded on one lane of an Illumina HiSeq 4000 flow cell and sequenced in a 2x150 bp paired end format. Base calling was performed by Illumina Real Time Analysis (RTA) v2.7.6 and output of the RTA was demultiplexed and converted to FastQ format with Illumina Bcl2fastq v2.18.0.

The Meta Genome Rapid Annotation using Subsystem Technology (MG-RAST) (35) version 4.0.2. was used for the taxonomic analysis of the metagenomes. The processing pipeline included merging paired end reads, SolexaQA (81) to trim low-quality regions from FASTQ data and dereplication to remove artificial duplicate reads. Gene calling was completed using FragGeneScan (82). For taxonomic profiles, the best hit classification at a maximum e-value of $1e^{-5}$, a minimum identity of 60% and a minimum alignment length of 15 against the ReqSeq database (83) were used. The MG-RAST plugin Krona was used to illustrate the taxonomic composition of each sample (84). MG-RAST was used to generate rarefaction curves. MG-RAST ID numbers and pre- and post- QC sequencing data have been summarized (Supplementary Table

2.2) and the datasets are publicly available on MG-RAST. The following chlorinated solvent degrading genera were investigated in the MG-RAST data: *Anaeromyxobacter* (85), *Dehalococcoides* (2, 4-6, 86, 87), *Polaromonas* (56, 72), *Nocardioides* (70, 88), *Desulfitobacterium* (47-50), *Geobacter* (52), *Sulfurospirillum* (89-91), *Dehalobacter* (92-94), *Desulfomonile* (95, 96), *Desulfuromonas* (97, 98), *Propionibacterium* (99), *Mycobacterium* (67, 100), *Dehalobacter* (93, 101), *Desulfomonile*, (102) and *Dehalogenimonas* (103-106).

DIAMOND (double index alignment of next-generation sequencing data) (37) was used as the alignment tool for all functional genes. The collected protein sequences from same species or with the same function were aligned to themselves for dereplication (removing sequences with 100% identity) and one representative sequence was left as the reference for that group. Then, low quality sequences and Illumina adapters sequences were removed using Trimmomatic (107). The shotgun reads were then aligned to the dereplicated references for each groundwater sample and SDC-9 using DIAMOND. Only reads that exhibited an identity of ≥ 90% and an alignment length ≥ 49 amino acids to the reference sequences were counted as aligned reads to each sequence. For each, relative abundance values were calculated using the number of aligned reads divided by the total number of sequences for each sample. The relative abundance values were then normalized by (divided by) the number of dereplicated reference sequences for each gene to produce normalized relative abundance values. Details concerning qPCR targeted towards *vcrA* are provided in the Supplementary Section.

Text A4. *vcrA* qPCR

The PCR tubes (20 µL reactions) contained 10 µL iTaq Universal Probe Supermix (Bio-Rad), 1.2 µL TaqMan probe (38, 108) and balance water to 18 µL. PCR amplifications were performed in three stages: 1) 95 °C for 15 min, 2) 40 cycle of 95 °C for 15 s, 58 °C for 1 min, 3) a slow

ramp of 1% to 95 °C for 15 s and 58 °C for 15 s. DNA templates and plasmid standards (containing a partial *vcrA* gene in a GenScript plasmid) were added to each reaction as 2 µL aliquots. All qPCR experiments were performed in a bench top thermal cycler (C1000 Touch Thermal Cycler, Bio-Rad).

Supplementary Table 2. 1. Groundwater and sampling data.

Site	Date Bioaugmented	Date Sampled	Months between Inoculation & Sampling	Basic Geochemistry	Carbon Source	cVOCs at time of sampling (µg/L)
Quantico, VA	12/2/2015	5/17/2016	6.5	pH 5-8 ORP: -100 to 0 Dis Fe: 0-140 mg/L	None, H_2 generated via proton reduction	VC – 0-60 Cis-DCE - 0-120
Tulsa, OK	8/2013	6/9/2015	22	pH 6.1-7.1 ORP: -64 to 248 Dis Fe: 0-110 mg/L	EOS	VC – 60-830 Cis-DCE – 110-1500 TCE – 200-9000 Trans-DCE – 3-17 1,1-DCE – 5-1400 1,2-DCA – 1-34 1,1-DCA – Trace levels (<5) 1,4-Dioxane – 78-220
Indian Head, MD	9/23/15	6/22/2016	9	pH 6.2-8.9 ORP: 37 to -326	lactate	VC – 0-29 Cis-DCE – 0-178 TCE – 0-40 Carbon Disulfide – 1-8 MEK – 0-8
San Antonio, TX	113 & 514: 10/7 & 10/8/2014 35: 10/17 & 10/30/2014	7/28/2016	20	pH 5.6-6.8 ORP: -32 to 237	EVO	VC – 2.5-8.4 Cis-DCE – 2.5-7.4 TCE – 0-2
Edison, NJ	7/8/2009	11/10/2015	76	Overall: pH 5-6.5 ORP: -100 to 3 Dis Fe: 0.7 – 13 mg/L MW-114: pH 6.0 ORP: -70 Dis Fe: 5.19 mg/L MW-303S: pH 6.1 ORP: -80 Dis Fe: 2.05 mg/L	Lactate, yeast extract, potassium bicarbonate	Overall: VC – 0-1170 Cis-DCE – 0-1190 TCE – 0-8 1,2,4-Trimethylbenzene – 0-8. Trace levels (<5) of benzene, MTBE, ethylbenzene, xylenes, isopropyl benzene, 1,3,5-trimethylbenzene, sec-butylbenzene, 1,1-DCE, tDCE MW-114: VC – 83 Cis-DCE – 70 TCE – 8 MW-303S: VC – 2J Cis-DCE – 4J

Supplementary Table 2. 2. Groundwater and SDC-9 MG-RAST sequence analysis data.

Location	Monitoring Well	MG-RAST ID #	Pre QC Sequence Count	Post QC Sequence Count	Post QC Mean Length
SDC-9	Culture-1	mgm4795922.3	6,845,624	5,090,799	236 ± 37 bp
	Culture-2	mgm4795924.3	6,181,247	4,478,198	229 ± 39 bp
San Antonio, TX	MW35	mgm4795328.3	5,301,996	4,513,530	239 ± 35 bp
	MW113	mgm4795332.3	6,185,927	5,404,716	239 ± 35 bp
	MW514	mgm4795329.3	5,934,109	4,847,401	240 ± 35 bp
Tulsa, OK	MW2	mgm4795334.3	5,691,547	4,714,786	239 ± 35 bp
	MW3	mgm4795333.3	6,872,780	5,425,995	227 ± 38 bp
	MW4	mgm4795342.3	6,200,534	5,327,773	238 ± 35 bp
	IW3	mgm4795340.3	5,889,710	4,891,993	241 ± 34 bp
	IW4	mgm4795341.3	6,938,129	5,228,938	229 ± 38 bp
	IW6	mgm4795673.3	7,800,767	6,112,693	230 ± 38 bp
Quantico, VA	MWCW2	mgm4795675.3	5,171,923	3,998,002	241 ± 35 bp
	PMW2	mgm4795339.3	662,422	471,513	231 ± 39 bp
	MWAW1	mgm4795335.3	233,177	157,539	226 ± 38 bp
	MW 15R	mgm4795679.3	6,710,609	5,018,326	241 ± 35 bp
	PMW4	mgm4795678.3	5,429,417	4,433,348	241 ± 35 bp
Edison, NJ	MW114	mgm4795676.3	6,174,464	5,242,089	240 ± 35 bp
	MW303S	mgm4795677.3	5,824,346	5,008,287	239 ± 35 bp
Indian Head, MD	IW5	mgm4795927.3	5,674,151	4,837,429	241 ± 34 bp
	IW7	mgm4795847.3	2,036,212	1,547,247	219 ± 41 bp
	MW38	mgm4795845.3	5,832,647	5,071,187	240 ± 35 bp
	MW40	mgm4795846.3	5,265,951	4,480,623	241 ± 35 bp

Supplementary Table 2. 3. Genomes used for collecting functional protein sequences.

Organism/Name	Strain	Size (Mb)	GC%	Replicons	WGS	Gene	Protein	# of RDase	Release Date
Dehalococcoides mccartyi	195	1.46972	48.9	chromosome:NC_002936.3/CP000027.1	-	1582	1497	19	10/3/2001
Dehalococcoides mccartyi	CG5	1.36215	47.2	chromosome:NZ_CP006951.1/CP006951.1	-	1459	1395	25	8/4/2014
Dehalococcoides mccartyi	CBDB1	1.3955	47	chromosome:NC_007356.1/AJ965256.1	-	1479	1412	32	8/19/2005
Dehalococcoides mccartyi	BAV1	1.34189	47.2	chromosome:NC_009455.1/CP000688.1	-	1444	1374	10	5/7/2007
Dehalococcoides mccartyi	VS	1.41346	47.3	chromosome:NC_013552.1/CP001827.1	-	1505	1432	37	12/3/2009
Dehalococcoides mccartyi	GT	1.36015	47.3	chromosome:NC_013890.1/CP001924.1	-	1468	1399	20	2/17/2010
Dehalococcoides mccartyi	DCMB5	1.4319	47.1	chromosome:NC_020386.1/CP004079.1	-	1524	1461	23	2/22/2013
Dehalococcoides mccartyi	BTF08	1.45233	47.3	chromosome:NC_020387.1/CP004080.1	-	1556	1485	20	2/22/2013
Dehalococcoides mccartyi	GY50	1.40742	47	chromosome:NC_022964.1/CP006730.1	-	1499	1427	26	11/26/2013
Dehalococcoides mccartyi	CG4	1.38231	48.7	chromosome:NZ_CP006950.1/CP006950.1	-	1470	1401	13	8/4/2014
Dehalococcoides mccartyi	CG1	1.48668	46.9	chromosome:NZ_CP006949.1/CP006949.1	-	1600	1527	32	8/4/2014
Dehalococcoides mccartyi	IBARAKI	1.45106	47	chromosome Unknown:NZ_AP014563.1/AP014563.1	-	1556	1471	28	9/9/2015
Dehalococcoides mccartyi	11a5	1.46791	46.87	chromosome:NZ_CP011127.1/CP011127.1 plasmid pDhc6:NZ_CP011128.1/CP011128.1	-	1587	1521	30	4/5/2016
Dehalococcoides mccartyi	CG3	1.52129	46.9	chromosome:NZ_CP013074.1/CP013074.1	-	1657	1589	20	12/6/2016
Dehalococcoides mccartyi	KBTCE2	1.3292	49.1	chromosome:NZ_CP019865.1/CP019865.1	-	1431	1364	4	2/28/2017
Dehalococcoides mccartyi	KBDCA1	1.42846	47.4	chromosome:NZ_CP019867.1/CP019867.1	-	1563	1483	6	2/28/2017
Dehalococcoides mccartyi	KBDCA2	1.39432	47.5	chromosome:NZ_CP019868.1/CP019868.1	-	1523	1443	6	2/28/2017
Dehalococcoides mccartyi	KBTCE3	1.2716	49.3	chromosome:NZ_CP019866.1/CP019866.1	-	1361	1295	4	2/28/2017
Dehalococcoides mccartyi	KBDCA3	1.33749	47.6	chromosome:NZ_CP019946.1/CP019946.1	-	1441	1372	7	3/6/2017
Dehalococcoides mccartyi	KBVC2	1.33773	47.2	chromosome:NZ_CP019969.1/CP019969.1	-	1440	1378	16	3/9/2017
Dehalococcoides mccartyi	KBVC1	1.3599	47.3	chromosome:NZ_CP019968.1/CP019968.1	-	1456	1393	21	3/9/2017
Dehalococcoides mccartyi	KBTCE1	1.38891	47.3	chromosome:NZ_CP019999.1/CP019999.1	-	1502	1441	16	3/15/2017
Dehalococcoides mccartyi	UCH-ATV1	1.38778	48.8	chromosome:NZ_AP017649.1/AP017649.1	-	1489	1408	15	7/8/2017
Dehalococcoides mccartyi	MB	1.57151	48.3	-	JGYD01	1711	1614	27	12/4/2015
Dehalococcoides mccartyi	11a	1.32452	47.2	-	JGVX01	1415	1339	8	12/4/2015
Dehalococcoides mccartyi	JNA	1.46251	47.1	-	JSWM01	1582	1515	26	1/15/2016
Dehalococcoides mccartyi	SG1	1.42874	47.1	-	JPRE01	1535	1467	28	8/18/2014
Dehalococcoides mccartyi	WBC-2	1.37458	47.4	chromosome:CP017572.1	-	1466	1386	15	10/12/2016
Dehalococcoides mccartyi	EV-VC	1.4716	46.8	-	LZFK01	1535	1475	28	7/1/2016
Dehalococcoides mccartyi	EV-TCE	1.3573	48.5	-	LZFJ01	1451	1369	11	7/1/2016
Dehalogenimonas lykanthroporepellens	BL-DC-9	1.68651	55.5	chromosome:NC_014314.1/CP002084.1	-	1732	1650	20	1/28/2014
Dehalogenimonas formicexedens	NSZ-14	2.092789	54	chromosome:NZ_CP018258.1/CP018258.1	-	2176	2091	25	10/17/2017
Dehalogenimonas alkenigignens	IP3-3	1.85	55.9	-	LFDV01	1940	1856	29	6/23/2016
Dehalogenimonas	WBC-2	1.72573	49.2	chromosome:CP011392.1	-	1800	1721	22	5/8/2015
Geobacter lovleyi	SZ	0.077113	52.97	chromosome:NC_010815.1/CP001090 plasmid pGLOV01: NC_010815.1/CP001090.1	-	3640	3552	2	6/26/2016
Sulfurospirillum	SL2-1	2.87654	38.7	chromosome:NZ_CP021416.1/CP021416.1	-	2943	2701	2	5/31/2017
Sulfurospirillum	JPD-1	2.81409	38.8	chromosome:NZ_CP023275.1/CP023275.1	-	2882	2793	2	9/18/2017

Supplementary Table 2. 3. (continued)

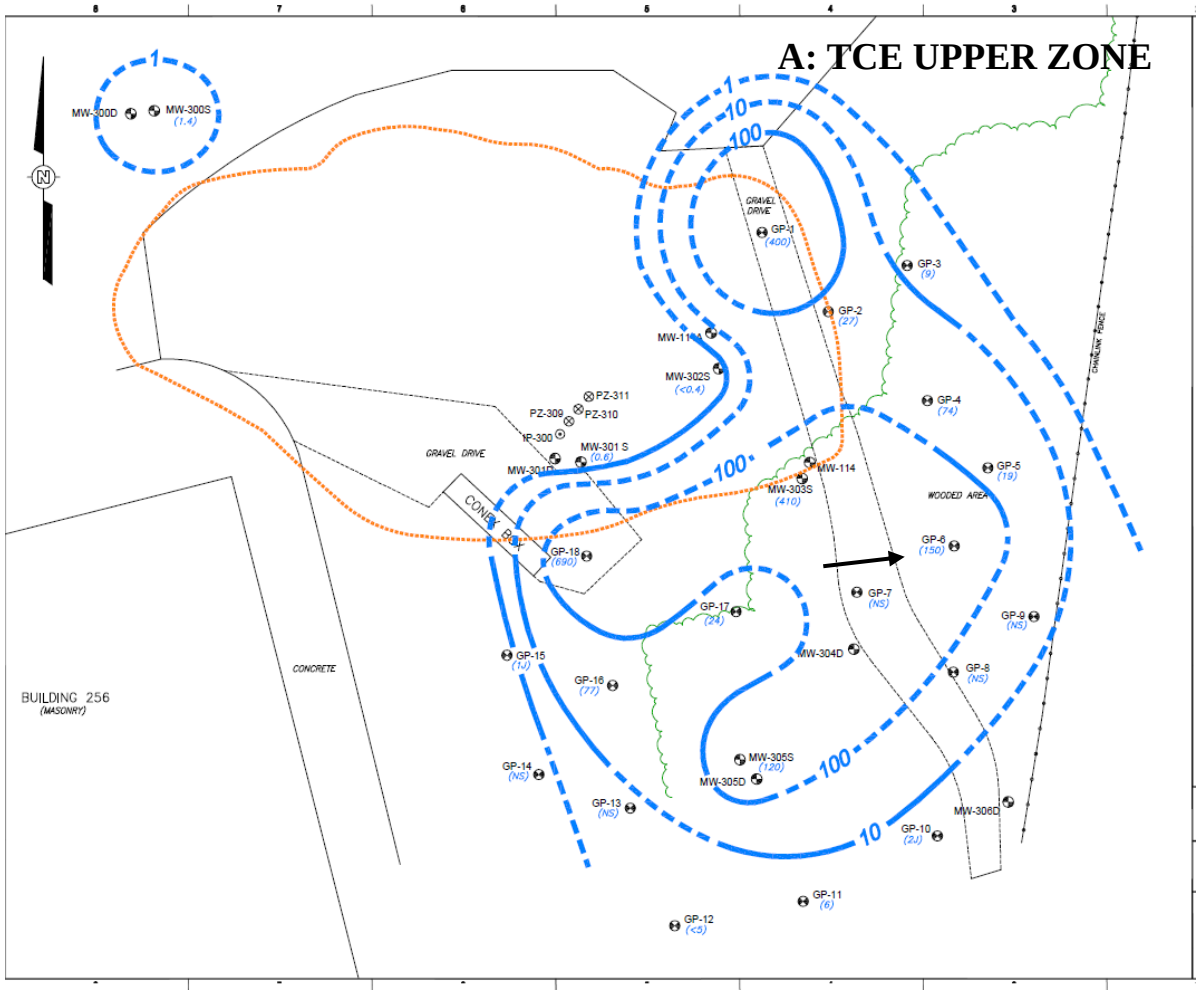
Sulfurospirillum	SL2-2	2.87661	38.7	chromosome:NZ_CP021979.1/CP021979.1	-	2924	2699	2	6/22/2017
Sulfurospirillum halorespirans	DSM 13726	3.03	41.3	chromosome:NZ_CP017111.1/CP017111.1	-	3035	2967	2	9/8/2017
Sulfurospirillum multivorans	DSM 12446	3.18	40.9	chromosome:NZ_CP007201.1/CP007201.1	-	3288	3186	2	4/29/2015
Anaeromyxobacter dehalogenans	2CP-C	5.01348	74.9	chromosome:NC_007760.1/CP000251.1	-	4522	4416	2	1/27/2006
Dehalobacter restrictus	DSM 9455	2.94	44.6	chromosome:NZ_CP007033.1/CP007033.1	-	2848	2647	23	5/14/2014
Dehalobacter	DCA	3.06995	44.6	chromosome:NC_018866.1/CP003869.1	-	2974	2848	18	10/16/2012
Dehalobacter	CF	3.09205	44.3	chromosome:NC_018867.1/CP003870.1	-	2985	2882	18	10/16/2012
Dehalobacter	E1	2.95026	43.8	-	CANE01	2866	2719	7	9/19/2012
Dehalobacter	FTH1	6.32936	58.9	-	AQYY01	5934	5727	33	4/19/2013
Dehalobacter	UNSWDHB	3.20156	44.9	-	AUUR01	3105	2944	19	8/9/2013
Dehalobacter	TeCB1	3.13322	44	-	MCHF01	3106	2961	24	8/18/2016
Desulfitobacterium hafniense	DCB-2	5.27913	47.5	chromosome:NC_011830.1/CP001336.1	-	5038	4821	7	1/5/2009
Desulfitobacterium hafniense	Y51	5.72753	47.4	chromosome:NC_007907.1/AP008230.1	-	5484	5227	2	3/10/2006
Desulfitobacterium hafniense	TCP-A	4.96723	47.3	-	AQZD01	4839	4556	5	4/22/2013
Desulfitobacterium hafniense	PCP-1	5.56321	47.5	-	ARAZ01	5327	5095	7	4/22/2013
Desulfitobacterium hafniense	PCE-S	5.6667	47.3	-	-	5490	5417	6	-
Desulfitobacterium	PCE1	4.22	45	-	AQZF01	4070	3873	6	9/16/2013
Desulfitobacterium chlororespirans	DSM 11544	5.61	47.3	-	FRDN01	5367	5282	2	12/2/3016
Desulfitobacterium dichloroeliminans	LMG P-21439	3.62	44.2	chromosome:NC_019903.1/CP003344.1	-	3463	3300	1	6/17/2013
Desulfitobacterium dehalogenans	ATCC 51507	4.32	45	chromosome:NC_018017.1/CP003348.1	-	4212	3974	7	9/10/2015
Polaromonas JS666	JS666	5.89868	62.00	chromosome:NC_007948.1/CP000316.1 plasmid 1:NC_007949.1/CP000317.1 plasmid 2:NC_007950.1/CP000318.1	-	5660	5485	1[a]	2006/04/10

a: Cytochrome P450 (ABE47160.1) from *Polaromonas* JS666 is not a RDase but catalyzes the initial step of cDCE degradation.

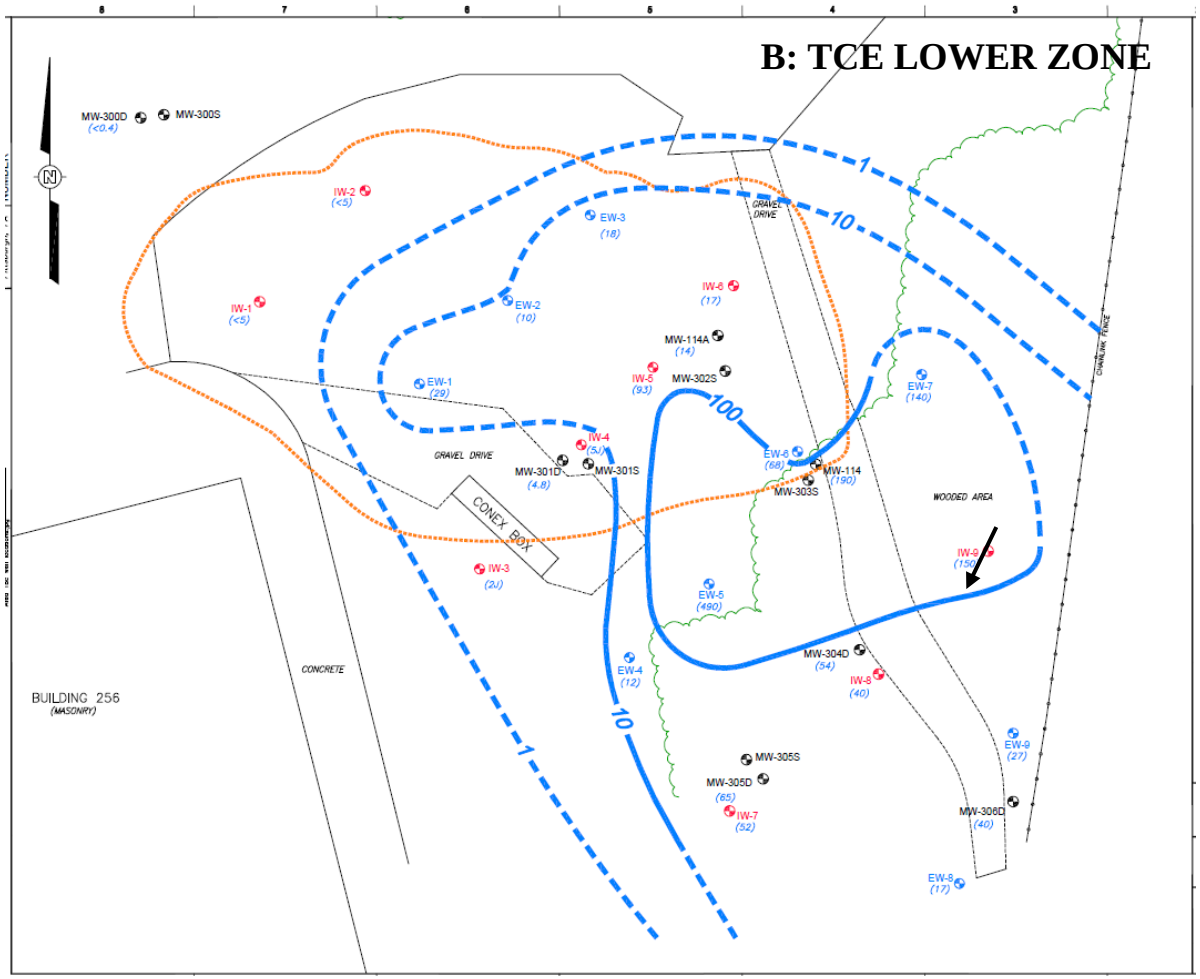
Supplementary Table 2. 4. Number of collected genomes and dereplicated RDases.

Microorganism	Number of collected genome	Dereplicated RDase number
Dehalococcoides mccartyi	30	317
Dehalogenimonas	4	91
Anaeromyxobacter	1	2
Dehalobacter	7	103
Geobacter	1	2
Sulfurospirillum	5	6
Desulfitobacterium	9	36
Polaromonas	1	1 (not an RDase)

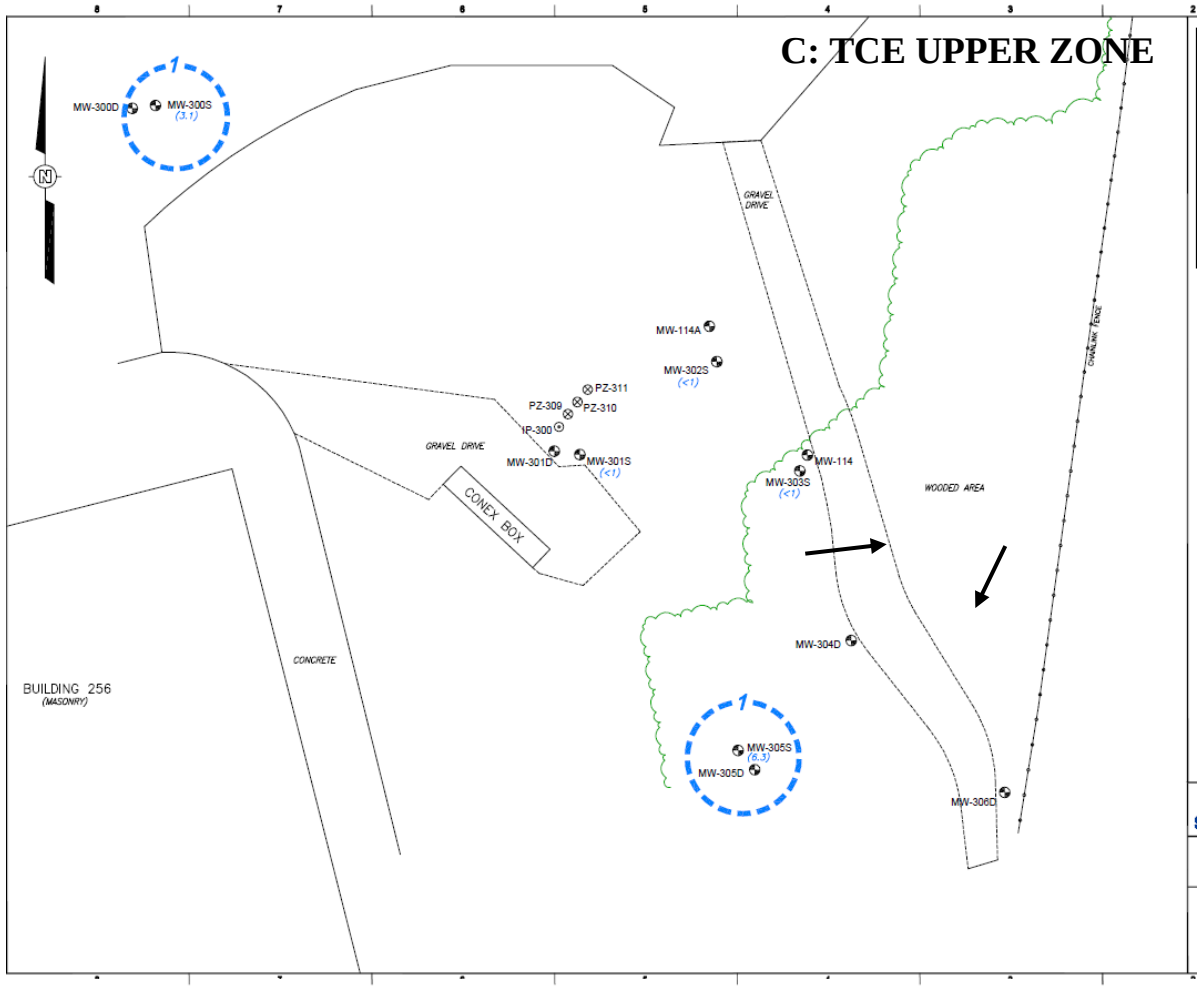
Supplementary Figure 2. 1. TCE plume maps for the Edison, NJ site. TCE contour maps for the site prior to addition of emulsified oil and dehalogenating culture SDC-9 in 2009 are provided for the shallow zone (A) and deep zone at the site (B). Well 303S is located in the shallow zone and well 114 is located in the deep zone. Post-treatment contour maps in 2010 for the shallow zone (C) and deep zone (D) are also provided. All values are in µg/L. The wells from which samples were collected and analyzed are indicated with arrows.



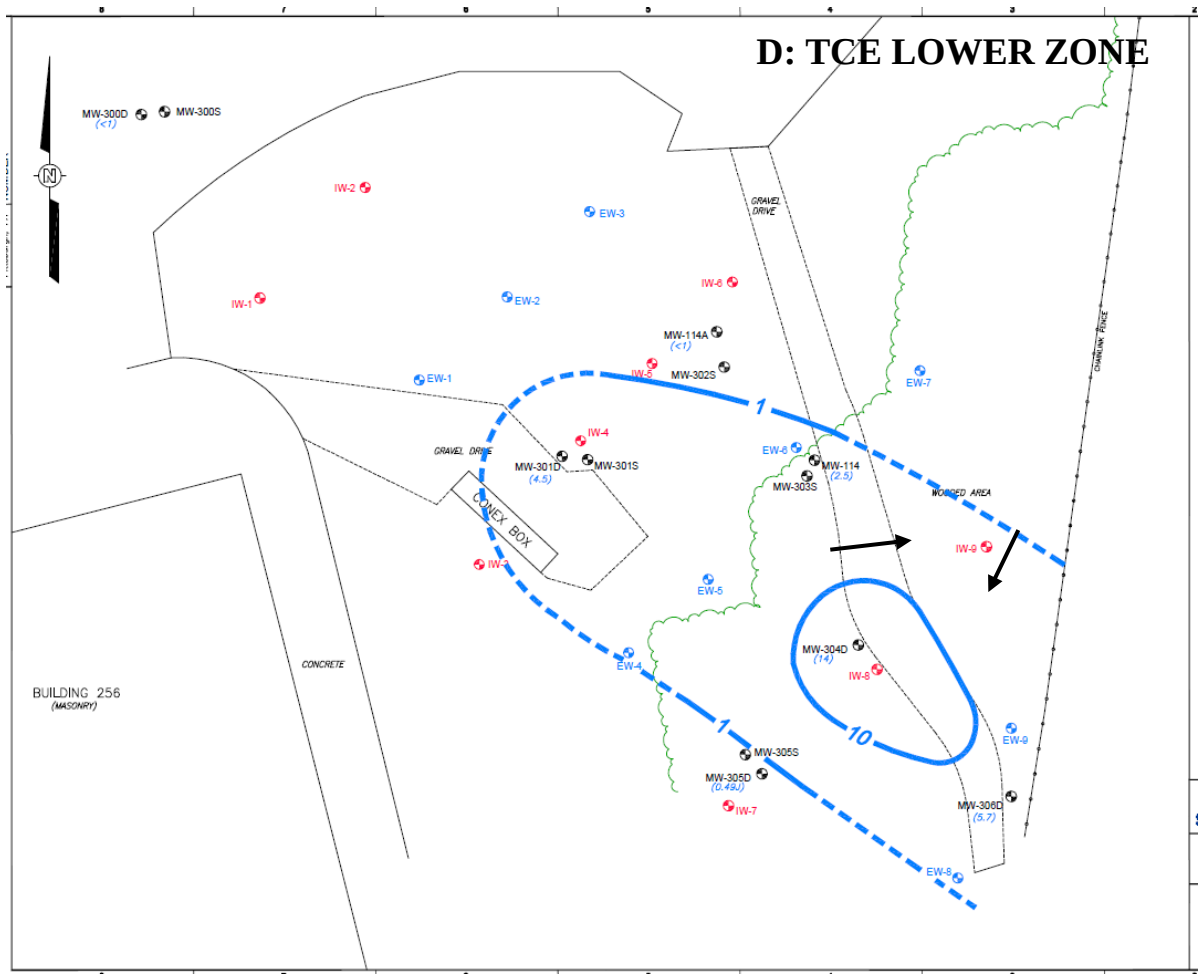
Supplementary Figure 2. 1. (continued)



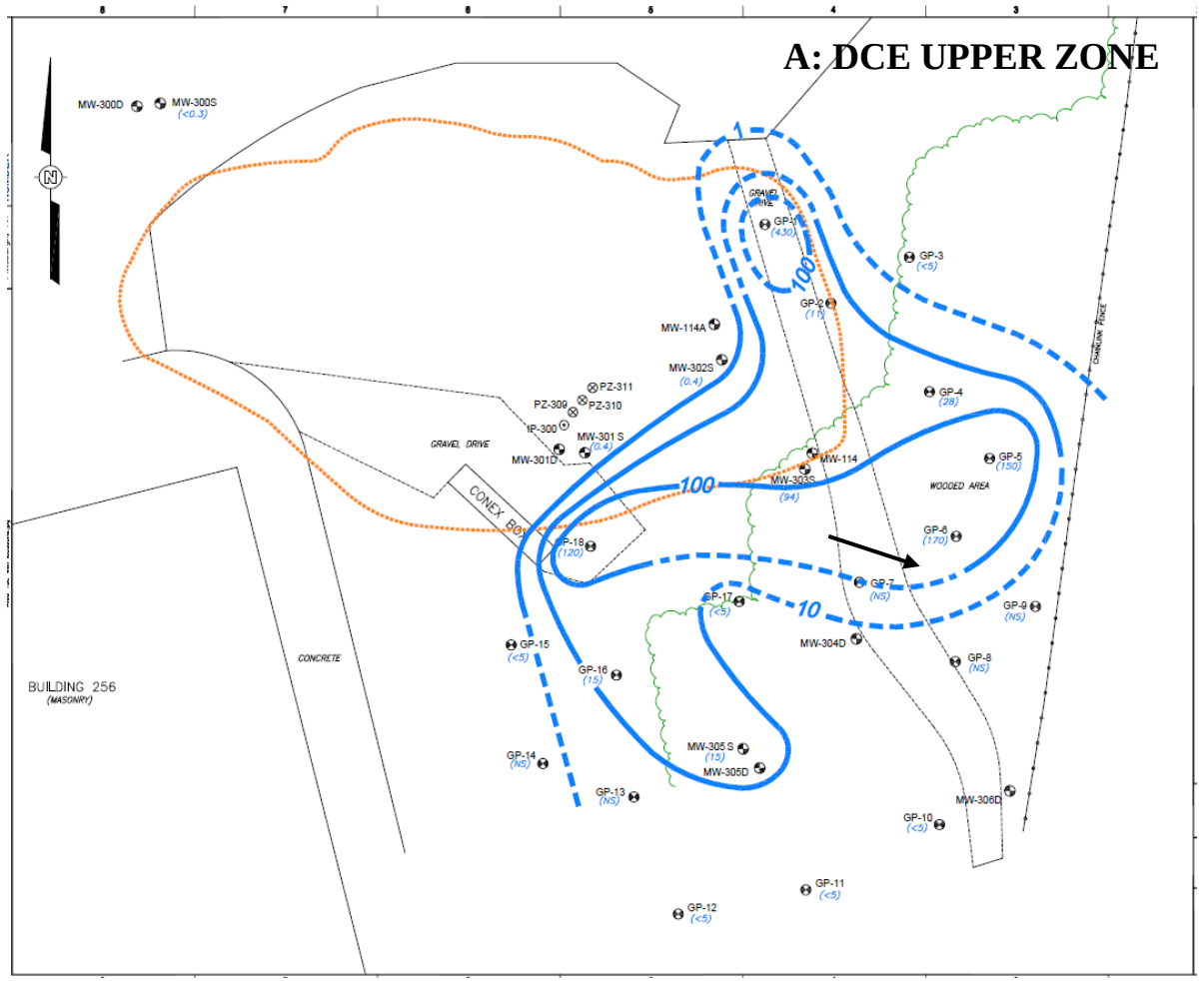
Supplementary Figure 2. 1. (continued)



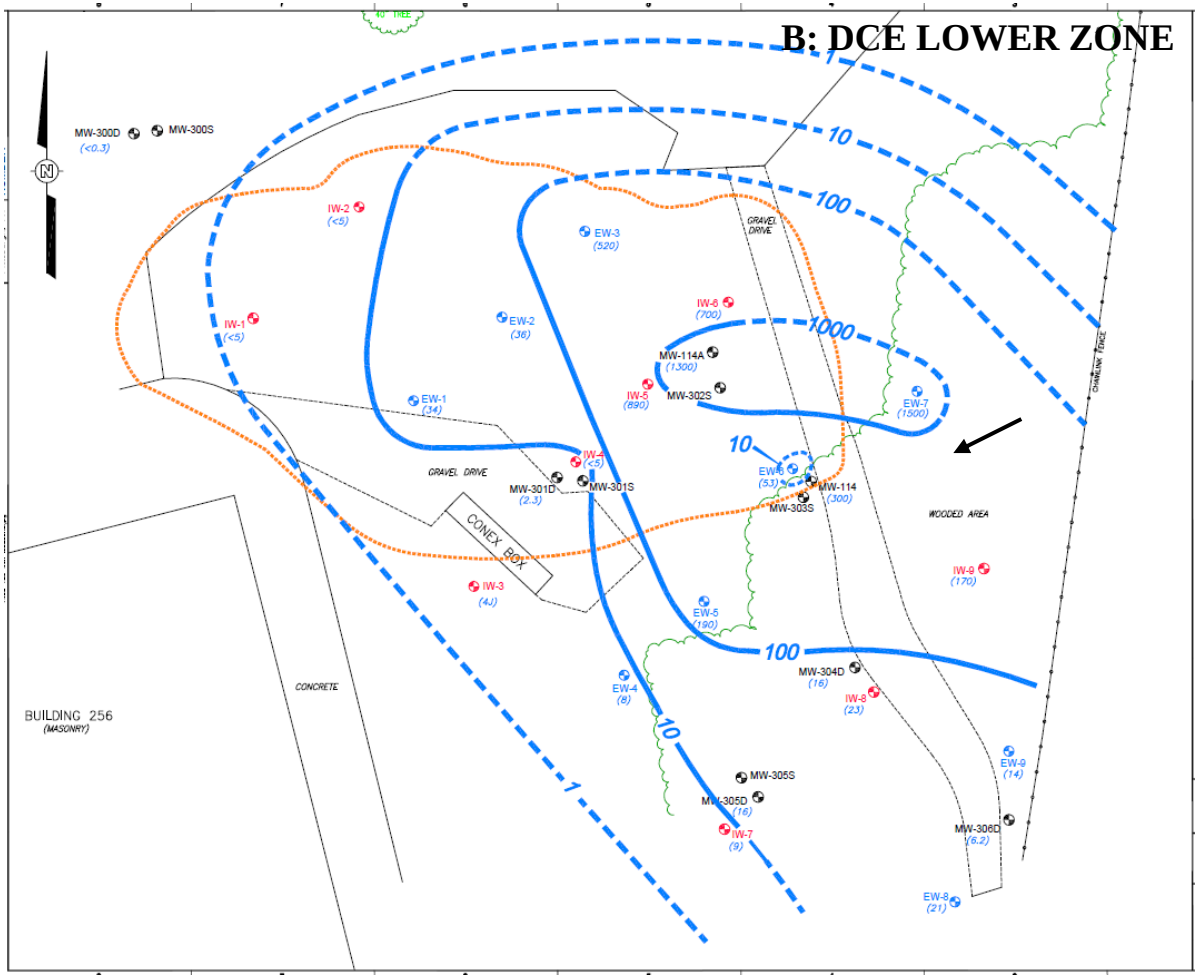
Supplementary Figure 2. 1. (continued)



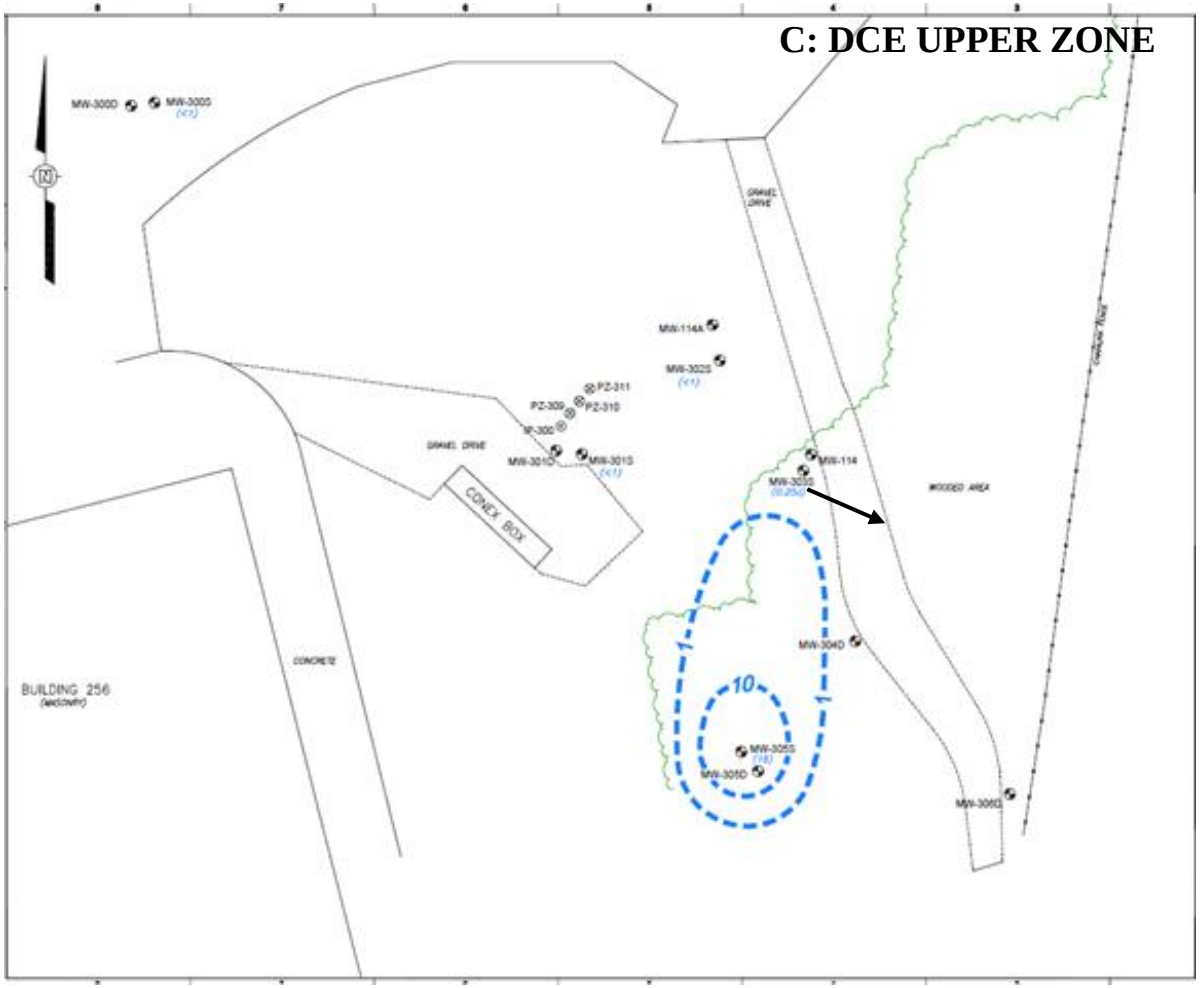
Supplementary Figure 2. 2. Cis-DCE Plume maps for the Edison, NJ site. Cis-DCE contour maps for the site prior to addition of emulsified oil and dehalogenating culture SDC-9 in 2009 are provided for the shallow zone (A) and deep zone at the site (B). Well 303S is located in the shallow zone and well 114 is located in the deep zone. Post-treatment contour maps in 2010 for the shallow zone (C) and deep zone (D) are also provided. All values are in µg/L. The wells from which samples were collected and analyzed are indicated with arrows.



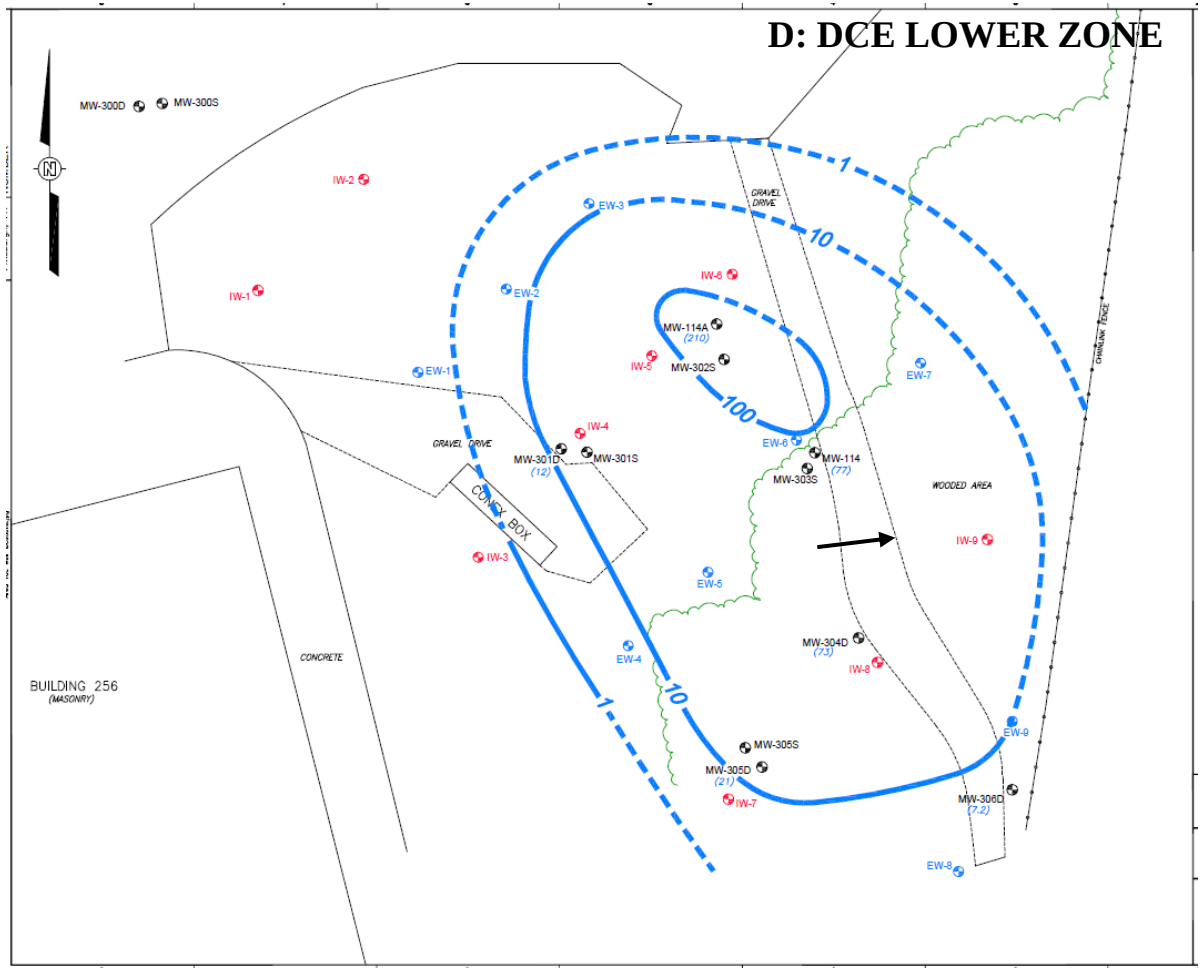
Supplementary Figure 2. 2. (continued)



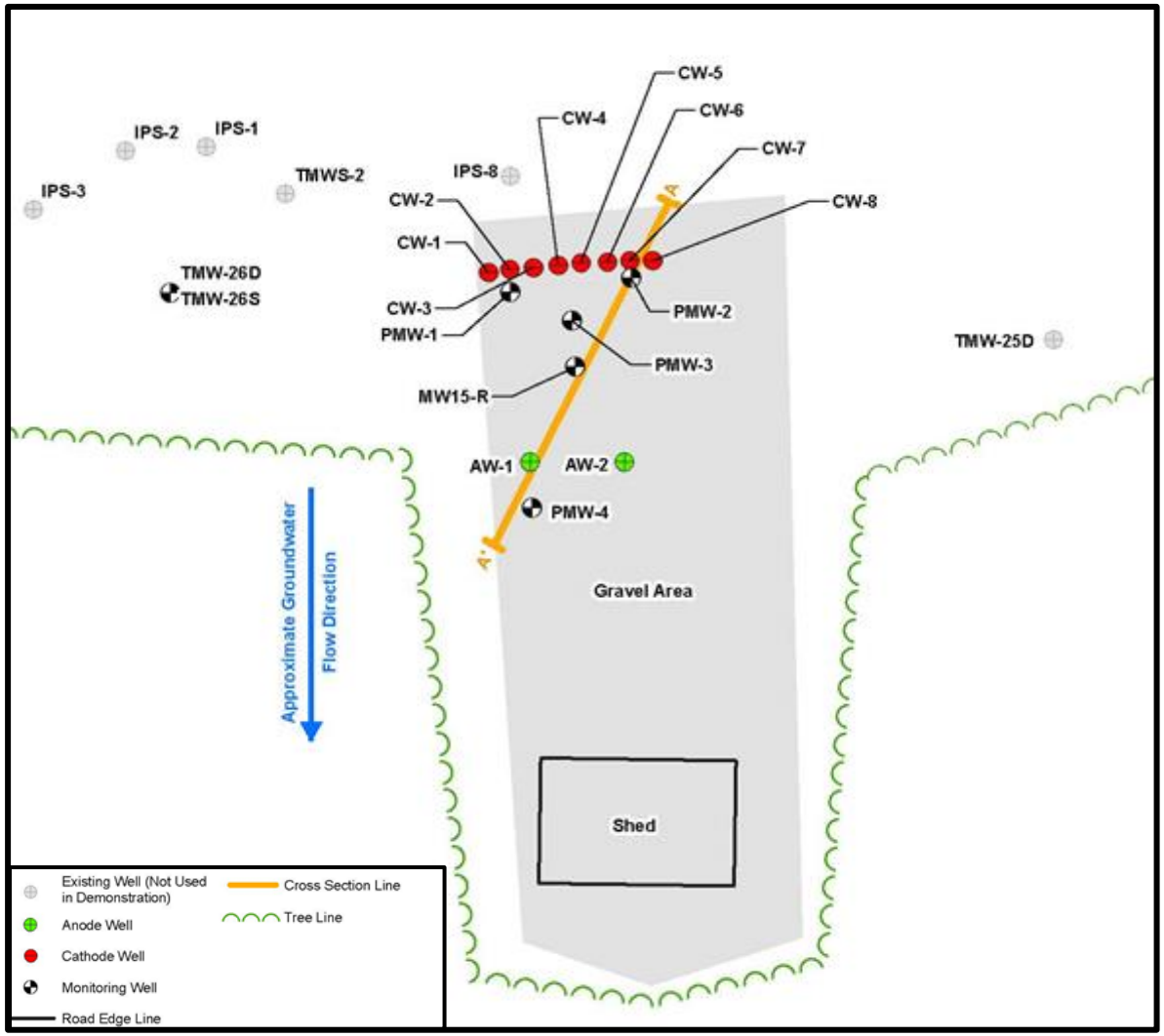
Supplementary Figure 2. 2. (continued)



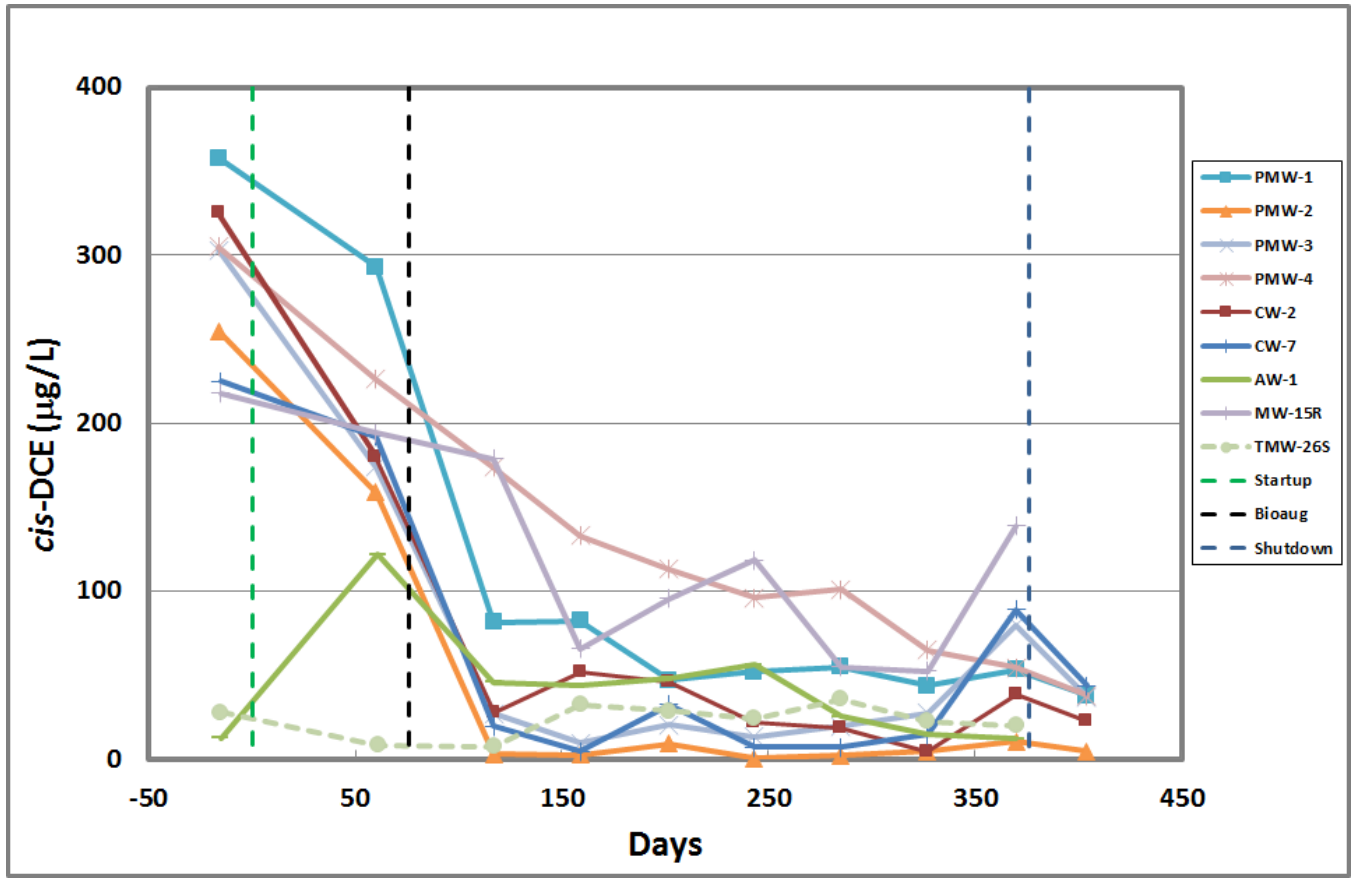
Supplementary Figure 2. 2. (continued)



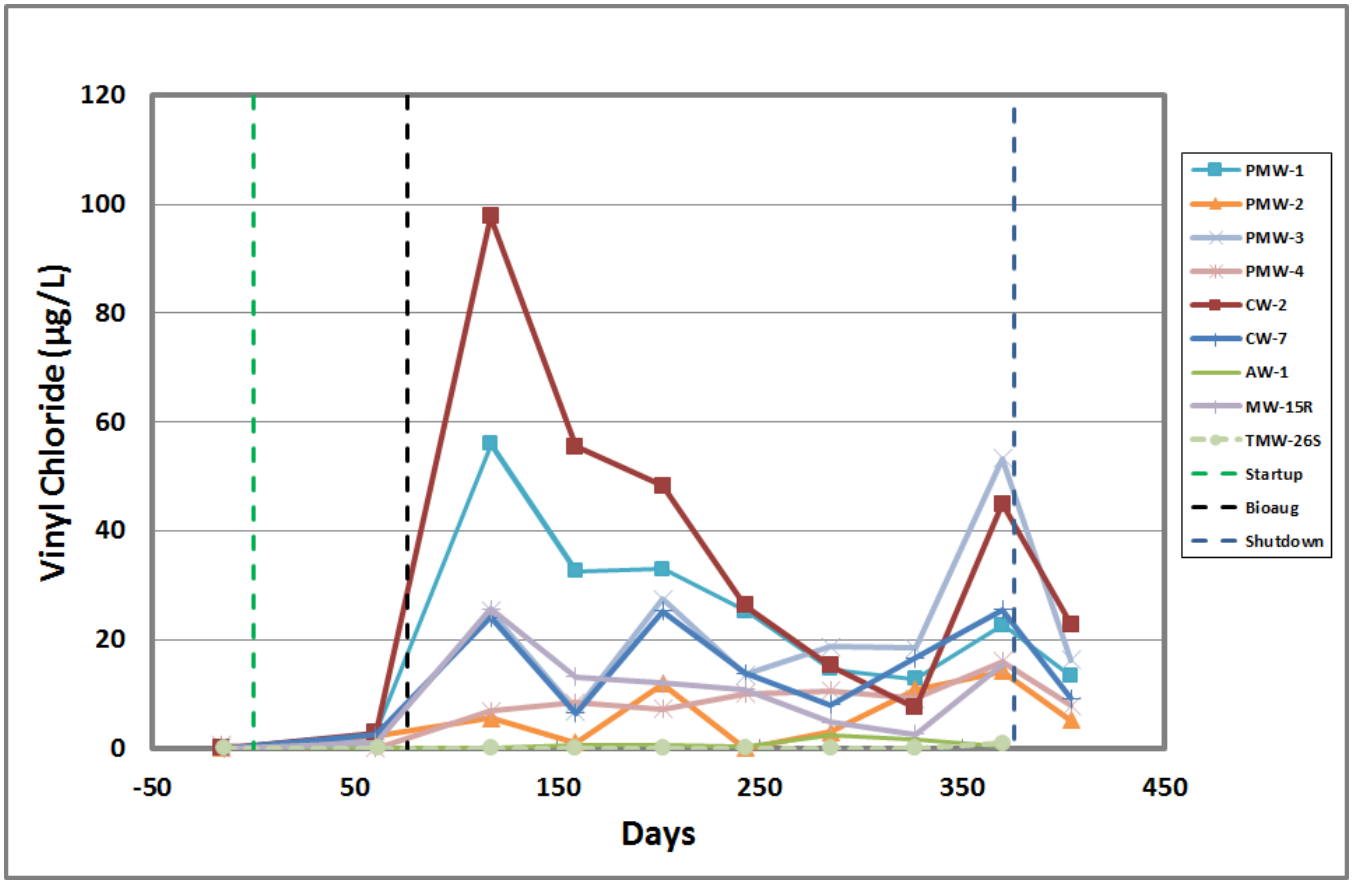
Supplementary Figure 2. 3. Demonstration plot layout at the Quantico, VA site. The cathode and anode wells are indicated by red and green symbols, respectively. This system was used to supply H_2 to support reductive dechlorination of cis-DCE downgradient of a landfill. See data in Supplementary Figures 17-19.



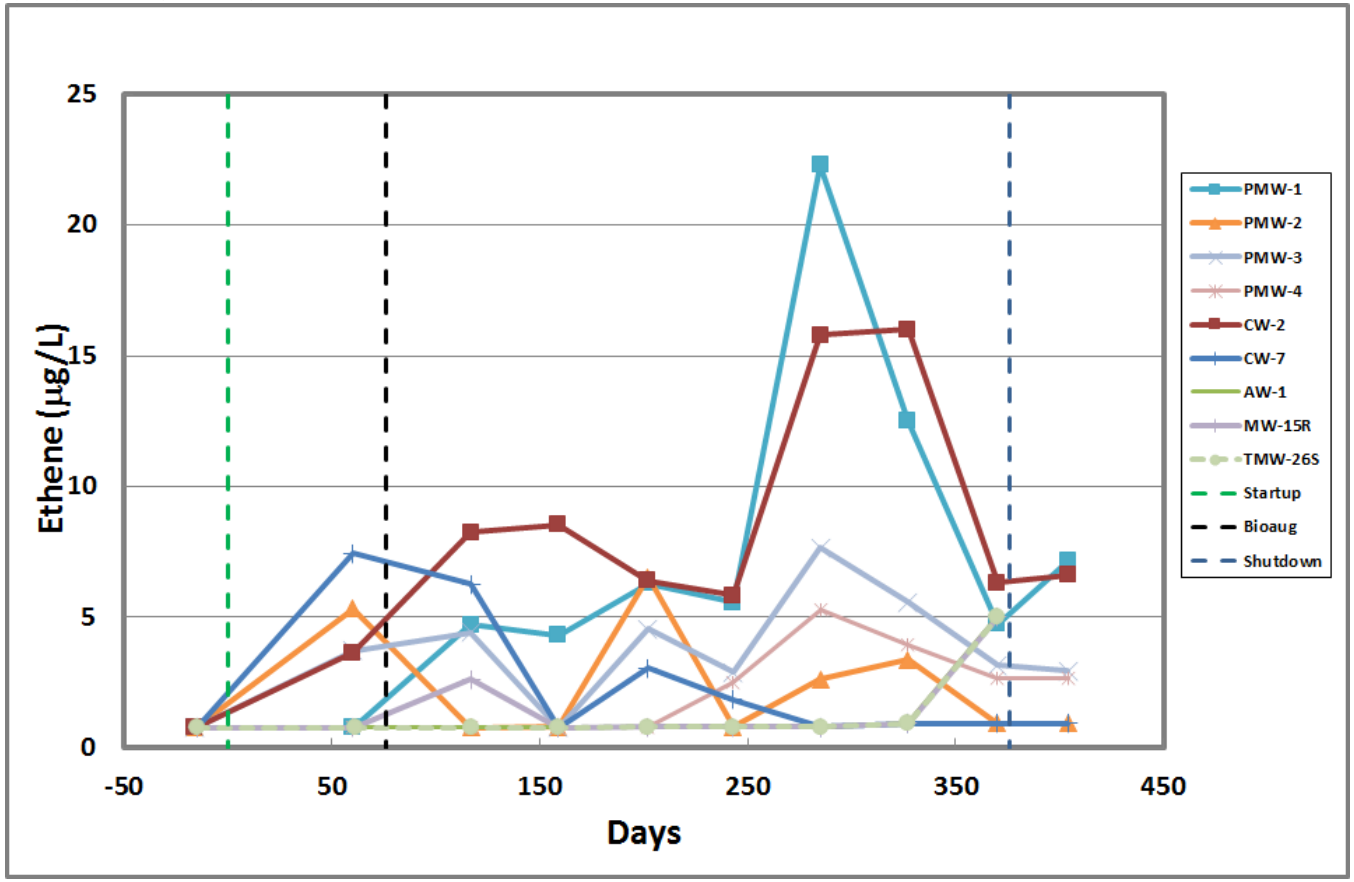
Supplementary Figure 2. 4. Concentration data for cis-DCE at the Quantico, VA site. The groundwater samples were collected on Day 243 from wells CW-2, PMW-2, CW-2, AW-1, MW-15R, and PMW-4.



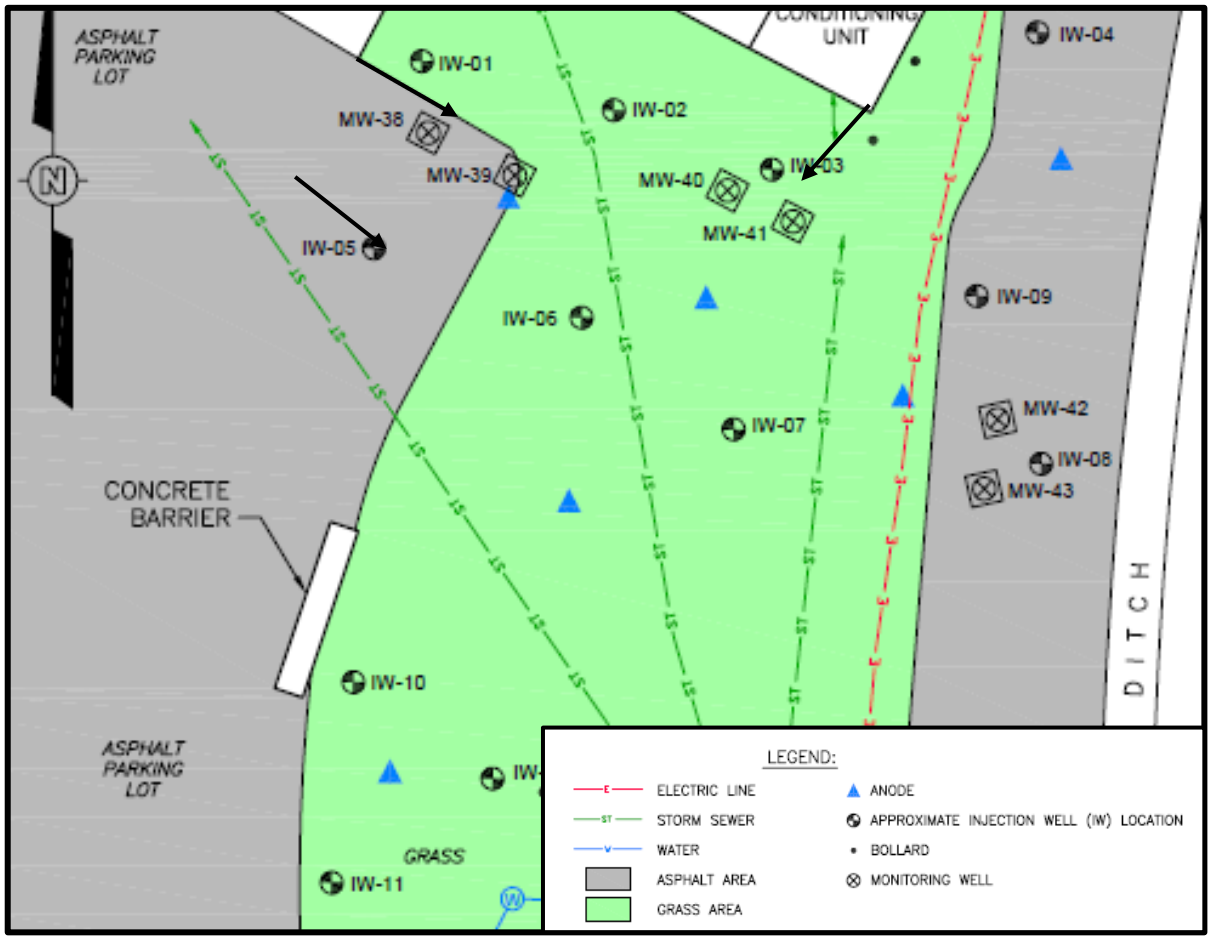
Supplementary Figure 2. 5. Concentration data for vinyl chloride at the Quantico, VA site. The groundwater samples were collected on Day 243 from wells CW-2, PMW-2, CW-2, AW-1, MW-15R, and PMW-4.



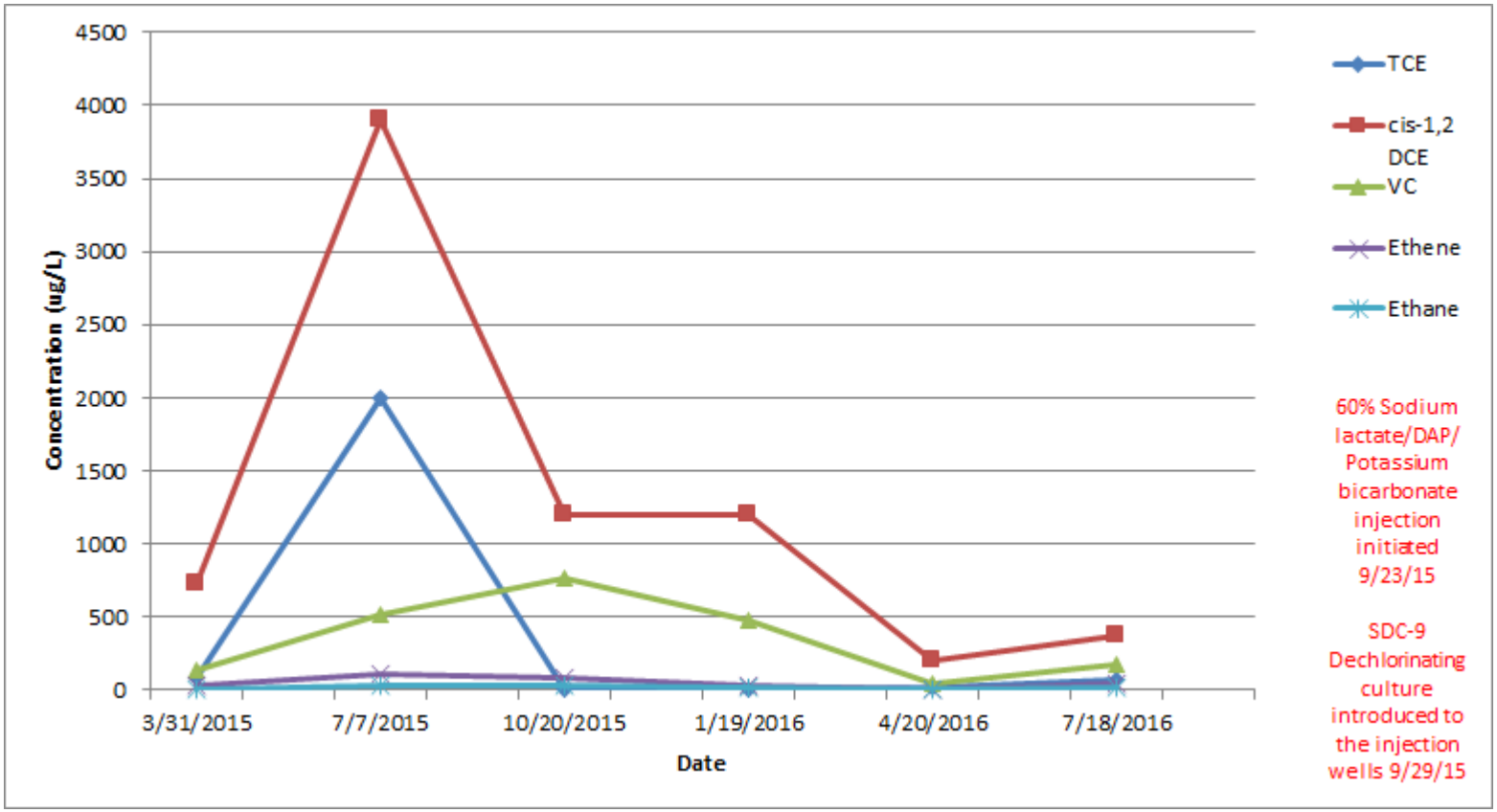
Supplementary Figure 2. 6. Concentration data for ethene at the Quantico, VA site. The groundwater samples were collected on Day 243 from wells CW-2, PMW-2, CW-2, AW-1, MW-15R, and PMW-4.



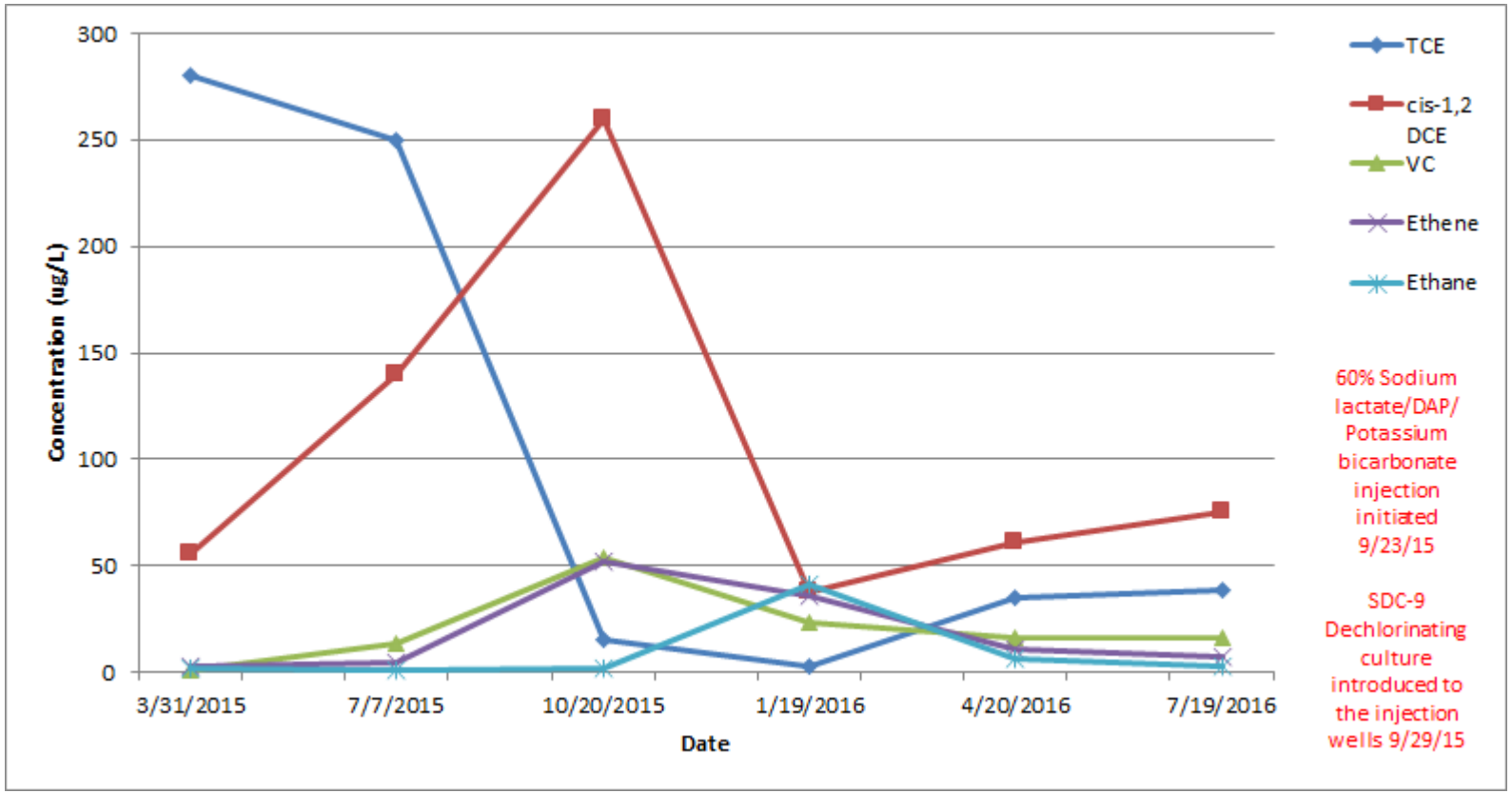
Supplementary Figure 2. 7. Demonstration plot layout at the Indian Head, Md site. Injection wells (IWs) were amended with lactate, diammonium phosphate, potassium bicarbonate (for pH adjustment) and dehalogenating culture SDC-9. Monitoring wells (MWs) were used to measure system performance. A low voltage was used to maintain system pH. Anodes for this system are shown in the figure. Wells that were sampled are indicated by arrows. See MW data in Supplementary Figures 21-22. No analytical data are available for the IWs.



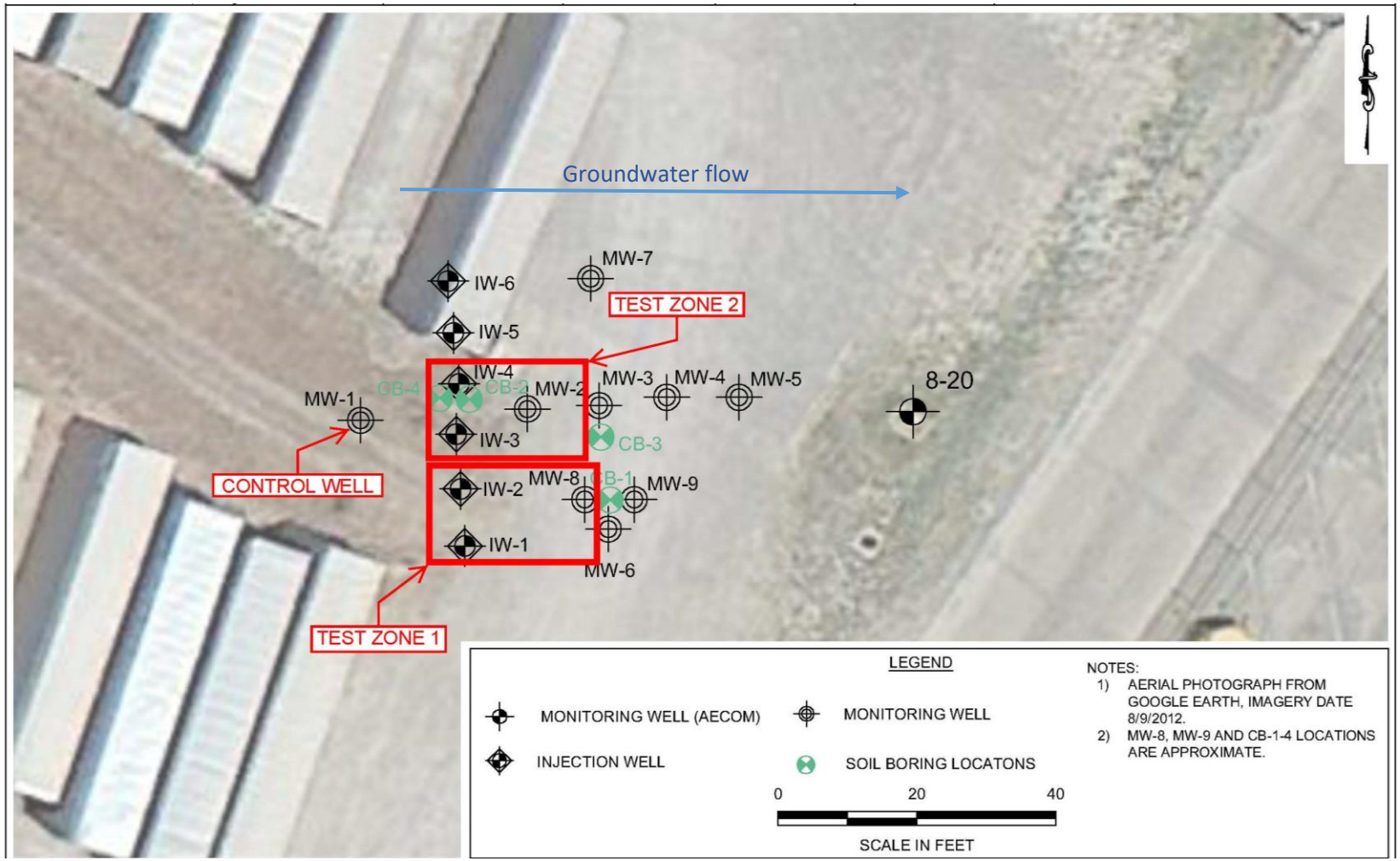
Supplementary Figure 2. 8. Concentration data for cVOCs, ethene and ethane in well MW38 at the Indian Head, Md site. The groundwater samples were collected on 6/22/16.



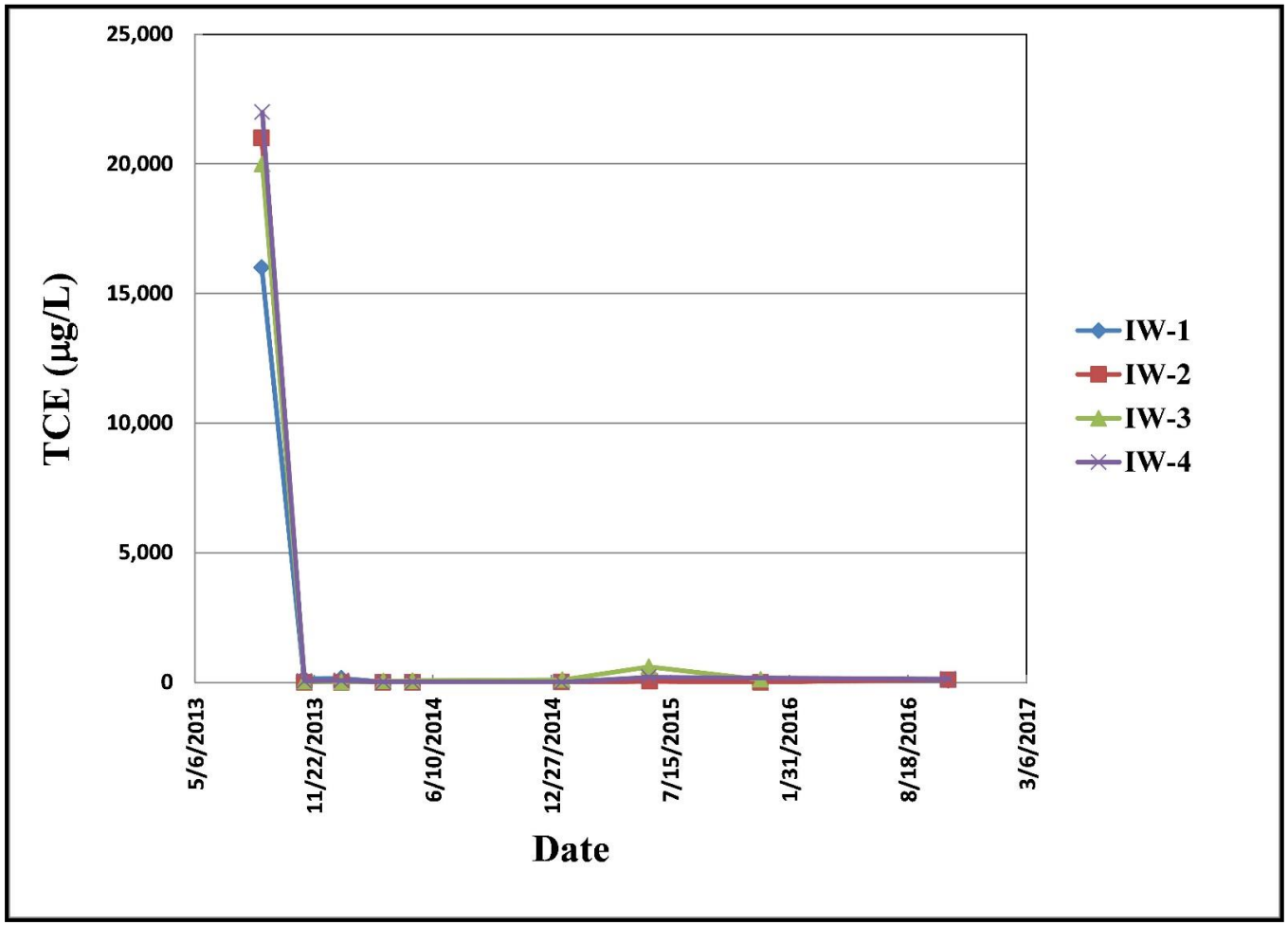
Supplementary Figure 2. 9. Concentration data for cVOCs, ethene and ethane in well MW40 at the Indian Head, Md site. The groundwater samples were collected on 6/22/16.



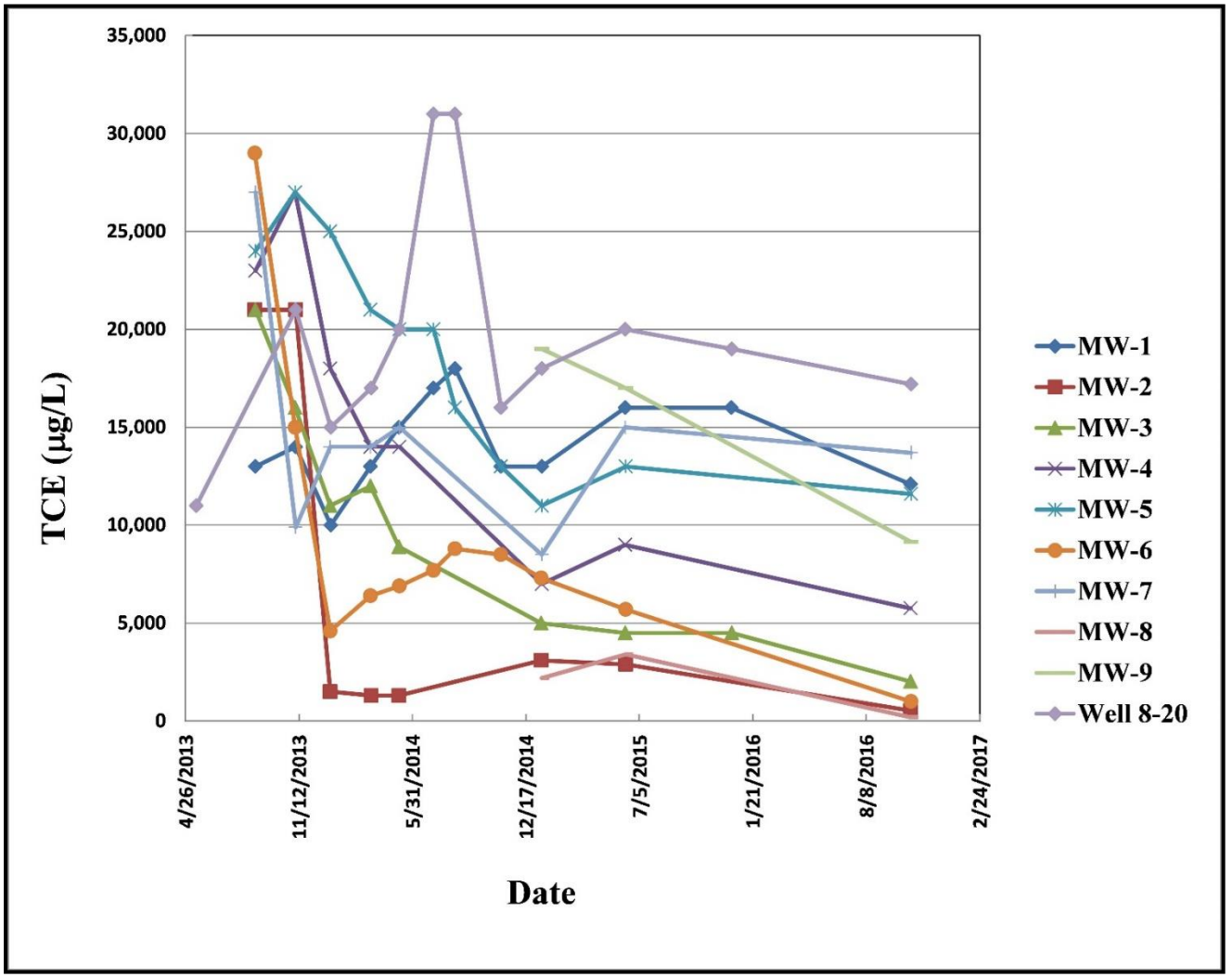
Supplementary Figure 2. 10. Demonstration Plot layout at the Tulsa, Ok site. IWs are emulsified oil and dehalogenating culture SDC-9 injection wells and MWs are groundwater monitoring wells. See data in Supplementary Figures 24-26.



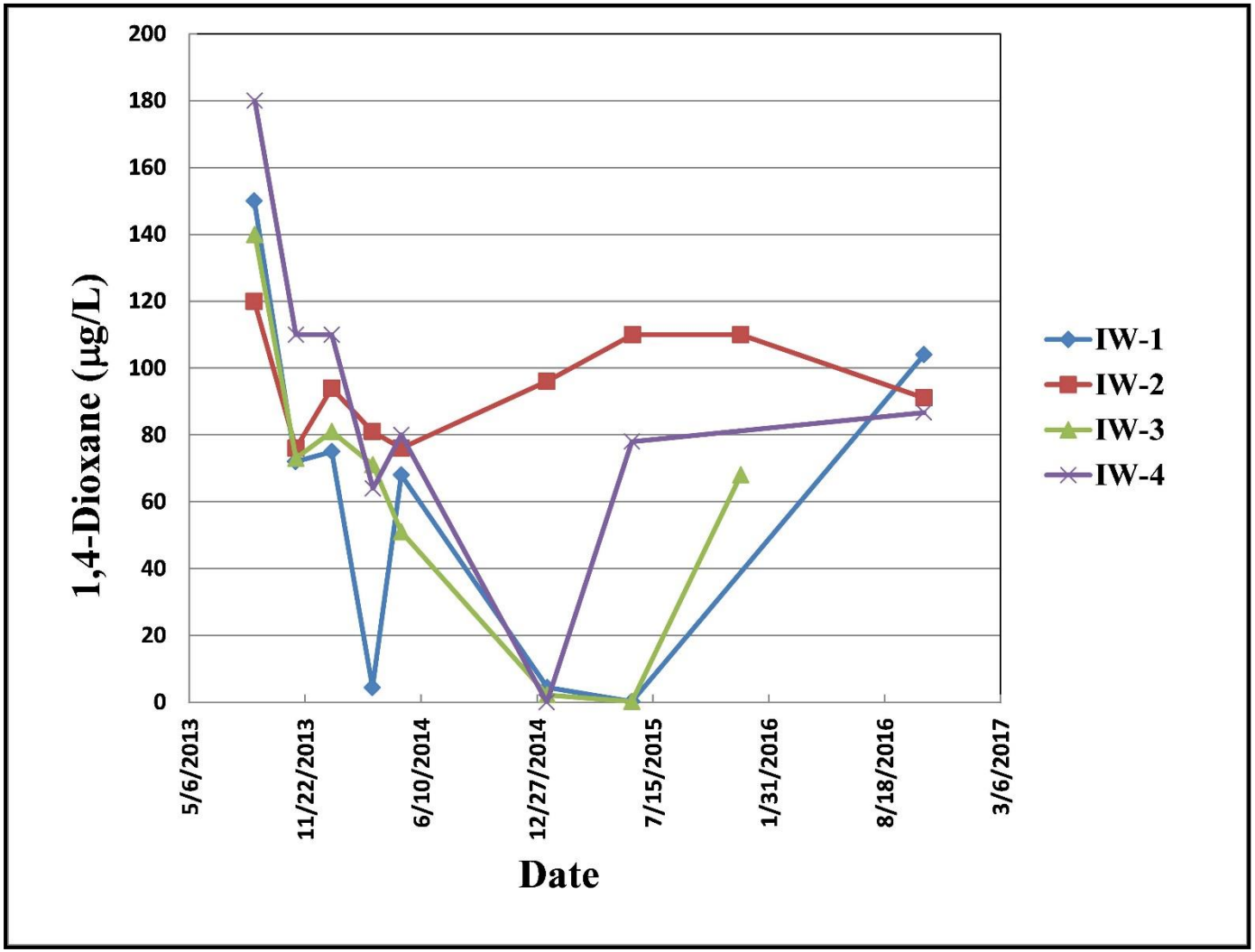
Supplementary Figure 2. 11. Concentration data for TCE in injection wells (IWs) at the Tulsa, OK Site. The groundwater samples were collected on 6/09/15.



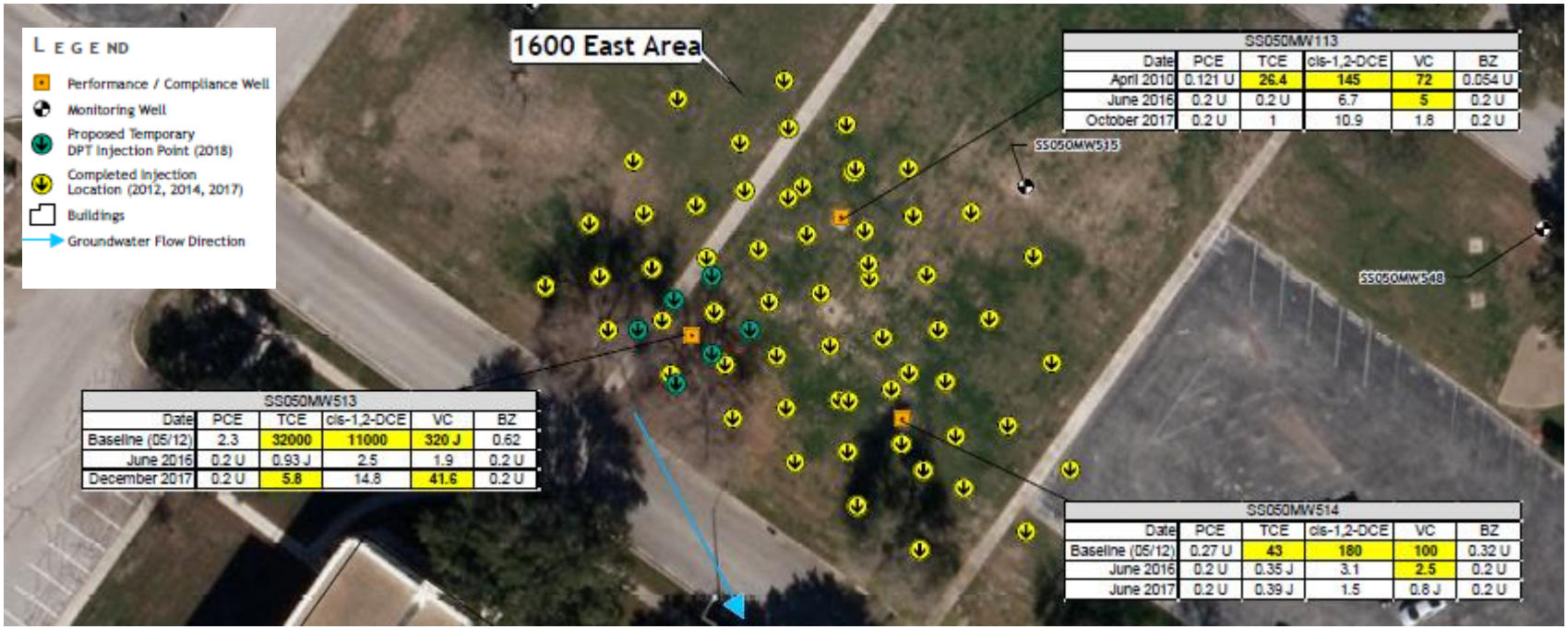
Supplementary Figure 2. 12. Concentration data for TCE in monitoring wells (MWs) at the Tulsa, OK Site. The groundwater samples were collected on 6/09/15.



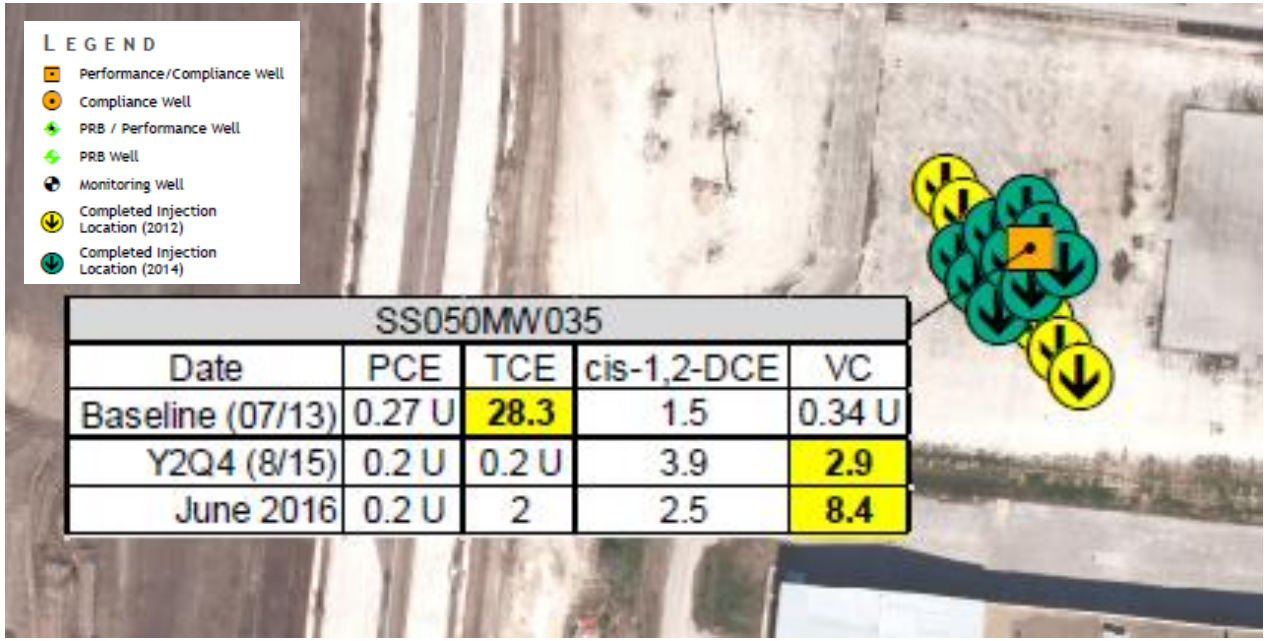
Supplementary Figure 2. 13. Concentration data for 1,4-dioxane in injection wells (IWs) at the Tulsa, OK Site. The groundwater samples were collected on 6/09/15.



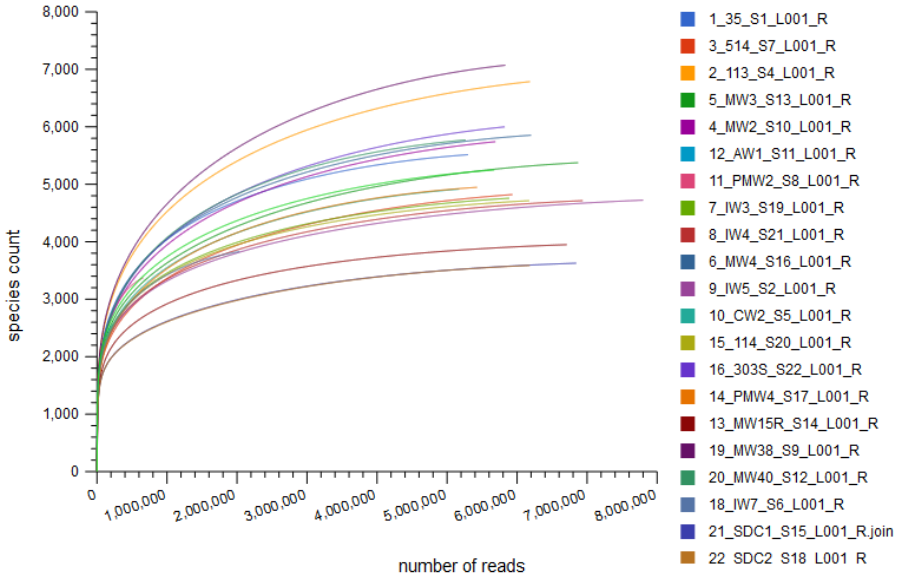
Supplementary Figure 2. 14. Injection points and locations of monitoring wells SS050MW113 (113) and SS050MW514 (514) at the San Antonio, TX, Site. Analytical data are provided for each well. Groundwater samples were collected on 7/28/16. BZ = benzene.



Supplementary Figure 2. 15. Injection points and location of monitoring well SS050MW035 (35) at the San Antonio, TX, Site. Analytical data are provided. Groundwater samples were collected on 7/28/16.



SS050MW035				
Date	PCE	TCE	cis-1,2-DCE	VC
Baseline (07/13)	0.27 U	**28.3**	1.5	0.34 U
Y2Q4 (8/15)	0.2 U	0.2 U	3.9	**2.9**
June 2016	0.2 U	2	2.5	**8.4**

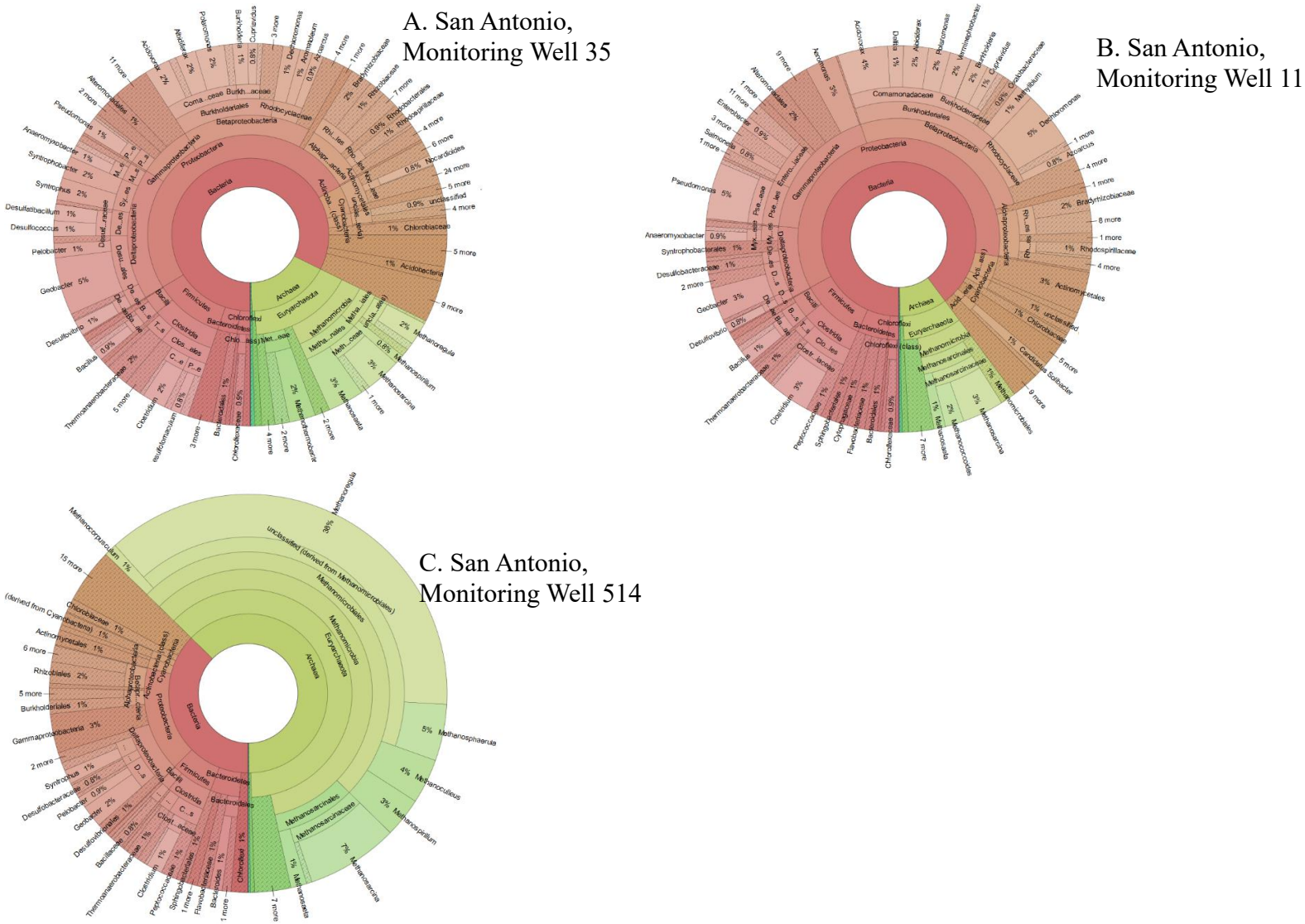


Supplementary Figure 2. 16. Rarefaction curves for microbial communities in groundwater and in SDC-9.

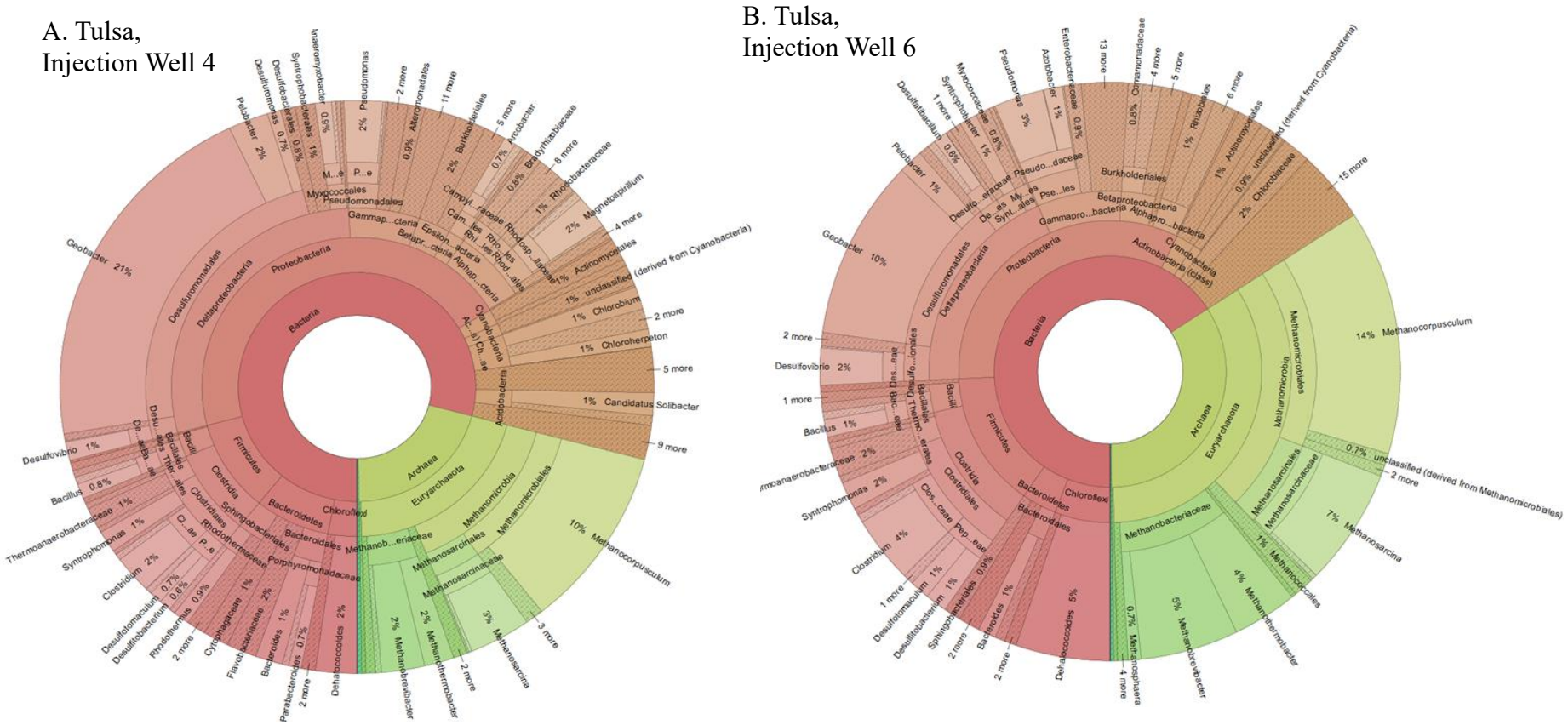
A. SDC-9, replicate 1

B. SDC-9, replicate 2

Supplementary Figure 2. 17. Classification of microbial communities in two samples of SDC-9 (data analyzed with MG-RAST).



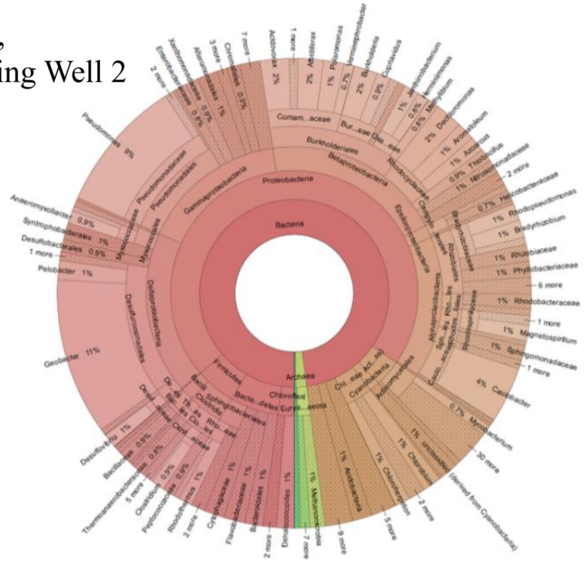
Supplementary Figure 2. 18. Classification of microbial communities in three monitoring well groundwater samples from San Antonio (data analyzed with MG-RAST).



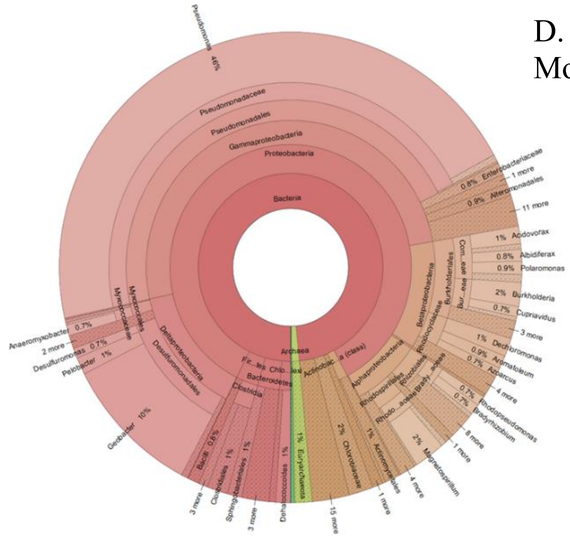
Supplementary Figure 2. 19. Classification of microbial communities in injection well (A and B) and monitoring well (C, D and E) groundwater samples from Tulsa (data analyzed with MG-RAST).

Supplementary Figure 2. 19. (continued)

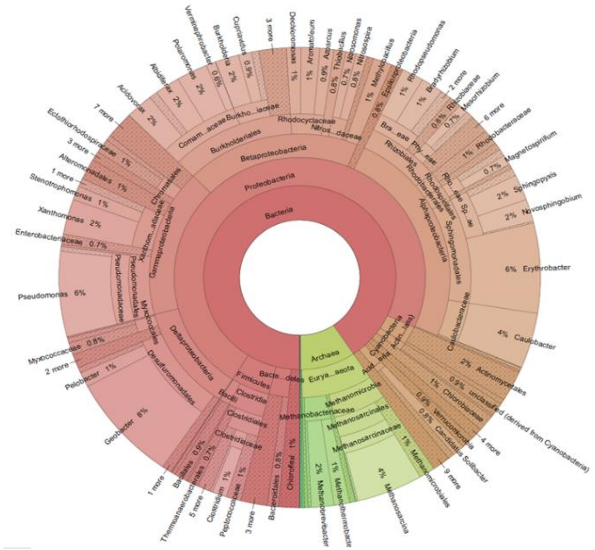
C. Tulsa,
Monitoring Well 2

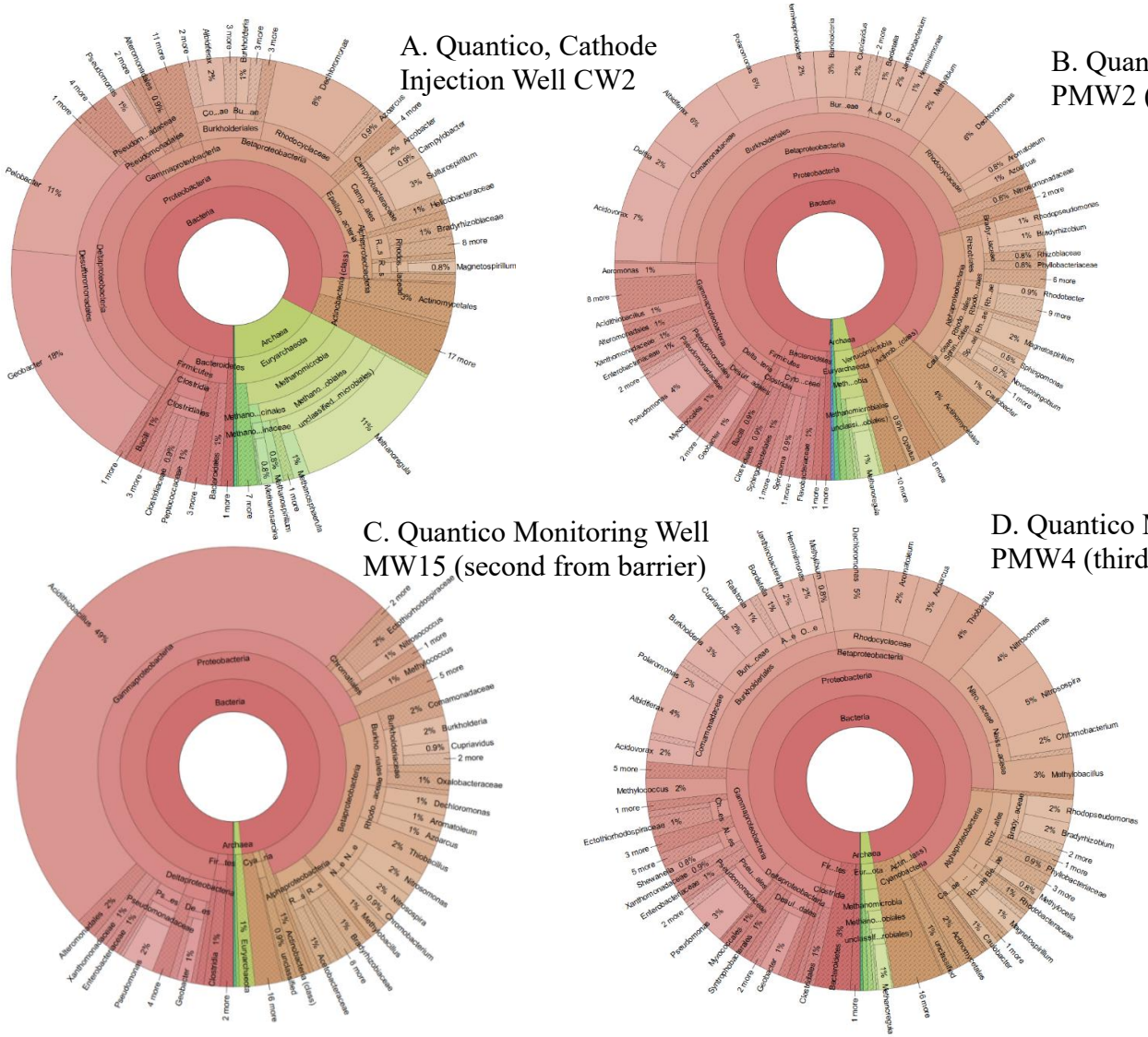


D. Tulsa,
Monitoring Well 3

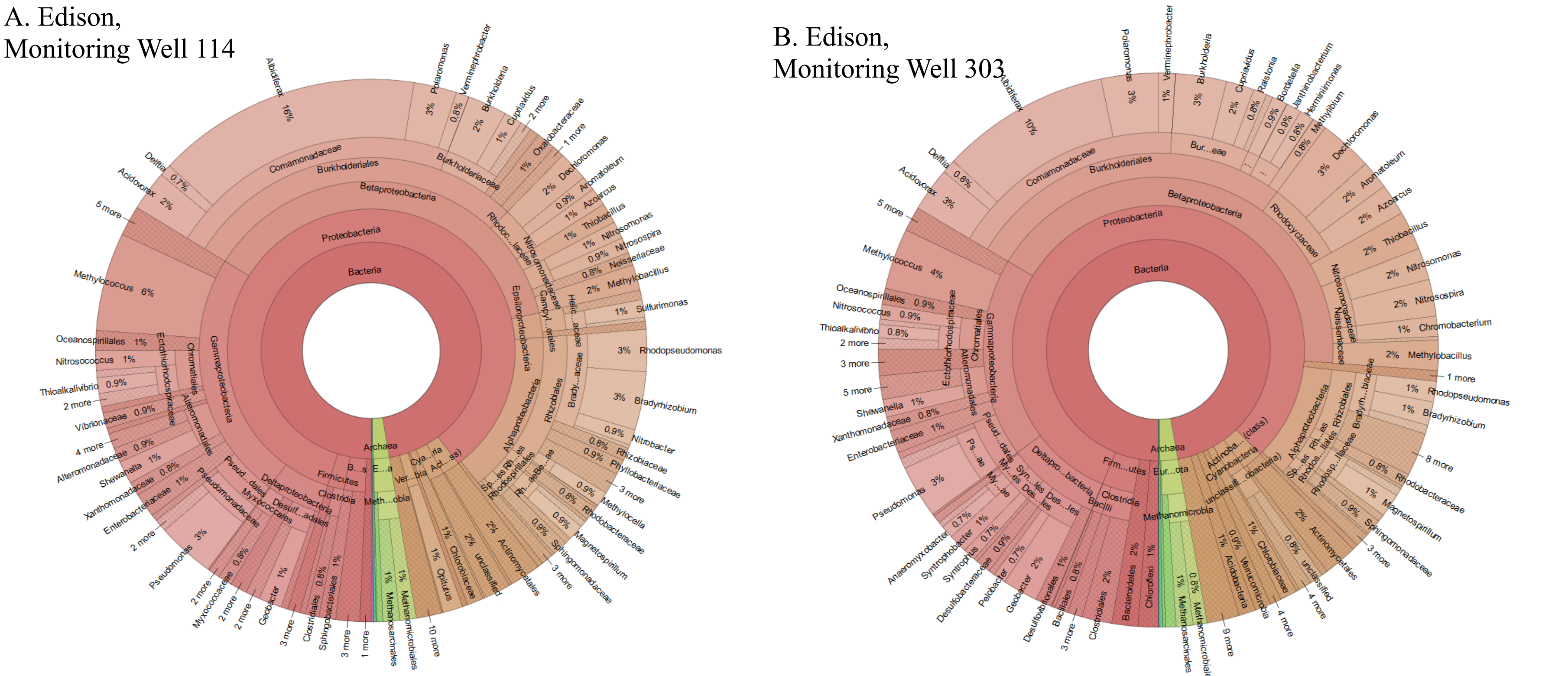


E. Tulsa,
Monitoring Well 4

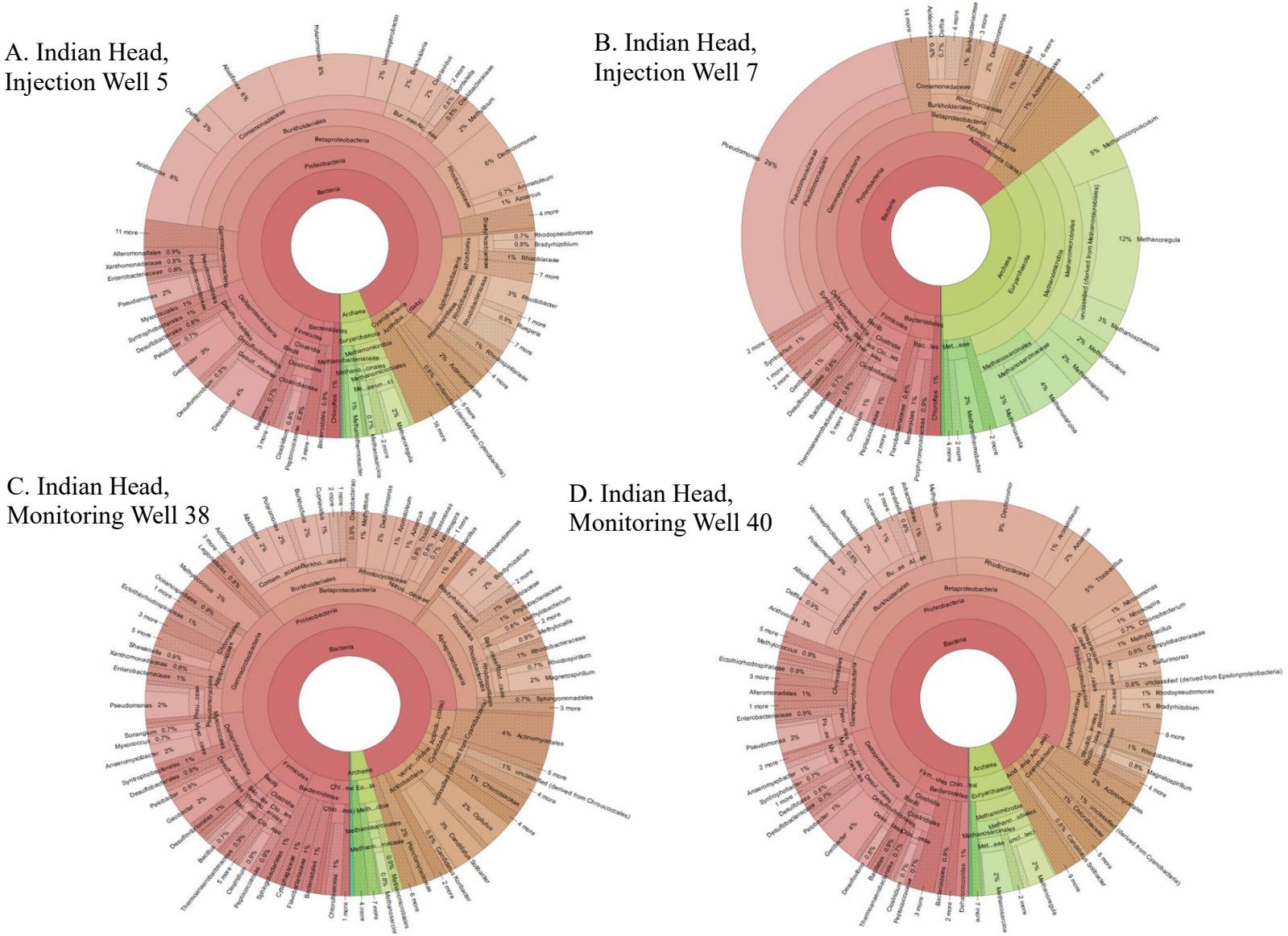




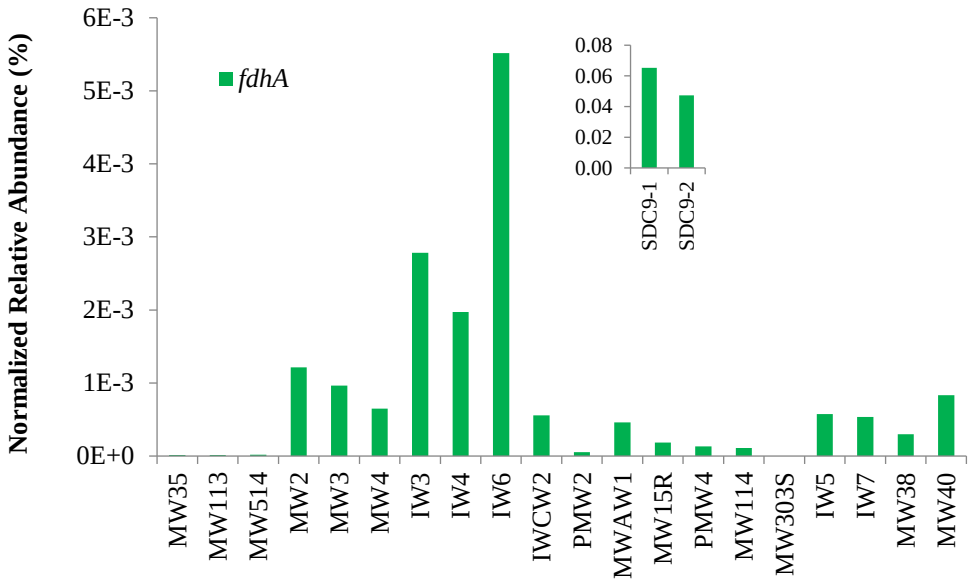
Supplementary Figure 2. 20. Classification of microbial communities in groundwater injection well (A) and monitoring well (B, C, D) samples from Quantico (data analyzed with MG-RAST).



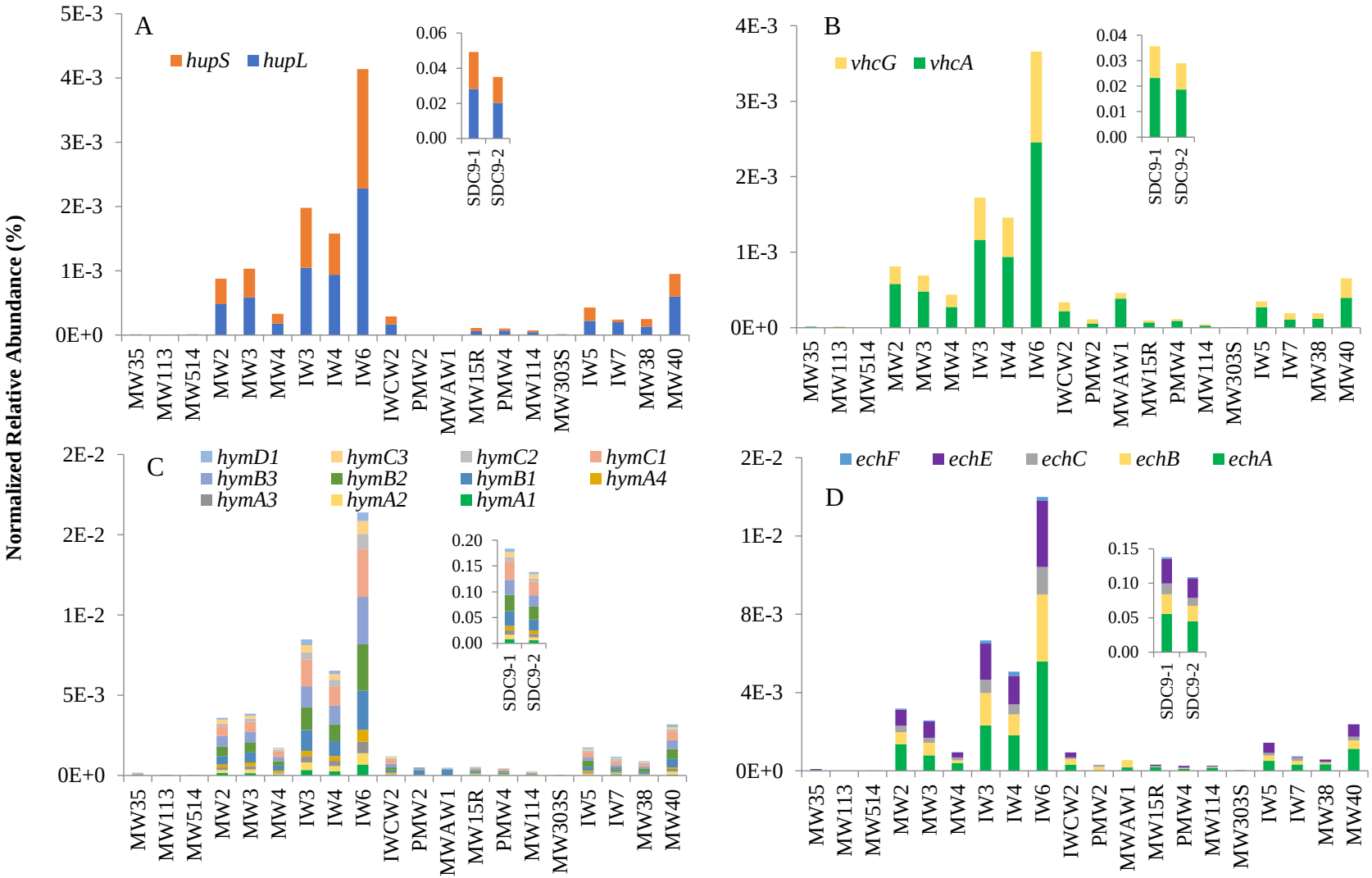
Supplementary Figure 2. 21. Classification of microbial communities in groundwater monitoring well samples from Edison (data analyzed with MG-RAST).



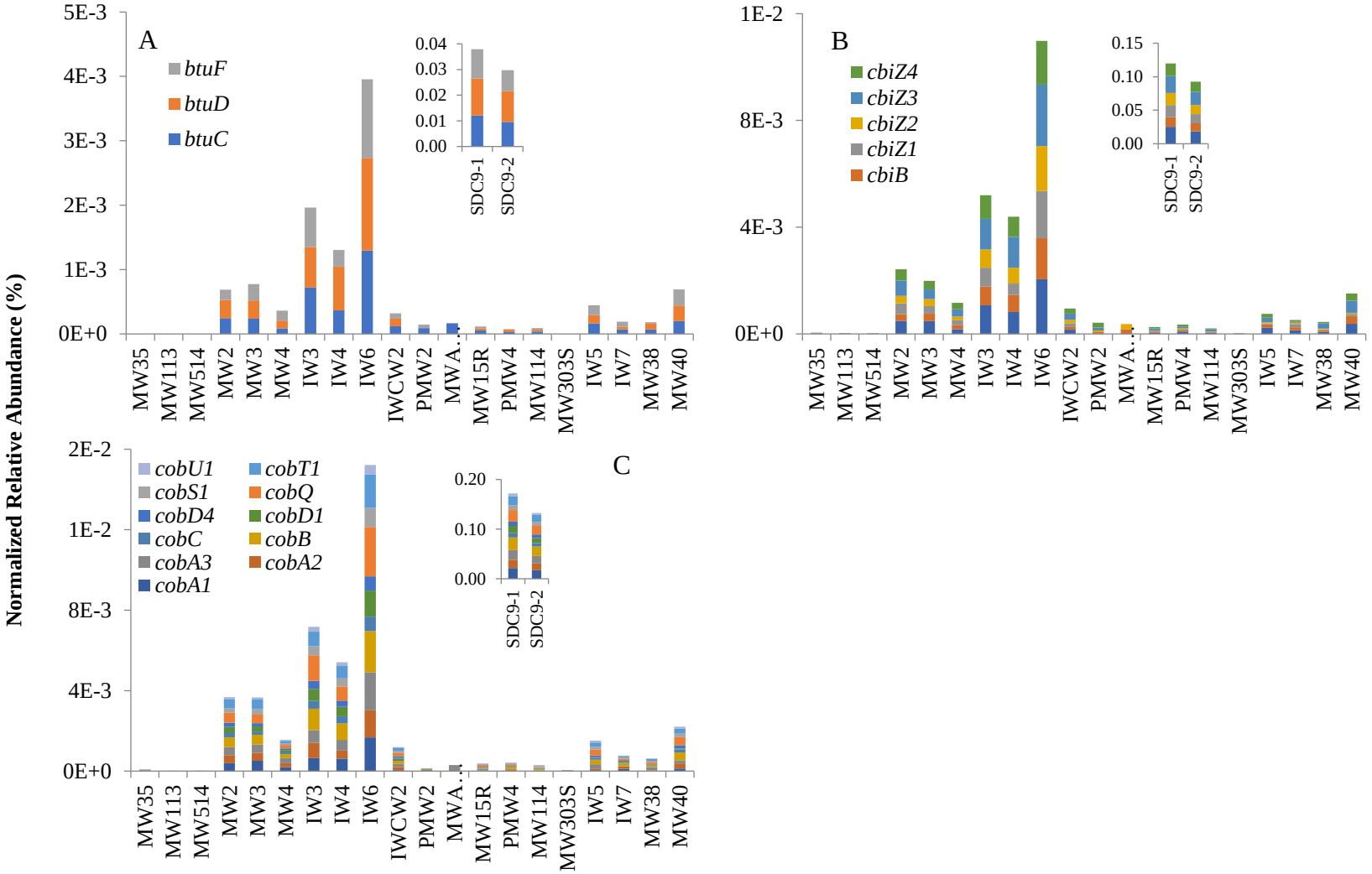
Supplementary Figure 2. 22. Classification of microbial communities in groundwater injection (A, B) and monitoring well (C, D) samples from Indian Head (data analyzed with MG-RAST).



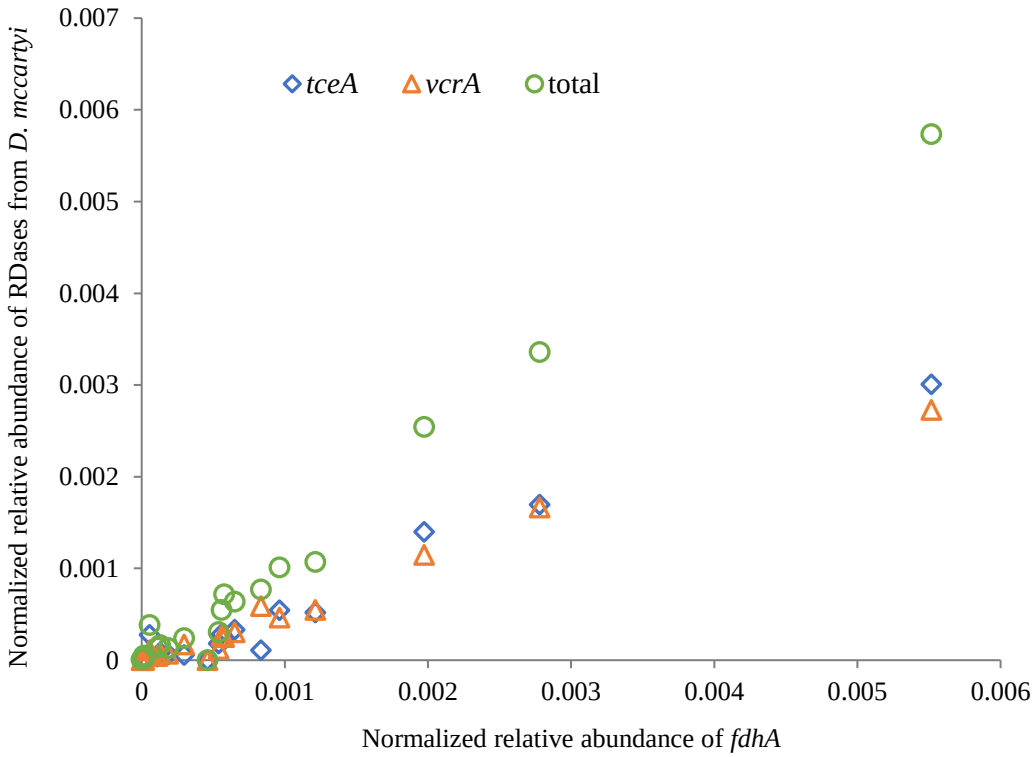
Supplementary Figure 2. 23. Normalized relative abundance (%) of *fdhA* in SDC-9 (insert) and in groundwater from the five chlorinated solvent sites (data analyzed with DIAMOND).



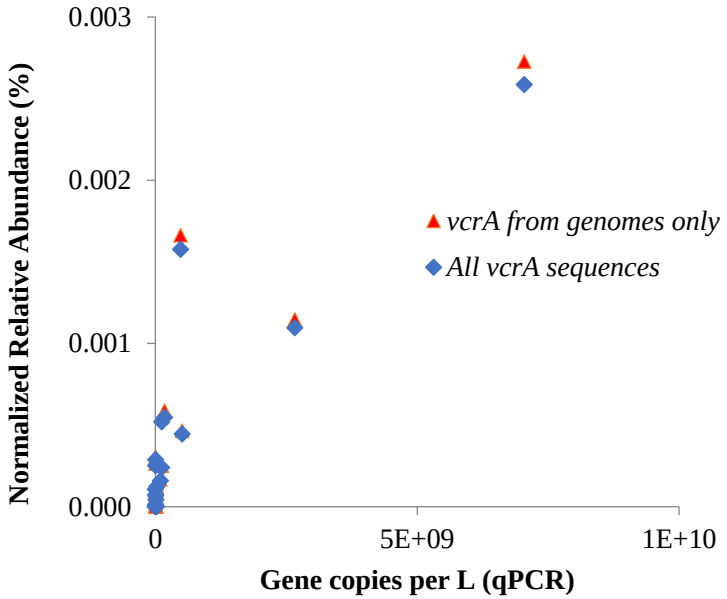
Supplementary Figure 2. 24. Normalized relative abundance (%) of *Dehalococcoidies mccartyi* hydrogenase genes *hupLS* (A), *vhcAG* (B), *hymABCD* (C) and *echABCEF* (D) in SDC-9 (inserts) and in groundwater from the five chlorinated solvent sites (data analyzed with DIAMOND).



Supplementary Figure 2. 25. Normalized relative abundance (%) of *Dehalococcoidies mccartyi* corrinoid metabolism genes *btuFCD* (A), *cbiA*, *cbiB*, *cbiZ* (B) and *cobA*, *cobB*, *cobC*, *cobD*, *cobQ*, *cobS*, *cobT*, *cobU* (C) in SDC-9 (inserts) and in groundwater from the five chlorinated solvent sites (data analyzed with DIAMOND).



Supplementary Figure 2. 26. Comparison between normalized relative abundance of *vcrA*, *tceA* and sum of RDases to *fdhA* (data analyzed with DIAMOND).



Supplementary Figure 2. 27. Comparison between *vcrA* gene copies (per L) determined via qPCR and shotgun sequencing (normalized relative abundance, %, MG-RAST). The results from two shotgun sequencing quantification methods are shown (as discussed in the text).

REFERENCES

REFERENCES

1. Steffan RJ, Vainberg S. 2013. Production and handling of *Dehalococcoides* bioaugmentation cultures, p 89-115. *In* Stroo HF, Leeson A, Ward CH (ed), Bioaugmentation for Groundwater Remediation. Springer, New York.

2. Muller JA, Rosner BM, von Abendroth G, Meshulam-Simon G, McCarty PL, Spormann AM. 2004. Molecular identification of the catabolic vinyl chloride reductase from *Dehalococcoides* sp strain VS and its environmental distribution. Applied and Environmental Microbiology 70:4880-4888.

3. He JZ, Ritalahti KM, Aiello MR, Loffler FE. 2003. Complete detoxification of vinyl chloride by an anaerobic enrichment culture and identification of the reductively dechlorinating population as a *Dehalococcoides* species. Applied and Environmental Microbiology 69:996-1003.

4. Maymo-Gatell X, Anguish T, Zinder SH. 1999. Reductive dechlorination of chlorinated ethenes and 1,2-dichloroethane by "*Dehalococcoides ethenogenes*" 195. Applied and Environmental Microbiology 65:3108-3113.

5. He J, Sung Y, Krajmalnik-Brown R, Ritalahti KM, Loffler FE. 2005. Isolation and characterization of *Dehalococcoides* sp strain FL2, a trichloroethene (TCE)- and 1,2-dichloroethene-respiring anaerobe. Environmental Microbiology 7:1442-1450.

6. Krajmalnik-Brown R, Holscher T, Thomson IN, Saunders FM, Ritalahti KM, Loffler FE. 2004. Genetic identification of a putative vinyl chloride reductase in *Dehalococcoides* sp strain BAV1. Applied and Environmental Microbiology 70:6347-6351.

7. Lyon DY, Vogel TM. 2013. Bioaugmentation for groundwater remediation: an overview, p 1-38. *In* Stroo HF, Leeson A, Ward CH (ed), Bioaugmentation for groundwater remediation. Springer, New York.

8. Kanitkar YH, Stedtfeld RD, Hatzinger PB, Hashsham SA, Cupples AM. 2017. Development and application of a rapid, user-friendly, and inexpensive method to detect *Dehalococcoides* sp reductive dehalogenase genes from groundwater. Applied Microbiology and Biotechnology 101:4827-4835.

9. Perez-de-Mora A, Zila A, McMaster ML, Edwards EA. 2014. Bioremediation of chlorinated ethenes in fractured bedrock and associated changes in dechlorinating and nondechlorinating microbial populations. Environmental Science & Technology 48:5770-5779.

10. Petrovskis EA, Amber WR, Walker CB. 2013. Microbial monitoring during bioaugmentation with *Dehalococcoides*, p 171-197. *In* Stroo HF, Leeson A, Ward CH (ed), Bioaugmentation for groundwater remediation. Springer, New York.

11. Hatt JK, Loffler FE. 2012. Quantitative real-time PCR (qPCR) detection chemistries affect enumeration of the *Dehalococcoides* 16S rRNA gene in groundwater. Journal of Microbiological Methods 88:263-270.

12. Lee PKH, Macbeth TW, Sorenson KS, Deeb RA, Alvarez-Cohen L. 2008. Quantifying genes and transcripts to assess the in situ physiology of "*Dehalococcoides*" spp. in a trichloroethene-contaminated groundwater site. Applied and Environmental Microbiology 74:2728-2739.

13. Liang Y, Liu XK, Singletary MA, Wang K, Mattes TE. 2017. Relationships between the abundance and expression of functional genes from vinyl chloride (VC)-degrading bacteria and geochemical parameters at VC-contaminated sites. Environmental Science & Technology 51:12164-12174.

14. van der Zaan B, Hannes F, Hoekstra N, Rijnaarts H, de Vos WM, Smidt H, Gerritse J. 2010. Correlation of *Dehalococcoides* 16S rRNA and chloroethene-reductive dehalogenase genes with geochemical conditions in chloroethene-contaminated groundwater. Applied and Environmental Microbiology 76:843-850.

15. Munro JE, Kimyon O, Rich DJ, Koenig J, Tang SH, Low A, Lee M, Manefield M, Coleman NV. 2017. Co-occurrence of genes for aerobic and anaerobic biodegradation of dichloroethane in organochlorine-contaminated groundwater. Fems Microbiology Ecology 93.

16. Kotik M, Davidova A, Voriskova J, Baldrian P. 2013. Bacterial communities in tetrachloroethene-polluted groundwaters: A case study. Science of the Total Environment 454:517-527.

17. Nemecek J, Dolinova I, Machackova J, Spanek R, Sevcu A, Lederer T, Cernik M. 2017. Stratification of chlorinated ethenes natural attenuation in an alluvial aquifer assessed by hydrochemical and biomolecular tools. Chemosphere 184:1157-1167.

18. Simsir B, Yan J, Im J, Graves D, Loffler FE. 2017. Natural attenuation in streambed sediment receiving chlorinated solvents from underlying fracture networks. Environmental Science & Technology 51:4821-4830.

19. Atashgahi S, Lu Y, Zheng Y, Saccenti E, Suarez-Diez M, Ramiro-Garcia J, Eisenmann H, Elsner M, Stams AJM, Springael D, Dejonghe W, Smidt H. 2017. Geochemical and microbial community determinants of reductive dechlorination at a site biostimulated with glycerol. Environmental Microbiology 19:968-981.

20. Kao CM, Liao HY, Chien CC, Tseng YK, Tang P, Lin CE, Chen SC. 2016. The change of microbial community from chlorinated solvent-contaminated groundwater after biostimulation using the metagenome analysis. Journal of Hazardous Materials 302:144-150.

21. Dugat-Bony E, Biderre-Petit C, Jaziri F, David MM, Denonfoux J, Lyon DY, Richard JY, Curvers C, Boucher D, Vogel TM, Peyretaillade E, Peyret P. 2012. In situ TCE degradation mediated by complex dehalorespiring communities during biostimulation processes. Microbial Biotechnology 5:642-653.

22. Nemecek J, Pokorny P, Lhotsky O, Knytl V, Najmanova P, Steinova J, Cernik M, Filipova A, Filip J, Cajthaml T. 2016. Combined nano-biotechnology for in-situ remediation of mixed contamination of groundwater by hexavalent chromium and chlorinated solvents. Science of the Total Environment 563:822-834.

23. Kocur CMD, Lomheim L, Molenda O, Weber KP, Austrins LM, Sleep BE, Boparai HK, Edwards EA, O'Carroll DM. 2016. Long-term field study of microbial community and dechlorinating activity

following carboxymethyl cellulose-stabilized nanoscale zero-valent iron injection. Environmental Science & Technology 50:7658-7670.

24. Nemecek J, Steinova J, Spanek R, Pluhar T, Pokorny P, Najmanova P, Knytl V, Cernik M. 2018. Thermally enhanced in situ bioremediation of groundwater contaminated with chlorinated solvents - A field test. Science of the Total Environment 622:743-755.

25. Badin A, Broholm MM, Jacobsen CS, Palau J, Dennis P, Hunkeler D. 2016. Identification of abiotic and biotic reductive dechlorination in a chlorinated ethene plume after thermal source remediation by means of isotopic and molecular biology tools. Journal of Contaminant Hydrology 192:1-19.

26. Reiss RA, Guerra P, Makhnin O. 2016. Metagenome phylogenetic profiling of microbial community evolution in a tetrachloroethen-contaminated aquifer responding to enhanced reductive dechlorination protocols. Standards in Genomic Sciences 11.

27. Adetutu EM, Gundry TD, Patil SS, Golneshin A, Adigun J, Bhaskarla V, Aleer S, Shahsavari E, Ross E, Ball AS. 2015. Exploiting the intrinsic microbial degradative potential for field-based in situ dechlorination of trichloroethene contaminated groundwater. Journal of Hazardous Materials 300:48-57.

28. Weigold P, El-Hadidi M, Ruecker A, Huson DH, Scholten T, Jochmann M, Kappler A, Behrens S. 2016. A metagenomic-based survey of microbial (de)halogenation potential in a German forest soil. Scientific Reports 6.

29. Hug LA, Beiko RG, Rowe AR, Richardson RE, Edwards EA. 2012. Comparative metagenomics of three *Dehalococcoides*-containing enrichment cultures: the role of the non-dechlorinating community. Bmc Genomics 13.

30. Brisson VL, West KA, Lee PKH, Tringe SG, Brodie EL, Alvarez-Cohen L. 2012. Metagenomic analysis of a stable trichloroethene-degrading microbial community. Isme Journal 6:1702-1714.

31. Puls RW, Barcelona MJ. 1996. Low-flow (minimal drawdown) ground-water sampling procedures, EPA/540/S-95/504. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, DC.

32. Schaefer CE, Lippincott DR, Steffan RJ. 2010. Field-scale evaluation of bioaugmentation dosage for treating chlorinated ethenes. Ground Water Monitoring and Remediation 30:113-124.

33. Schaefer CE, Condee CW, Vainberg S, Steffan RJ. 2009. Bioaugmentation for chlorinated ethenes using *Dehalococcoides* sp.: Comparison between batch and column experiments. Chemosphere 75:141-148.

34. Kanitkar YH, Stedtfeld RD, Steffan RJ, Hashsham SA, Cupples AM. 2016. Loop-mediated isothermal amplification (LAMP) for rapid detection and quantification of *Dehalococcoides* biomarker genes in commercial reductive dechlorinating cultures KB-1 and SDC-9. Applied and Environmental Microbiology 82:1799-1806.

35. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards RA. 2008. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. Bmc Bioinformatics 9.

36. Fish JA, Chai BL, Wang Q, Sun YN, Brown CT, Tiedje JM, Cole JR. 2013. FunGene: the functional gene pipeline and repository. Frontiers in Microbiology 4.

37. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nature Methods 12:59-60.

38. Ritalahti KM, Amos BK, Sung Y, Wu QZ, Koenigsberg SS, Loffler FE. 2006. Quantitative PCR targeting 16S rRNA and reductive dehalogenase genes simultaneously monitors multiple *Dehalococcoides* strains. Applied and Environmental Microbiology 72:2765-2774.

39. DiSpirito A, J. , Gulledge JC, Murrell AK, Shiemke ME, Lidstrom, Krema CL. 1992. Trichloroethylene oxidation by the membrane-associated methane monooxygenase in type I, type II, and type x methanotrophs. Biodegradation 2:151-164.

40. Wymore RA, Lee MH, Keener WK, Miller AR, Colwell FS, Watwood ME, Sorenson KS. 2007. Field evidence for intrinsic aerobic chlorinated ethene cometabolism by methanotrophs expressing soluble methane monooxygenase. Biodegradation 11:125-139.

41. Lee SW, Keeney DR, Lim DH, Dispirito AA, Semrau JD. 2006. Mixed pollutant degradation by *Methylosinus trichosporium* OB3b expressing either soluble or particulate methane monooxygenase: Can the tortoise beat the hare? Applied and Environmental Microbiology 72:7503-7509.

42. He Y, Mathieu J, Yang Y, Yu PF, da Silva MLB, Alvarez PJJ. 2017. 1,4-Dioxane biodegradation by *Mycobacterium dioxanotrophicus* PH-06 is associated with a group-6 soluble di-iron monooxygenase. Environmental Science & Technology Letters 4:494-499.

43. Jin YO, Mattes TE. 2010. A quantitative PCR assay for aerobic, vinyl chloride- and ethene-assimilating microorganisms in groundwater. Environmental Science & Technology 44:9036-9041.

44. Morris RM, Fung JM, Rahm BG, Zhang S, Freedman DL, Zinder SH, Richardson RE. 2007. Comparative proteomics of *Dehalococcoides* spp. reveals strain-specific peptides associated with activity. Appl Environ Microbiol 73:320-6.

45. Turkowsky D, Jehmlich N, Diekert G, Adrian L, von Bergen M, Goris T. 2018. An integrative overview of genomic, transcriptomic and proteomic analyses in organohalide respiration research. Fems Microbiology Ecology 94.

46. Villemur R, Saucier M, Gauthier A, Beaudet R. 2002. Occurrence of several genes encoding putative reductive dehalogenases in *Desulfitobacterium hafniense/frappieri* and *Dehalococcoides ethenogenes*. Canadian Journal of Microbiology 48:697-706.

47. Zhao SY, Ding C, He JZ. 2015. Detoxification of 1,1,2-trichloroethane to ethene by *Desulfitobacterium* and identification of its functional reductase gene. Plos One 10.

48. Utkin I, Woese C, Wiegel J. 1994. Isolation and characterization of *Desulfitobacterium dehalogenans* gen-nov, sp-nov, an anaerobic bacterium which reductively dechlorinates chlorophenolic compounds. International Journal of Systematic Bacteriology 44:612-619.

49. Furukawa K, Suyama A, Tsuboi Y, Futagami T, Goto M. 2005. Biochemical and molecular characterization of a tetrachloroethene dechlorinating *Desulfitobacterium* sp strain Y51: a review. Journal of Industrial Microbiology & Biotechnology 32:534-541.

50. Gerritse J, Drzyzga O, Kloetstra G, Keijmel M, Wiersum LP, Hutson R, Collins MD, Gottschal JC. 1999. Influence of different electron donors and accepters on dehalorespiration of tetrachloroethene by *Desulfitobacterium frappieri* TCE1. Applied and Environmental Microbiology 65:5212-5221.

51. Yan J, Ritalahti KM, Wagner DD, Loffler FE. 2012. Unexpected specificity of interspecies cobamide transfer from *Geobacter* spp. to organohalide-respiring *Dehalococcoides mccartyi* strains. Applied and Environmental Microbiology 78:6630-6636.

52. Sung Y, Fletcher KF, Ritalaliti KM, Apkarian RP, Ramos-Hernandez N, Sanford RA, Mesbah NM, Loffler FE. 2006. *Geobacter lovleyi* sp nov strain SZ, a novel metal-reducing and tetrachloroethene-dechlorinating bacterium. Applied and Environmental Microbiology 72:2775-2782.

53. Doong RA, Lee CC, Lien CM. 2014. Enhanced dechlorination of carbon tetrachloride by *Geobacter sulfurreducens* in the presence of naturally occurring quinones and ferrihydrite. Chemosphere 97:54-63.

54. Men YJ, Yu K, Baelum J, Gao Y, Tremblay J, Prestat E, Stenuit B, Tringe SG, Jansson J, Zhang T, Alvarez-Cohen L. 2017. Metagenomic and metatranscriptomic analyses reveal the structure and dynamics of a dechlorinating community containing *Dehalococcoides mccartyi* and corrinoid-providing microorganisms under cobalamin-limited conditions. Applied and Environmental Microbiology 83.

55. Men YJ, Feil H, VerBerkmoes NC, Shah MB, Johnson DR, Lee PKH, West KA, Zinder SH, Andersen GL, Alvarez-Cohen L. 2012. Sustainable syntrophic growth of *Dehalococcoides ethenogenes* strain 195 with *Desulfovibrio vulgaris* Hildenborough and *Methanobacterium congolense*: global transcriptomic and proteomic analyses. Isme Journal 6:410-421.

56. Mattes TE, Alexander AK, Richardson PM, Munk AC, Han CS, Stothard P, Coleman NV. 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. Applied and Environmental Microbiology 74:6405-6416.

57. Wagner DD, Hug LA, Hatt JK, Spitzmiller MR, Padilla-Crespo E, Ritalahti KM, Edwards EA, Konstantinidis KT, Loffler FE. 2012. Genomic determinants of organohalide-respiration in *Geobacter lovleyi*, an unusual member of the *Geobacteraceae*. Bmc Genomics 13.

58. Liang Y, Cook LJ, Mattes TE. 2017. Temporal abundance and activity trends of vinyl chloride (VC)-degrading bacteria in a dilute VC plume at Naval Air Station Oceana. Environmental Science and Pollution Research 24:13760-13774.

59. Mahendra S, Alvarez-Cohen L. 2006. Kinetics of 1,4-dioxane biodegradation by monooxygenase-expressing bacteria. Environmental Science & Technology 40:5435-5442.

60. Hatzinger PB, Banerjee R, Rezes R, Streger SH, McClay K, Schaefer CE. 2017. Potential for cometabolic biodegradation of 1,4-dioxane in aquifers with methane or ethane as primary substrates. Biodegradation 28:453-468.

61. Li MY, Mathieu J, Liu YY, Van Orden ET, Yang Y, Fiorenza S, Alvarez PJJ. 2014. The abundance of tetrahydrofuran/dioxane monooxygenase genes (*thmA/dxmA*) and 1,4-dioxane degradation activity are significantly correlated at various impacted aquifers. Environmental Science & Technology Letters 1:122-127.

62. Gedalanga P, Madison A, Miao Y, Richards T, Hatton J, DiGuiseppi WH, Wilson J, Mahendra S. 2016. A multiple lines of evidence framework to evaluate intrinsic biodegradation of 1,4-dioxane. Remediation-the Journal of Environmental Cleanup Costs Technologies & Techniques 27:93-114.

63. da Silva MLB, Woroszylo C, Castillo NF, Adamson DT, Alvarez PJJ. 2018. Associating potential 1,4-dioxane biodegradation activity with groundwater geochemical parameters at four different contaminated sites. Journal of Environmental Management 206:60-64.

64. Rahm BG, Richardson RE. 2008. *Dehalococcoides*' gene transcripts as quantitative bioindicators of tetrachloroethene, trichloroethene, and cis-1,2-dichloroethene dehalorespiration rates. Environmental Science & Technology 42:5099-5105.

65. Rahm BG, Richardson RE. 2008. Correlation of respiratory gene expression levels and pseudo-steady-state PCE respiration rates in *Dehalococcoides ethenogenes*. Environmental Science & Technology 42:416-421.

66. Heavner GLW, Mansfeldt CB, Debs GE, Hellerstedt ST, Rowe AR, Richardson RE. 2018. Biomarkers' responses to reductive dechlorination rates and oxygen stress in bioaugmentation culture KB-1 (TM). Microorganisms 6.

67. Hartmans S, Debont JAM. 1992. Aerobic vinyl chloride metabolism in *Mycobacterium aurum* L1. Applied and Environmental Microbiology 58:1220-1226.

68. Danko AS, Saski CA, TomkinS JP, Freedman DL. 2006. Involvement of coenzyme M during aerobic biodegradation of vinyl chloride and ethene by *Pseudomonas putida* strain AJ and *Ochrobactrum* sp strain TD. Applied and Environmental Microbiology 72:3756-3758.

69. Coleman NV, Spain JC. 2003. Epoxyalkane: Coenzyme M transferase in the ethene and vinyl chloride biodegradation pathways of *Mycobacterium* strain JS60. Journal of Bacteriology 185:5536-5545.

70. Mattes TE, Coleman NV, Spain JC, Gossett JM. 2005. Physiological and molecular genetic analyses of vinyl chloride and ethene biodegradation in *Nocardioides* sp strain JS614. Archives of Microbiology 183:95-106.

71. Coleman NV, Spain JC. 2003. Distribution of the coenzyme m pathway of epoxide metabolism among ethene- and vinyl chloride-degrading *Mycobacterium* strains. Applied and Environmental Microbiology 69:6041-6046.

72. Nishino SF, Shin KA, Gossett JM, Spain JC. 2013. Cytochrome P450 initiates degradation of cis-dichloroethene by *Polaromonas* sp strain JS666. Applied and Environmental Microbiology 79:2263-2272.

73. Yoon S, Im J, Bandow N, DiSpirito AA, Semrau JD. 2011. Constitutive expression of pMMO by *Methylocystis* strain SB2 when grown on multi-carbon substrates: implications for biodegradation of chlorinated ethenes. Environmental Microbiology Reports 3:182-188.

74. Chang HL, Alvarez-Cohen L. 1996. Biodegradation of individual and multiple chlorinated aliphatic hydrocarbons by methane-oxidizing cultures. Applied and Environmental Microbiology 62:3371-3377.

75. DeRosa CT, Wilbur S, Holler J, Richter P, Stevens YW. 1996. Health evaluation of 1,4-dioxane. Toxicology and Industrial Health 12:1-43.

76. Mohr T, Stickney J, DiGuiseppi W. 2010. Environmental Investigation and Remediation: 1,4-Dioxane and Other Solvent Stabilizers. CRC Press, Boca Raton, ML.

77. Adamson DT, Anderson RH, Mahendra S, Newell CJ. 2015. Evidence of 1,4-dioxane attenuation at groundwater sites contaminated with chlorinated solvents and 1,4-dioxane. Environmental Science & Technology 49:6510-6518.

78. Adamson DT, Mahendra S, Walker KL, Rauch SR, Sengupta S, Newell CJ. 2014. A multisite survey to identify the scale of the 1,4-dioxane problem atcontaminated groundwater sites. Environmental Science & Technology Letters 1:254-258.

79. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. Journal of Molecular Biology 215:403-410.

80. Yang Y, Higgins SA, Yan J, Simsir B, Chourey K, Iyer R, Hettich RL, Baldwin B, Ogles DM, Loffler FE. 2017. Grape pomace compost harbors organohalide-respiring *Dehalogenimonas* species with novel reductive dehalogenase genes. Isme Journal 11:2767-2780.

81. Cox MP, Peterson DA, Biggs PJ. 2010. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. Bmc Bioinformatics 11:485:1-6.

82. Rho MN, Tang HX, Ye YZ. 2010. FragGeneScan: predicting genes in short and error-prone reads. Nucleic Acids Research 38.

83. Pruitt KD, Tatusova T, Maglott DR. 2005. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Research 33:D501-D504.

84. Ondov BD, Bergman NH, Phillippy AM. 2011. Interactive metagenomic visualization in a Web browser. Bmc Bioinformatics 12.

85. Sanford RA, Cole JR, Tiedje JM. 2002. Characterization and description of *Anaeromyxobacter dehalogenans* gen. nov., sp nov., an aryl-halorespiring facultative anaerobic myxobacterium. Applied and Environmental Microbiology 68:893-900.

86. Kube M, Beck A, Zinder SH, Kuhl H, Reinhardt R, Adrian L. 2005. Genome sequence of the chlorinated compound respiring bacterium *Dehalococcoides* species strain CBDB1. Nature Biotechnology 23:1269-1273.

87. Loffler FE, Yan J, Ritalahti KM, Adrian L, Edwards EA, Konstantinidis KT, Muller JA, Fullerton H, Zinder SH, Spormann AM. 2015. *Dehalococcoides mccartyi* gen. nov., sp. nov., obligately organohalide-respiring anaerobic bacteria relevant to halogen cycling and bioremediation, belong to a novel bacterial class, *Dehalococcoidia c*lassis nov., order *Dehalococcoidales* ord. nov. and family *Dehalococcoidaceae* fam. nov., within the phylum *Chloroflexi* (vol 63, pg 625, 2013). International Journal of Systematic and Evolutionary Microbiology 65:2015-2015.

88. Coleman NV, Wilson NL, Barry K, Brettin TS, Bruce DC, Copeland A, Dalin E, Detter JC, del Rio TG, Goodwin LA, Hammon NM, Han SS, Hauser LJ, Israni S, Kim E, Kyrpides N, Land ML, Lapidus A, Larimer FW, Lucas S, Pitluck S, Richardson P, Schmutz J, Tapia R, Thompson S, Tice HN, Spain JC, Gossett JG, Mattes TE. 2011. Genome sequence of the ethene- and vinyl chloride-oxidizing *Actinomycete Nocardioides* sp strain JS614. Journal of Bacteriology 193:3399-3400.

89. Luijten MLGC, de Weert J, Smidt H, Boschker HTS, de Vos WM, Schraa G, Stams AJM. 2003. Description of *Sulfurospirillum halorespirans* sp nov., an anaerobic, tetrachloroethene-respiring bacterium, and transfer of *Dehalospirillum multivorans* to the genus *Sulfurospirillum* as *Sulfurospirillum multivorans* comb. nov. International Journal of Systematic and Evolutionary Microbiology 53:787-793.

90. Goris T, Schiffmann CL, Gadkari J, Schubert T, Seifert J, Jehmlich N, Von Bergen M, Diekert G. 2015. Proteomics of the organohalide-respiring *Epsilonproteobacterium Sulfurospirillum multivorans* adapted to tetrachloroethene and other energy substrates. Scientific Reports 5.

91. Goris T, Schubert T, Gadkari J, Wubet T, Tarkka M, Buscot F, Adrian L, Diekert G. 2014. Insights into organohalide respiration and the versatile catabolism of *Sulfurospirillum multivorans* gained from comparative genomics and physiological studies. Environmental Microbiology 16:3562-3580.

92. Zhang S, Wondrousch D, Cooper M, Zinder SH, Schuurmann G, Adrian L. 2017. Anaerobic dehalogenation of chloroanilines by *Dehalococcoides mccartyi* strain CBDB1 and *Dehalobacter* strain 14DCB1 via different pathways as related to molecular electronic structure. Environmental Science & Technology 51:3714-3724.

93. Holliger C, Hahn D, Harmsen H, Ludwig W, Schumacher W, Tindall B, Vazquez F, Weiss N, Zehnder AJB. 1998. *Dehalobacter restrictus* gen. nov. and sp. nov., a strictly anaerobic bacterium that reductively dechlorinates tetra- and trichloroethene in an anaerobic respiration. Archives of Microbiology 169:313-321.

94. Grostern A, Edwards EA. 2006. Growth of *Dehalobacter* and *Dehalococcoides* spp. during degradation of chlorinated ethanes. Applied and Environmental Microbiology 72:428-436.

95. Deweerd KA, Mandelco L, Tanner RS, Woese CR, Suflita JM. 1990. *Desulfomonile tiedjei* gen-nov and sp-nov, a novel anaerobic, dehalogenating, sulfate-reducing bacterium. Archives of Microbiology 154:23-30.

96. Sun BL, Cole JR, Tiedje JM. 2001. *Desulfomonile limimaris* sp nov., an anaerobic dehalogenating bacterium from marine sediments. International Journal of Systematic and Evolutionary Microbiology 51:365-371.

97. Krumholz LR. 1997. *Desulfuromonas chloroethenica* sp. nov. uses tetrachloroethylene and trichloroethylene as electron acceptors. International Journal of Systematic Bacteriology 47:1262-1263.

98. Sung Y, Ritalahti KM, Sanford RA, Urbance JW, Flynn SJ, Tiedje JM, Loffler FE. 2003. Characterization of two tetrachloroethene-reducing, acetate-oxidizing anaerobic bacteria and their description as *Desulfuromonas michiganensis* sp nov. Applied and Environmental Microbiology 69:2964-2974.

99. Chang YC, Ikeutsu K, Toyama T, Choi D, Kikuchi S. 2011. Isolation and characterization of tetrachloroethylene- and cis-1,2-dichloroethylene-dechlorinating propionibacteria. Journal of Industrial Microbiology & Biotechnology 38:1667-1677.

100. Coleman NV, Mattes TE, Gossett JM, Spain JC. 2002. Phylogenetic and kinetic diversity of aerobic vinyl chloride-assimilating bacteria from contaminated sites. Applied and Environmental Microbiology 68:6162-6171.

101. Wild A, Hermann R, Leisinger T. 1997. Isolation of an anaerobic bacterium which reductively dechlorinates tetrachloroethene and trichloroethene. Biodegradation 7:507-511.

102. Fathepure BZ, Tiedje JM. 1994. Reductive dechlorination of tetrachloroethylene by a chlorobenzoate-enriched biofilm reactor. Environmental Science & Technology 28:746-752.

103. Mortan SH, Martin-Gonzalez L, Vicenta T, Caminal G, Nijenhuis I, Adrian L, Marco-Urrea E. 2017. Detoxification of 1,1,2-trichloroethane to ethene in a bioreactor co-culture of *Dehalogenimonas* and *Dehalococcoides mccartyi* strains. Journal of Hazardous Materials 331:218-225.

104. Moe WM, Yan J, Nobre MF, da Costa MS, Rainey FA. 2009. *Dehalogenimonas lykanthroporepellens* gen. nov., sp nov., a reductively dehalogenating bacterium isolated from chlorinated solvent-contaminated groundwater. International Journal of Systematic and Evolutionary Microbiology 59:2692-2697.

105. Bowman KS, Nobre MF, da Costa MS, Rainey FA, Moe WM. 2013. *Dehalogenimonas alkenigignens* sp nov., a chlorinated-alkane-dehalogenating bacterium isolated from groundwater. International Journal of Systematic and Evolutionary Microbiology 63:1492-1498.

106. Molenda O, Quaile AT, Edwards EA. 2016. *Dehalogenimonas* sp strain WBC-2 genome and identification of its trans-dichloroethene reductive dehalogenase, *TdrA*. Applied and Environmental Microbiology 82:40-50.

107. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-20.

108. Kanitkar YH, Stedtfeld RD, Steffan RJ, Hashsham SA, Cupples AM. 2016. Loop-Mediated Isothermal Amplification (LAMP) for Rapid Detection and Quantification of Dehalococcoides Biomarker Genes in Commercial Reductive Dechlorinating Cultures KB-1 and SDC-9. Appl Environ Microbiol 82:1799-1806.

CHAPTER 3 Diversity and Abundance of the Functional Genes and Bacteria Associated with RDX Degradation at a Contaminated Site Pre- and Post- Biostimulation

This chapter is being prepared for submission to a peer reviewed journal: Hongyu Dang and Alison M. Cupples. Diversity and Abundance of the Functional Genes and Bacteria Associated with RDX Degradation at a Contaminated Site Pre- and Post- Biostimulation.

Abstract

Bioremediation is becoming an increasingly popular approach for the remediation of sites contaminated with the explosive hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX). Multiple lines of evidence are often needed to assess the success of such approaches, with molecular studies frequently providing important information on the abundance of key biodegrading species. Towards this goal, the current study utilized shotgun sequencing to determine the abundance and diversity of functional genes (*xenA, xenB, xplA, diaA, pnrB, nfsI*) and species previously associated with RDX biodegradation in groundwater before and after biostimulation at an RDX contaminated Navy Site. For this, DNA was extracted from four and seven groundwater wells pre- and post- biostimulation, respectively. From a set of 65 previously identified RDX degraders, 31 were found within the groundwater samples, with the most abundant species being *Variovorax* sp. JS1663, *Pseudomonas fluorescens*, *Pseudomonas putida* and *Stenotrophomonas maltophilia*. Further, 9 RDX degrading species significantly ($p<0.05$) increased in abundance following biostimulation. Both the sequencing data and qPCR indicated *xenA* and *xenB* exhibited the highest relative abundance among the six genes. Several genes (*diaA, nsfI, xenA* and *pnrB)* exhibited higher relative abundance values in some wells following biostimulation. The study provides a comprehensive approach for assessing biomarkers during RDX bioremediation and

provides evidence that biostimulation generated a positive impact on a set of key species and genes.

1. Introduction

Hexahydro-1,3,5-trinitro-1,3,5-triazine, also known as Royal Demolition Explosive or RDX, is a synthetic product and commonly used explosive (1). This chemical is classified as a possible human carcinogen and has different impacts on human health (1). Due to the denotation of ordnance, firing of munitions on military training ranges, and the manufacturing and transport of munitions, RDX has caused soil, groundwater and sediment contamination, with 39 identified sites currently on the National Priority List (2). Although a range of approaches (incineration, activated carbon columns and hydrogen peroxide) have been used to remediate RDX contaminated sites (1, 3, 4), more recently, interest has turned to bioremediation because of the potential cost savings. Successful pilot- and full-scale explosives bioremediation demonstrations have occurred at multiple DoD facilities including the Nebraska Ordnance Plant, Cornhusker Army Ammunition Plant, and Milan Army Ammunition Plant (5-10). In a study involving push-pull tests to measure *in situ* RDX degradation rates following the addition of corn syrup, lactose, emulsified oil, and ethanol, biostimulation with corn syrup was the most successful, illustrating 82% RDX removal (11). Bioaugmentation with *Gordonia* sp. strain KTR9 has also been shown to be an effective RDX bioremediation approach (8-10).

A number of bacteria and functional genes have been associated with the biodegradation of RDX under aerobic or anaerobic conditions. A type I nitroreductase (encoded by *nfsI)* present in both *Morganella morganii* strain B2 and *Enterobacter cloacae* strain 96-3 was responsible for the nitroreduction of RDX (12). Type I, or oxygen insensitive, nitroreductases, use a two-electron reduction mechanism to reduce nitro groups under aerobic conditions (13, 14). Another

functional gene, *pnrB*, from a *Pseudomonas* sp. and *Stenotrophomonas maltophilia*, was also associated with RDX degradation (15). The enzyme encoded by the well-studied *xplA* gene has been linked with nitro group removal and ring cleavage by the genera *Rhodococcus, Gordonia, Williamsia* and *Microbacterium* (16-19). Under anaerobic conditions, the enzyme encoded by *diaA* from *Clostridium kluyveri* initiated RDX transformation through nitro group denitration (20, 21). Finally *xenA* and *xenB*, associated with the genus *Pseudomonas*, encode enzymes that primarily transform RDX to methylenedinitramine (22). Monitoring of these genes has previously been deployed at contaminated sites as evidence for RDX degradation (8, 9, 23).

Current approaches to detect RDX degraders in groundwater have typically focused on PCR (9, 10). Although this approach has a high level of sensitivity, it is often limited by the number of genes that can be targeted. More recently, our group used high throughput qPCR to quantify key RDX degrading genes in groundwater at a contaminated site before and after biostimulation (24). However, the specificity of the qPCR primers was not evaluated and the possibility of false positives could not be ruled out. Building on our previous work, the overall aim of the current study was to quantify and explore the diversity of the genes and bacteria previously associated with RDX at an RDX contaminated site both before and after biostimulation with fructose. Although others have used shotgun sequencing to investigate these functional genes in ovine rumen, the approach, to our knowledge, has yet to be used on groundwater from an RDX contaminated site (25). The specific objectives were 1) to determine the relative abundance of each functional gene, 2) to ascertain the taxonomy of the microorganisms associated with each functional gene, 3) to investigate changes in gene abundance following biostimulation and 4) to ascertain if previously identified RDX degraders were present at the site and if their abundance changed following biostimulation. The approach

has the potential to provide a greater depth of knowledge compared to commonly used methods and represents a promising tool for evaluating biodegradation potential at RDX contaminated sites.

2. Materials and Methods

2.1 Sample Collection and DNA Extraction

Groundwater was collected between 2017 and 2019 from an RDX contaminated site, both before and after biostimulation. Site and remediation details as well as the collection of pre-biostimulation samples were described in a previous study (24). Samples collected in 1-L amber glass bottles were shipped to the laboratory overnight on ice and were stored in the dark at 4 °C prior to DNA extraction. Samples were collected both before biostimulation (injection of fructose amended groundwater) and post biostimulation (~1.5 yr. later). All post biostimulation samples used for the shotgun sequencing analysis were newly collected since the previous study. Approximately 1L of groundwater was flowed through a 47-mm-diameter 0.22-µm filter (GSWG047S6, Millipore) using a vacuum pump and then the filter was put into the PowerBead tube from DNeasy PowerWater kit (Qiagen, Germany). The rest of the DNA extraction followed the manufacturer's protocol. Extractions were performed in triplicate, were eluted in 50 µL and stored at -20 °C for further use.

2.2 High Throughput Sequencing

DNA extracts (Table 3.1) were submitted for library generation and sequencing to the Research Technology Support Facility (RTSF) Genomics Core at Michigan State University (MSU). The libraries were prepared using the Takara SMARTer ThruPLEX DNA-Seq Kit and SMARTer DNA HT Dual Index Kit following manufacturer's recommendations. Completed libraries were quantified using a combination of Qubit dsDNA HS and Agilent 4200 TapeStation HS

DNA1000 assays. The libraries were pooled in equimolar proportions and the pool was quantified using the Kapa Biosystems Illumina Library Quantification qPCR kit. This pool was loaded onto one lane of an Illumina HiSeq 4000 flow cell and sequencing was performed in a 2x150bp paired end format using a HiSeq 4000 300 cycle SBS reagent kit. Base calling was performed by Illumina Real Time Analysis (RTA) v2.7.7 and output of RTA was demultiplexed and converted to FastQ format with Illumina Bcl2fastq v2.19.1.

Table 3. 1. Collection dates, RDX concentrations, locations, DNA concentrations and names of groundwater sampling wells.

Source	Well name	Biostimulation status[a]	Date of collection	RDX concentration (µg/L)	Concentration (ng/µL)	
					Rep 1	Rep 2
Perched aquifer	MW22	-	Not collected	-	-	-
		Post	April 2019	<1 (4/30/18)	35.1	44.9
Shallow aquifer	MW32	Pre	Nov. 2017	8 (9/11/17)	11.8	15.2
		Post	April 2019	<1 (4/30/18)	48.0	48.3
Perched aquifer	MW48	-	Not collected	59 (1/27/18)	-	-
		Post	April 2019	3 (4/28/18)	46.8	46.2
Perched aquifer	MW60R	-	Not collected	133 (1/27/18)	-	-
		Post	April 2019	99 (4/28/18)	18.7	29.3
Shallow aquifer	MW62	Pre	Nov. 2017	23 (9/11/17)	101.0	133.0
		Post	April 2019	<1 (4/30/18)	55.0	54.0
Perched aquifer	MW66	Pre	Jan. 2018	103 (1/27/18)	19.8	18.1
		Post	April 2019	3 (4/28/18)	44.5	48.4
Perched aquifer	MW67	Pre	Jan. 2018	24 (1/27/18)	4.9	5.6
		Post	April 2019	4 (4/28/18)	40.7	33.5

[a]All pre-bioaugmentation samples were collected in a previous study(24)

2.3 Taxonomic Analysis

Taxonomic analysis of the metagenomes was achieved using the Meta Genome Rapid Annotation using Subsystem Technology (MG-RAST) (26) (Version 4.0.3.). The processing pipeline involved merging paired end reads, trimming low-quality regions with SolexaQA (27) and removing the artificial duplicate reads with dereplication. Gene calling was performed using FragGeneScan (28). Default setting (best hit classification, 10^{-5} e-value, 60% identity, and a minimal alignment length of 15 amino acids) with the databases ReqSeq (29) and KEGG Orthologs (KO) (30) were used for taxonomic and functional gene profiling. MG-RAST ID numbers and sequencing data have been summarized (Supplementary Table 3.1) and the datasets

are publicly available on MG-RAST. The MG-RAST data files were downloaded and analyzed in Microsoft Excel 2016 to generate the most abundant phylotypes in each sample.

2.4 Functional Gene Analysis

Reference sequences for the functional genes relevant to RDX degradation (*diaA* (21), *nfsI* (12), *pnrB* (31), *xenA* (22), *xenB* (22) and *xplA* (16, 19)) were collected from FunGene (32) using a minimum HMM coverage of 70% (Supplementary Table 3.2). FunGene filters (Supplementary Table 3.2) were set for collecting reference sequences with no less than 60% identity to the consensus sequence of that gene. Unaligned protein sequences were downloaded and dereplicated by the function Clustering.jar derep developed by Ribosomal Database Project (https://github.com/rdpstaff/RDPTools). Dereplicated reference sequences were used to create the database in DIAMOND (double index alignment of next-generation sequencing data) (33), which was the alignment tool for all of functional genes. Before alignment, low quality sequences and Illumina adapters were removed using Trimmomatic (34) (Version 0.36) with the Paired End Mode settings, as described in the Trimmomatic manual (http://www.usadellab.org/cms/uploads/supplementary/Trimmomatic/TrimmomaticManual_V0.32.pdf). As stated above, the processed datasets were then aligned to the dereplicated references using DIAMOND. Only reads that exhibited an identity of ≥ 60 % and an alignment length ≥ 49 amino acids to the reference sequences were retained as aligned reads to each sequence. For each, relative abundance values were calculated using the number of aligned reads divided by the total number of sequences for each sample. The relative abundance values were then normalized by (divided by) the number of dereplicated reference sequences for each gene to produce normalized relative abundance values.

2.5 Functional Gene Phylogenetic Trees

Trees were generated using the 20 most abundant sequences for each target gene (averaged across all samples). For this, text files with the appropriate accession numbers were uploaded to COBALT: constraint-based alignment tool for multiple protein sequences (35). The downloaded alignments (fasta plus gaps) from COBALT were submitted for MAFFT (multiple alignment using fast Fourier transform) alignment using an online server (36) (Version 7). The website was also used to generate trees (by the Neighbor-Joining method) and the trees were exported in Newick format. The downloaded tree files were then uploaded to Interactive Tree of Life (37) (Version 5.5.1). Sequences were colored by phylum (and in some classes, by class) and relative abundance values were added using the Datasets function called simple bar chart.

Additionally, the total relative abundance values of all aligned reference sequences for each gene were summed across all samples to generate a pie chart at the most appropriate classification level for each dataset (phylum, order, genus). The taxonomic information for all aligned reference sequences was obtained using the R package taxonomizr (38), RStudio (39) (Version 0.9.24) and R (40) (Version 4.0.2).

2.6 Co-occurrence Network of Genera

Genera found by MG-RAST with at least 0.1% average relative abundance and 50% occurrence in post biostimulated wells were selected for building the correlation network. The correlation of the filtered genera was calculated with R packages Hmisc (41) (Version 4.4-1) and Matrix (42) (Version 1.2-18) using Spearman ranking correlation. Strong correlations (correlation coefficient ≥ 0.85) and Benjamini-Hochberg method adjusted p value ($p < 0.01$) were set to filter the correlation results. The filtered correlation results were used to build occurrence network with the R package igraph (43) (Version 1.2.6) and this was then visualized in Gephi (44). Both the

identified genera associated with RDX degradation and the genera previously associated with the functional genes generation were marked with different colors.

2.7 Analysis for Species Associated with RDX Biodegradation

Shotgun sequences processed by Trimmomatic were merged with fastq-join (45) (Version 1.3.1). Previously identified RDX degraders were searched for in the National Center for Biotechnology Information (NCBI) taxonomy browser to find their lowest ranks (primarily rank of species) and taxonomy IDs (Supplementary Table 3.3). The merged reads were then aligned to the NCBI protein database (nr) with the taxids option in DIAMOND (33) (version 2.0.6). The searched results were restricted to maximum of 10 sequences with identity ≥ 85 % and query coverage ≥ 85 % and the resulting files were then imported into Megan (46) (community edition Version 6.19.7) for species taxonomy assignment. The phylogenetic tree of those species and normalized counts of reads were then visualized in Interactive Tree of Life (37) (Version 5.5.1).

2.8 Statistical Analysis

The software, Statistical Analysis of Taxonomic and Functional Profiles, (STAMP) (47) (Version 2.1.3) was used to analyze both the metagenomic data (from MG-RAST) and the output of the RDX degrading species comparison across all wells from Megan. Specifically, STAMP was used to detect differences in the relative proportions of the taxonomic and functional profiles between various samples. STAMP analysis included Welch's two-sided t-test for two groups (samples and live controls) ($p < 0.05$) to generate extended error bar figures. The parameters for the generation of these figures are listed (Supplementary Table 3.4). It was also used to generate heatmaps for the most abundant genera (relative abundance > 1.5%) and functional genes. Principle component analysis (PCA) based on genus was also completed in STAMP.

2.9 High Throughput Quantitative PCR

The SmartChip Real Time PCR system was used to quantify the functional genes associated with RDX degradation in the DNA extracts using 12 primer sets developed previously by our group (24) in a 12 assay X 384 sample configuration. A subset of 12 primers were selected based our previous research on the assays' combined theoretical coverage, their performance on the SmartChip and the need for only one plasmid per gene (24). Standards involved 10-fold serial plasmid dilutions (10^1-10^7 copies/reaction) with plasmids described in the same previous work. Gene copy numbers for the plasmids were calculated following the work of Ritalahti et al (48). Primers and plasmids were manufactured by Integrated DNA Technologies (IDT, Coralville, IA) and GenScript Biotech (Piscataway, NJ). Samples and assays were dispensed into a 72 × 72 nanowell chip with the Multisample Nanodispenser. On the chip, the final 100 nL individual reactions consisted of 1× LightCycler 480 SYBR Green I Master (Roche Applied Sciences, Indianapolis, IN), 0.5 μM each of the forward and reverse primers, DNA, and balance PCR grade water. Plasmid dilution series for each gene were in triplicate to generate the standard curves. DNA extracts were run on the chip primarily in triplicates or duplicates and negative controls contained water. Reactions that did not amplify or were identified as false positives were considered as missing data for all analyses. Reactions with a Ct value higher than 28 were assigned as false positives, as recommended for the SmartChip System (49). The gene copy number per milliliter or per gram of starting material was transformed from the gene copy number per reaction calculated based on the standard curve for the corresponding primer set and plasmid. Heatmaps were generated with $\log_{10}$ gene copy number per milliliter or gram using R package ggplot2 (50) (Version 3.3.2).

3. Results

3.1 RDX Concentrations

RDX concentrations before and after biostimulation indicated the approach was successful for reducing contaminant concentrations. For example, a reduction from 103 µg/L to 3 µg/L was observed for MW66. The majority of the other wells (except MW60R) illustrated RDX reductions to 4 µg/L or below.

3.2 Microbial Community Analysis based on MG-RAST

The overall similarity of microbial profiling of samples were tested with PCA based on the genus level results from MG-RAST (Supplementary Figure 3.1). The samples from different wells with similar microbial profiling were clustered together. Wells MW32 and 62 post clustered together (red circle). MW32 pre, MW66 pre and MW67 post clustered (green circle) as did MW62 pre, MW67 pre, MW22 post and MW60R (blue circle). The pre and post well samples clustered separately, indicating the microbial communities changed following biostimulation.

The main phylotypes from the microbial profiles were determined for all samples at the class, order, family and genus levels (Supplementary Figure 3.2). The communities were primarily composed of *Beta-*, *Alpha-*, *Gamma-*, *Delta-* and *Epsilon Proteobacteria* (total average abundance > 38%, 11%, 9% and 5%) (Supplementary Figure 3.2A). However, in MW48, the class *Clostridia* (abundance >14 %) was also important. At order level, the majority of sequences classified as *Burkholderiales* (average abundance ~ 31.7%) (*Beta Proteobacteria*), *Rhizobiales* (average abundance ~ 5.5%) (*Alpha Proteobacteria*), *Desulfuromonadales* (average abundance ~ 4.6%) (*Delta Proteobacteria*), *Bacteroidales* (average abundance ~ 3.7%) (*Bacteroidia*), *Pseudomonadales* (average abundance ~ 3.7%) (*Gamma Proteobacteria*), *Actinomycetales* (average abundance ~ 3.7%) (*Actinobacteria*), *Rhodocyclales* (average abundance ~ 3.5%)

(*Alpha Proteobacteria*) and *Clostridiales* (average abundance ~ 3.0%) (*Clostridia*) (Supplementary Figure 3.2B). *Campylobacterales* (abundance >22%) (*Epsilon Proteobacteria*) in MW32 and MW62 and *Sphingomonadales* (average abundance > 5%) (*Alpha Proteobacteria*) in MW62 and MW67 were also particularly abundant. At the family level, phylotypes primarily classified within the *Comamonadaceae* (average abundance ~ 18.3%) (*Burkholderiales*), *Burkholderiaceae* (average abundance ~ 6.5%) (*Burkholderiales*), *Oxalobacteraceae* (average abundance ~ 3.8%) (*Burkholderiales*), *Rhodocyclaceae* (average abundance ~ 3.5%) (*Rhodocyclales*), *Geobacteraceae* (average abundance ~ 3.5%) (*Desulfuromonadales*) and *Pseudomonadaceae* (average abundance ~ 3.3%) (*Pseudomonadales*), unclassified *Burkholderiales* (average abundance ~ 2.2%) and *Bradyrhizobiaceae* (average abundance ~ 2.1%) (*Rhizobiales*) (Supplementary Figure 3.2C). At the genus level, the dominant genera were *Polaromonas* (average abundance ~ 5.0%) (*Comamonadaceae*), *Acidovorax* (average abundance ~ 5.0%) (*Comamonadaceae*), *Albidiferax* (average abundance ~ 3.9%) (*Comamonadaceae*), *Geobacter* (average abundance ~ 3.5%) (*Geobacteraceae*), *Arcobacter* (average abundance ~ 3.5%) (*Campylobacteraceae*), *Pseudomonas* (average abundance ~ 3.0%) (*Pseudomonadaceae*) and *Burkholderia* (average abundance ~ 2.8%) (*Burkholderiaceae*) (Supplementary Figure 3.2D).

The 20 most abundant genera were characterized for each well. Eleven genera were present in at least 6 wells, including: *Polaromonas*, *Acidovorax*, *Albidiferax*, *Geobacter*, *Pseudomonas*, *Burkholderia*, *Bacteroides*, *Cupriavidus*, *Dechloromonas*, *Variovorax* and *Leptothrix* (Supplementary Figure 3.3). Six genera were more abundant following biostimulation (*Acrobacter, Geobacter, Bacteroides, Clostridium, Paludibacter* and *Pelobacter*) and 6 were less abundant (*Leptothrix, Variovorax, Methylibium, Cupriavidus, Verminephrobacter,*

Bradyrhizobium and *Caulobacter*) (t-test, $p < 0.05$, targeting >1.5% abundance threshold) (Supplementary Figure 3.4).

The most abundant genera from pre- and post- biostimulation of individual wells were also determined to investigate the changes in relative abundance for individual wells (Supplementary Figure 3.5). In MW32, *Albidiferax, Bacteroides, Sulfuricurvum, Sulfurimonas, Paludibacter, Parabacteroides, Clostridium, Dechloromonas* and unclassified *Campylobacterales* became abundant. In MW62, genera that increased after biostimulation were *Geobacter, Pelobacter Bacteroides, Paludibacter* and *Desulfovibrio*. Only two genera, *Dechloromonas* and *Clostridium* showed significant increases after biostimulation in MW66. Among those wells, *Acidovorax, Poloromonas* and *Variovorax* significantly decreased in post biostimulated samples.

3.3 Functional Genes Associated with RDX Biodegradation

All six functional genes previously linked to RDX biodegradation were detected in the groundwater samples from this site (Supplementary Figure 3.6). Normalized relative abundance values were highest for *xenB,* followed by *xenA,* and both were present in all samples analyzed. Following this, *diaA, nsfI* and *pnrB* all illustrated similar normalized relative abundance values. The normalized relative abundance of *xplA* was low or absent in the majority of samples except for a pre-biostimulation sample (MW32).

The taxonomic classifications, along with relative abundance values (pre- and post-biostimulation), of the most aligned reference sequences (top 20) for each gene are shown with phylogenetic trees (Figure 3.1). The most diverse set of sequences was obtained from *diaA* and the relative abundance of each varied across wells. Sequences classifying with *Clostridium kluyveri* (RDX degrader known to contain *diaA*) were present in the majority of wells (Figure

3.1A). The most abundant (again, top 20) *nsfI* (nitroreductase) sequences were almost exclusively classified with the *Gamma Proteobacteria* (Figure 3.1B). The RDX degrading nitroreductases from *Morganella morganii* strain B2 and *Enterobacter cloacae* strain 96-3 (12) were not found. Three functional genes (*pnrB, xenA* and *xenB*) all classified within the *Alpha, Beta, Delta,* or *Gamma Proteobacteria* (Figure 3.1C, D &E). Previous studies have associated the RDX degrading *pnrB* gene with *Pseudomonas* and *Stenotrophomonas* (31, 51) and although *Stenotrophomonas* was not detected here, the majority of *pnrB* sequences were classified as *Pseudomonas* (Figure 3.1C). The most abundant *pnrB* sequences classified as *Azotobacter*. The relative abundance of *pnrB* from *Azotobacter* was particularly high in post biostimulation wells (MW66 post, MW67 post,

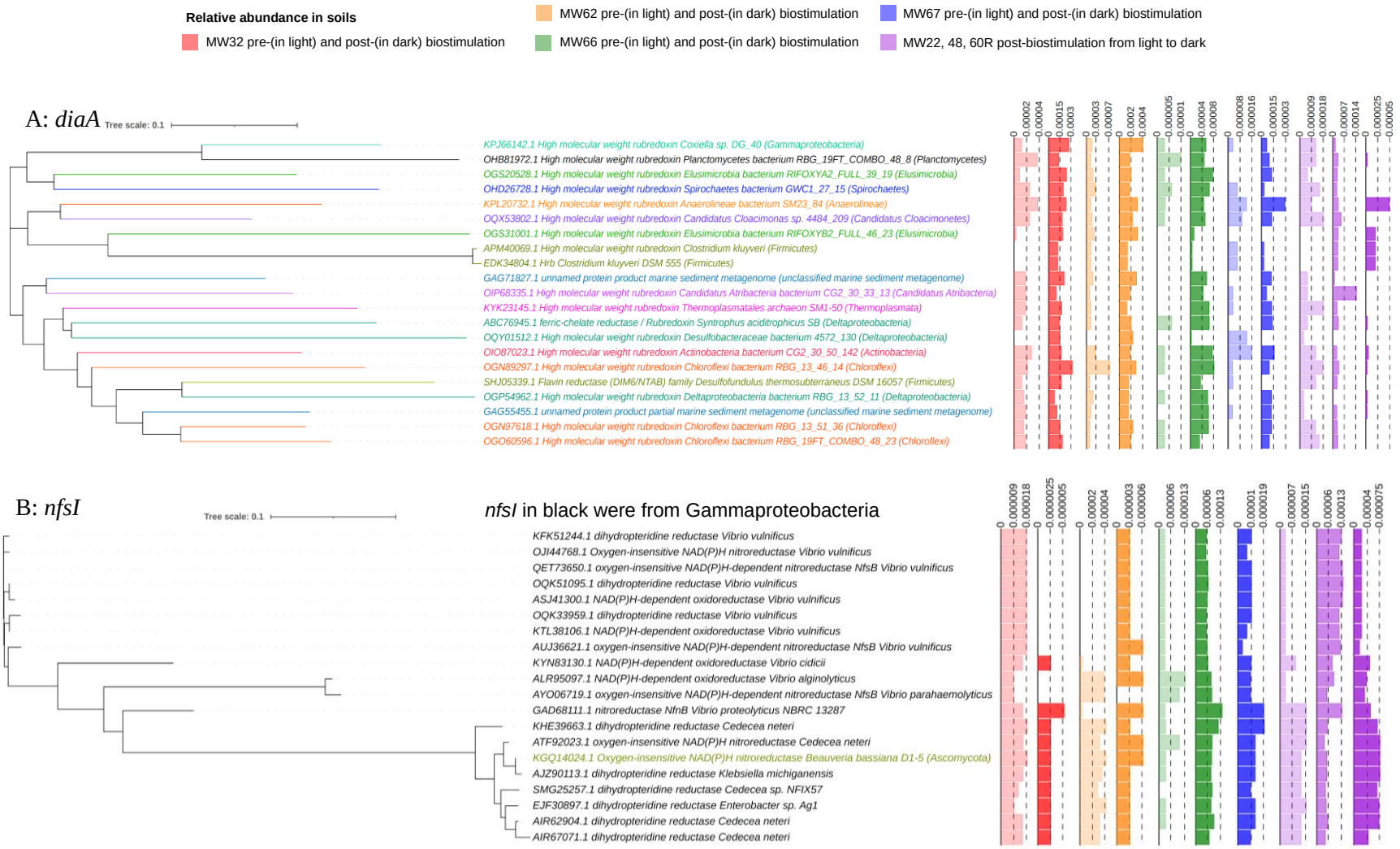
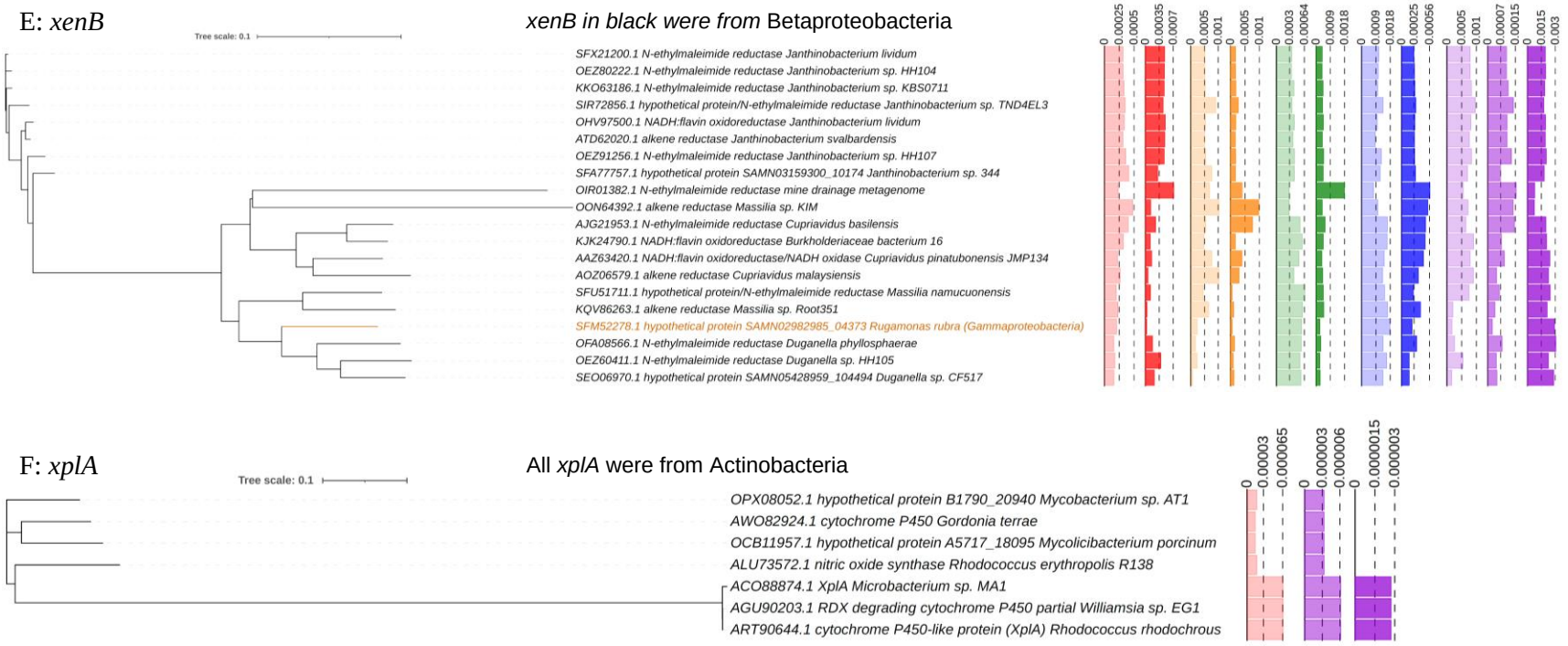


Figure 3. 1. Phylogenetic trees were built for the aligned reference sequences of functional genes (A-F), the reference sequences were colored by phylum or class. The bars on the right illustrated the relative abundance (%) of aligned reference sequences in different samples, light and dark red denoted pre- and post-biostimulation from MW32. light and dark orange denoted pre- and post-biostimulation from MW62, light and dark green denoted pre- and post-biostimulation from MW66, light and dark blue denoted pre- and post-biostimulation from MW67, purple denoted different wells with only post-biostimulated samples.

Figure 3. 1. (continued)

Figure 3. 1. (continued)



MW48 and MW60R). For *xenA* (previously associated with RDX degradation by *Pseudomonas* spp.(22)), the most abundance sequences were classified as *Sorangium cellulosum* (Delta Proteobacteria) (Figure 3.1D) and two sequences were classified within the genus *Pseudomonas*. Other classifications for *xenA* included *Acidovorax, Adenella, Methylobacillus, Methylovorus, Burkholderiales (*Beta Proteobacteria*)* and *Phyllobacterium (*Alpha Proteobacteria*).* For *xenB,* the most abundant sequences were classified within the Beta Proteobacteria with no sequences classifying as *Pseudomonas* (previously linked to RDX degradation by *xenB* (22)) (Figure 3.1E). The well-studied *xplA* gene has been associated with the genera *Rhodococcus, Gordonia, Williamsia* (suborder Corynebacterineae, phylum Actinobacteria) and *Microbacterium* (suborder Micrococcineae, phylum Actinobacteria) (16-19, 52-56). In the present study, sequences classifying with the genera *Gordonia, Rhodococcus, Williamsia* and *Microbacterium* were some of the most abundant *xplA* sequences observed (Figure 3.1E).

The relative abundance of the aerobic (*nfsI*, *pnrB* and *xplA*) and anaerobic (*diaA*, *xenA* and *xenB*) functional genes pre- and post- biostimulation was also investigated (Figure 3.2). The anaerobic genes illustrated greater levels of relative abundance (Figure 3.2B) compared to the aerobic genes (Figure 3.2A). In the four wells with both pre and post biostimulation data, two wells illustrated a dominance in aerobic genes pre biostimulation (MW32 and MW62) and two were more abundant post biostimulation (MW66 and MW67) (Figure 3.2A). For the anaerobic genes, three of the four wells illustrated a greater abundance post biostimulation (Figure 3.2B).

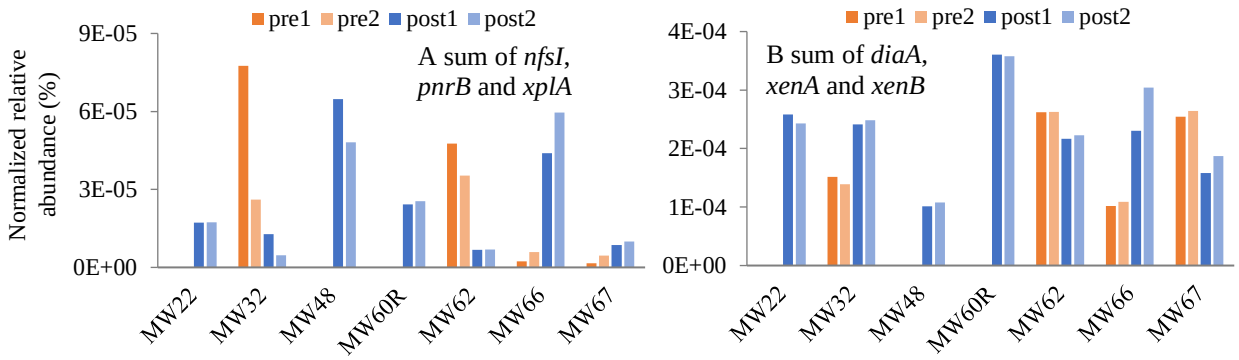


Figure 3. 2. Normalized relative abundance (%) of the total aerobic (*nfsI*, *pnrB* and *xplA*) (A) and anaerobic (*diaA*, *xenA* and *xenB*) (B) functional genes relevant to RDX biodegradation across all monitoring wells (MW) in replicate DNA extracts. The legend terms post and pre refer to the post- and pre-biostimulation samples, respectively.

To better illustrate the diversity of the taxonomic classifications associated with each gene, the taxonomic information of all functional gene reference sequences observed, across all wells, was determined (Supplementary Figure 3.7). The pie-charts generated were classified to the phylum, order or genus level, depending on the data for that gene (to allow a readable number of labels). Similar to the most abundant data for *diaA* (as discussed above) when all sequences were included for *diaA*, a large number of phyla were noted, with the majority classifying as Proteobacteria, Firmicutes and Chloroflexi. The majority of the *nsfI* sequences were classified within the genera *Vibrio, Klebsiella, Cedecea* and *Enterobacter*. For *pnrB*, the majority were classified within the genera *Pseudomonas, Azobacter, Pantoea, Massilia* and *Burkholderia*. The patterns for *xenA* and *xenB* were similar (classified to the order level), with *Burkholderiales, Pseudomonadales, Enterobacterales* and *Rhizobiales* being commonly found. The least amount of diversity was noted for *xplA* with sequences classifying within the genera *Rhodococcus, Williamsia, Gordonia, Mycobacterium, Mycolicibacterium* and *Microbacterium*. The genera associated with those functional genes that had not been identified for RDX degradation were denoted as potential degraders, which were used for following co-occurrence analysis.

3.4 Co-occurrence of Genera Associated with RDX Biodegradation

A co-occurrence network, with strong correlation between each node, was built for better illustrating the relationship of the main genera found with MG-RAST (Figure 3.3). A total of 16 identified genera (colored in red) associated with RDX degradation were found. Based on classifications of those functional genes, potential genera for RDX degradation were also denoted on the network. The majority of the potential genera (28, colored in green) were found for *xenA* or *xenB*, followed by *diaA* (3, colored in orange). Only one genus was shown for *pnrB* and *xplA*. The co-occurrence network was also processed to group those nodes into 7 different modules (Supplementary Figure 3.8). The modules colored in dark green (bottom right corner) and blue grouped *Clostridium* which was identified for generating *diaA*, all three potential genera for *diaA* and several other identified genera. *Rhodococcus* and the potential genera for *xplA* were grouped in purple. The modules colored in light green and orange grouped the rest of identified genera and most potential genera for *xenA* or *xenB*.

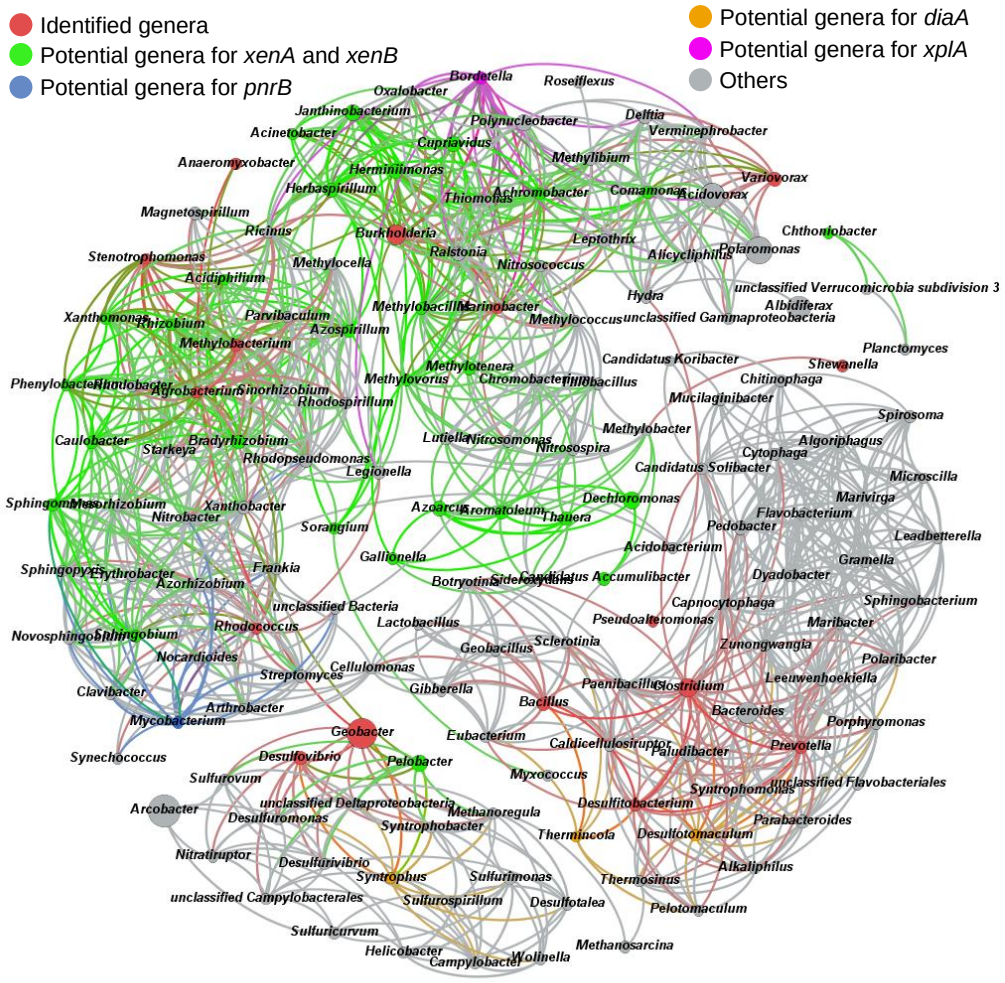


Figure 3. 3. Co-occurrence network based on spearman correlation (rho > 0.85 and p-value < 0.01) of main genera found in all samples from post-biostimulated wells. Only genus with an average abundance > 0.1% and present in at least 50% of samples were considered. Node size indicates the relative abundance (0.1% ~ 5.46%). Nodes colored in red: identified genus associated with RDX degradation. Nodes colored in orange: potential genus to generate *diaA*. Nodes colored in pink: potential genus to generate *pnrB*. Nodes colored in blue: potential genus to generate *xplA*. Nodes colored in green: potential genus to generate *xenA* or *xenB*. No potential genus to generate *nfsI* was found.

3.5 Presence of Known RDX Degraders

Based on the result of MG-RAST, a number of genera associated with RDX degradation were found. The classification to those species within those genera were performed with cutoffs of identity ≥ 85 % and query coverage ≥ 85 %. A phylogram tree for species previously identified as RDX degraders was generated (Figure 3.4). The analysis indicated the presence of 31 RDX degrading bacterial species across all samples. Two fungal species previously associated with RDX degradation were also detected (data not shown). From the 31 bacterial species identified, the genus *Variovorax* showed the highest number of alignments. Others with higher alignments included the genera *Pseudomonas*, *Stenotrophomonas*, *Geobacter* and *Agrobacterium*. The statistical analysis indicated 9 RDX degrading species demonstrated a significant increase in abundance ($p<0.05$) in the post-biostimulation samples compared to the pre-biostimulation samples (Figure 3.5).

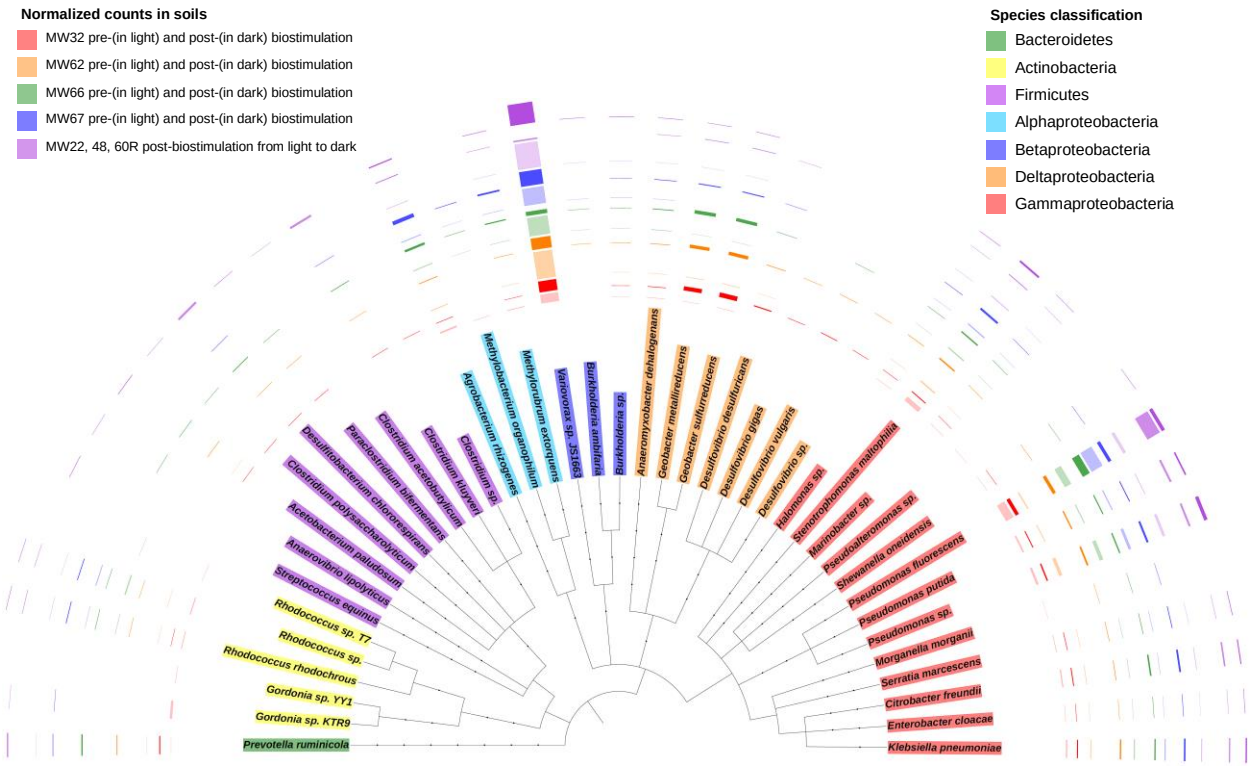


Figure 3. 4. Phylogram constructed with reads assigned (identity ≥ 85% and query coverage ≥ 85%) to the species associated with RDX degradation across all monitoring wells (MW) in replicate DNA extracts. Each species was colored with phylum or class from *Proteobacteria*. The bars in the outside indicated the normalized counts assigned to the species, missing bars meant zero counts.

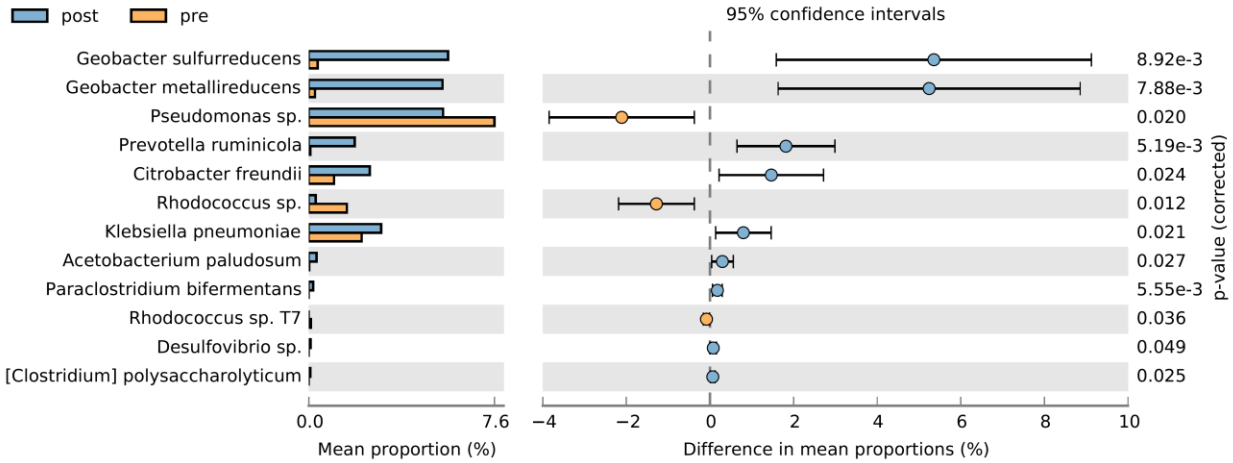


Figure 3. 5. Species associated with RDX degradation showed significant differences before and after biostimulation across all wells. The extended error bar was created using Welch's t-test (two sided) with the default CI option (Welch's inverted), default multiple test correction (no correction) and default *p* value filter of 0.05.

3.6 KEGG Pathways

The changes in the relative abundance of genes associated with xenobiotics biodegradation and metabolism were investigated using MG-RAST and STAMP (Supplementary Figure 3.9). After biostimulation, there was a significant increase for the genes involved in nitrotoluene degradation (Supplementary Figure 3.9 A). At function level, hydrogenase, carboxylesterase, N-ethylmaleimide reductase, 4-carboxymuconolactone decarboxylase and catechol 2,3-dioxygenase significantly were significantly more abundant after biostimulation (Supplementary Figure 3.9 B).

3.7 High Throughput qPCR

All samples were amplified with 16S rRNA specific primers as well as functional gene primers (Supplementary Figure 3.10 A and B). Samples from this study and a previous study (24) were included in the analysis. Three genes (*xenA, xenA* and *nsfI*) were commonly found in both the groundwater and soil samples, indicating their possible widespread occurrence in the environment. In both the groundwater and sediment DNA extracts, *diaA* was not detected. The gene *pnrB* (primer: pnrB_PS5) and *xplA* were only detected in a limited number of groundwater DNA extracts. The maximum copy number was correlated with the relative abundance of that gene by the Spearman's rank correlation test (Supplementary Table 3.5). Two genes (*xenA* and *xplA*) illustrated a significant correlation ($p<0.05$) between the two methods (shotgun sequencing analysis and qPCR).

4. Discussion

RDX concentration changes indicated the biostimulation approach was successful at the majority of groundwater wells. To our knowledge, this is the first investigation of the genes and phylotypes involved in RDX biodegradation using shotgun sequencing in groundwater from an

RDX contaminated site. Although several genes (*diaA, nsfI, xenA* and *pnrB)* exhibited higher relative abundance values in some wells following biostimulation, the differences of relative abundance between the pre- and post- biostimulation were not significant or consistent. The lack of a statistical difference in the current work may be related to the number of samples studied. When these genes were grouped by aerobic or anaerobic conditions, in the shallow aquifer (MW32 and MW62) there was a trend to transfer from aerobic to anaerobic functional genes. While in perched aquifer, although anaerobic genes were dominant, overall aerobic genes were still enriched (MW66), which may share similarity with a slow aerobic RDX degradation under microaerophilic (dissolved oxygen < 0.04 mg/L) condition (57). In a previous study, using environmental samples from two Navy sites, our group found that both *xplA* and *xenA* significantly increased during RDX biodegradation compared to the controls in both groundwater and sediment microcosms (58). Further, in a limited number of microcosms in the previous study, *xenB* gene copy numbers increased.

Comparing the current results to those previously obtained, must be performed with caution, as previous studies have used different detection methods and/or examined different sites. For example, one of the first studies on these functional genes used conventional PCR on groundwater from two sites (Picatinny Arsenal and Pueblo Chemical Depot) where RDX bioremediation (through the addition of organic substrates) was being examined (23). In that case, the researchers did not detect the targeted genes (*xplA*, *xenA xenB*, *onr* and *hydA*) in any of the groundwater samples. They suggested the lack of detection may have resulted from i) the absence of the genes, ii) low gene copy numbers or iii) limitations associated with the primers used. Several studies have targeted a subset of these genes during the evaluation of bioaugmentation for RDX remediation. In 2015, *xenB* and *xplA* were targeted at Umatilla

Chemical Depot (UMCD) in Umatilla, OR, as part of two forced-gradient bacterial transport tests of a mixed culture (strains in the genera *Gordonia, Rhodococcus, Psuedomonas*) or a single culture of *Gordonia* sp. strain KTR9. Through qPCR of *xplA, xenB* and a marker gene (kanamycin resistance gene), the researchers found that the three RDX-degrading strains were effectively introduced and transported within the aquifer (10). Another study at UMCD compared RDX removal rates under bioaugmentation with *Gordonia* sp. strain KTR9 to rates with biostimulation (low or high fructose) and also targeted *xplA* (8). They found that bioaugmentation achieved RDX concentration reductions comparable to those obtained by high carbon biostimulation while requiring substantially less fructose and thus resulting in cost benefits and less secondary water quality impacts.

More recently, the genes associated with RDX biodegradation were investigated at the same site as the current study (Naval Base Kitsap, Bangor Site F near Silverdale, WA) (9, 24). One such project investigated *xplA* and *xenB* during bioaugmentation with *Gordonia* sp. KTR9 and *Pseudomonas fluorescens* strain I-C cells and found that these strains were transported 13 m downgradient over 1 month. The research also demonstrated that bioaugmentation was a viable technology for accelerating RDX cleanup. The other study (by our group), designed new primers for high throughput quantitative PCR to target all six genes (24). The final 49 newly designed primer sets improved upon the theoretical coverage of published primer sets, and this improvement corresponded to more detections in the environmental samples. All genes, except *diaA*, were detected in the site samples, with *xenA* and *xenB* being the most common, agreeing with the results presented here.

Here, a key finding was the detection of a wide range of RDX degraders from numerous genera. Further, the identified genera appeared in the 20 most abundant genera (Supplementary

Figure 3.3) including: *Geobacter*, *Pseudomonas*, *Burkholderia*, *Clostridium*, *Variovorax*, *Desulfovibrio*, *Bacillus*, *Desulfitobacterium*, *Prevotella*, *Rhodococcus*. The potential genera for RDX degradation were dominant in at least 10 samples including: *Acidovorax*, *Cupriavidus*, *Janthinobacterium*, *Dechloromonas*, *Ralstonia*, *Bradyrhizobium* (all associated with *xenA* or *xenB*). Potential genera for generation *pnrB*, including: *Azotobacter* and *Bordetella* were only found in one sample for each as was *Mycobacterium* (associated with *xplA*). With the presence of the functional genes, the species associated with RDX degradation were explored based on relatively strict thresholds (identity ≥ 85 % and query coverage ≥ 85 %) so that the reads aligned very specifically to a taxon, even if the reads aligned to a less specific gene (46).

From the analysis of KEGG pathways in category of xenobiotics biodegradation and metabolism, nitrotoluene degradation was significantly increased after biostimulation. The nitroreductase (KEGG ID K10679) from nitrotoluene degradation was found to be more abundant after biostimulation while it was not statistically significant. Nitroreductase was identified from *Enterobacter cloacae* isolated from a munitions facility because of its ability to metabolize trinitrotoluene (59, 60). This was consistent with the result of relative abundance of *nfsI* (Supplementary Figure 3.6).

To date, previous studies have used other detection methods to explore microbial diversity at RDX contaminated sites. For example, the diversity of RDX degraders was examined at Picatinny Arsenal and Pueblo Chemical Depot using 16S rRNA gene amplicon sequencing (23). In that study, *Rhizobiales* and *Geobacter* were detected in nonbiostimulated samples and *Bacteroidetes* were detected in biostimulated samples. *Pseudomonas* and *Clostridium* were identified in both types of samples. In another project (using clone libraries), microbial communities were examined before and after the addition of acetate or lactate at

Picatinny Arsenal (61). In that project, *Beta Proteobacteria* accounted for more than half of the phylotypes after the addition of substrates. *Alpha Proteobacteria* and *Gamma Proteobacteria* decreased, while *Delta Proteobacteria* increased. *Actinobacteria* and *Firmicutes* were also detected, and *Clostridia* were enriched in samples following lactate addition. A similar pattern of *Beta Proteobacteria* dominance was also observed in groundwater samples at least two months after the addition of acetate at Iowa Army Ammunition Plant (62). Another study examined microbial community changes before and after biostimulation at Los Alamos National Laboratory. At that site, *Rhodococcus* (more than 28%) and *Pseudomonas* (about 6%) were abundant in the indigenous microbial community (63). The abundance of *Pseudomonas* increased to ~ 50% with the addition of safflower oil while that of *Rhodococcus* decreased to less than 5% with the addition of either acetate or safflower oil. In another experiment, waste glycerol (WG) was added to enhance in situ RDX biodegradation (64). *Geobacter*, *Clostridium*, *Klebsiella* and *Bacteroidales*, and *Sulfuricurvum* became enriched in WG impacted monitoring wells.

Previous cost-effective analysis indicated a cost of $79-254 achieved an average RDX transformation rate of 1.20/day in bioaugmentation while a cost of $4 achieved an average rate of 0.49/day in bioaugmentation (8). This low cost for biostimulation (20-63 times lower) will take only 2.5 times longer than bioaugmentation to remove the same amount of RDX. With both diverse functional genes and degraders detected in the indigenous microbial community, it has been suggested that biostimulation is a reasonable and effective alternative to bioaugmentation when cost is a major concern.

In summary, the functional genes and species associated with RDX were both detected in pre- and post-biostimulated samples. However, although sequences aligning with known RDX

degraders were present, it is unclear if these microorganisms were involved in RDX biodegradation at this site. The approach highlighted the importance of *xenA* and *xenB* and demonstrated that a large number of identified and potential RDX degraders were present both pre- and post-bioaugmentation. Further, a subset of these functional gene and degraders was significantly enriched following biostimulation, providing an additional line of evidence for assess the biodegradation potential and evaluating the success of the remediation approach. As the cost of the shotgun sequencing is likely to decrease, in the future, this approach has the potential to be deployed at a larger number of contaminated sites.

Acknowledgements

Thanks to M. M. Michalsen (U.S. Army Engineer Research Development Center) for providing the groundwater. Thanks to Malcolm Gander (Naval Facilities Engineering Command) for partially funding this project and thanks to Craig Tobias (University of Connecticut) for facilitating project funding.

APPENDIX

APPENDIX

Supplementary Table 3. 1. MG-RAST analysis data for datasets from DNA extracts of groundwater samples pre- and post-biostimulation.

Monitoring Well	MG-RAST ID	Upload: bp Count	Upload: Sequences Count	Upload: Mean Sequence Length	Upload: Mean GC percent	Artificial Duplicate Reads: Sequence Count	Post QC: bp Count	Post QC: Sequences Count	Post QC: Mean Sequence Length	Post QC: Mean GC percent
MW22_post1	mgm4886589.3	1,741,804,251 bp	7476242	233 ± 36 bp	53 ± 11 %	1292469	1,429,965,062 bp	6112829	234 ± 36 bp	53 ± 11 %
MW22_post2	mgm4886608.3	2,099,539,542 bp	9109772	230 ± 36 bp	52 ± 10 %	2270230	1,576,080,242 bp	6769037	233 ± 36 bp	52 ± 10 %
MW32_post1	mgm4886604.3	1,954,836,759 bp	8263883	237 ± 35 bp	46 ± 14 %	1328041	1,629,268,603 bp	6860711	237 ± 35 bp	47 ± 14 %
MW32_post2	mgm4886595.3	1,985,444,228 bp	8565024	232 ± 36 bp	47 ± 13 %	1350194	1,661,970,978 bp	7139317	233 ± 36 bp	48 ± 13 %
MW32_pre1	mgm4886601.3	1,661,003,748 bp	7047448	236 ± 35 bp	60 ± 11 %	759357	1,459,056,388 bp	6194085	236 ± 35 bp	60 ± 11 %
MW32_pre2	mgm4886594.3	1,527,083,893 bp	6406285	238 ± 34 bp	61 ± 10 %	749912	1,327,512,629 bp	5572112	238 ± 34 bp	61 ± 10 %
MW48_post1	mgm4886590.3	1,586,090,675 bp	6782319	234 ± 36 bp	43 ± 14 %	1322672	1,268,425,462 bp	5399369	235 ± 36 bp	44 ± 14 %
MW48_post2	mgm4886587.3	1,778,065,022 bp	7522747	236 ± 35 bp	41 ± 12 %	1291252	1,465,710,838 bp	6169619	238 ± 35 bp	41 ± 12 %
MW60R_post1	mgm4886592.3	1,899,735,011 bp	8131466	234 ± 36 bp	62 ± 10 %	1707435	1,496,289,863 bp	6337044	236 ± 35 bp	61 ± 10 %
MW60R_post2	mgm4886588.3	1,653,372,637 bp	7091470	233 ± 36 bp	62 ± 10 %	1291110	1,341,117,043 bp	5707179	235 ± 36 bp	61 ± 10 %
MW62_post1	mgm4886599.3	1,422,045,219 bp	5955796	239 ± 34 bp	49 ± 16 %	920874	1,192,591,889 bp	4975023	240 ± 34 bp	49 ± 16 %
MW62_post2	mgm4886607.3	1,534,652,060 bp	6486643	237 ± 35 bp	51 ± 16 %	991426	1,289,201,671 bp	5427940	238 ± 35 bp	51 ± 16 %
MW62_pre1	mgm4886598.3	1,189,280,134 bp	4946375	240 ± 34 bp	59 ± 9 %	612288	1,030,173,642 bp	4281608	241 ± 34 bp	59 ± 9 %
MW62_pre2	mgm4886605.3	1,101,957,157 bp	4558998	242 ± 33 bp	60 ± 8 %	546177	956,669,806 bp	3955392	242 ± 33 bp	60 ± 9 %
MW66_post1	mgm4886593.3	1,614,662,781 bp	6850124	236 ± 35 bp	54 ± 12 %	788212	1,411,464,163 bp	5985586	236 ± 35 bp	54 ± 12 %
xMW66_post2	mgm4886596.3	1,587,569,892 bp	6796355	234 ± 36 bp	52 ± 13 %	747488	1,397,130,817 bp	5981263	234 ± 36 bp	52 ± 13 %
MW66_pre1	mgm4886597.3	1,503,207,116 bp	6258351	240 ± 34 bp	58 ± 10 %	663943	1,323,804,007 bp	5513662	240 ± 34 bp	58 ± 10 %
MW66_pre2	mgm4886600.3	1,308,725,413 bp	5468213	239 ± 34 bp	58 ± 10 %	583529	1,153,919,440 bp	4823230	239 ± 34 bp	58 ± 10 %
MW67_post1	mgm4886591.3	1,322,169,825 bp	5542737	239 ± 34 bp	54 ± 13 %	716539	1,137,449,146 bp	4764623	239 ± 34 bp	54 ± 13 %
MW67_post2	mgm4886606.3	1,512,555,910 bp	6371805	237 ± 35 bp	55 ± 12 %	835319	1,298,629,352 bp	5465345	238 ± 35 bp	55 ± 12 %
MW67_pre1	mgm4886602.3	1,517,750,130 bp	6345506	239 ± 34 bp	60 ± 8 %	764462	1,320,419,861 bp	5517949	239 ± 34 bp	60 ± 8 %
MW67_pre2	mgm4886603.3	1,392,698,023 bp	5787985	241 ± 34 bp	61 ± 8 %	696637	1,211,531,946 bp	5029835	241 ± 33 bp	60 ± 8 %

Supplementary Table 3. 2. FunGene filters for obtaining the reference sequences and the number of collected sequences before and after dereplication.

Gene	Webpage name	Minimum HMM coverage	Minimum score	Number of sequences collected	Number of sequences after dereplication
diaA	diaA_new	70%	350	116	90
nfsI	nfsI	70%	302	15742	653
pnrB	pnrB	70%	309	275	131
xenA	xenA	70%	497	3085	957
xenB	xenB	70%	475	6336	1371
xplA	xplA	70%	1000	11	7

Supplementary Table 3. 3. Identified RDX degraders with the lowest rank name and taxonomy ID from NCBI.

Strain or species for RDX or its metabolites degradtion	reference	Lowest rank name in NCBI	NCBI Rank	NCBI taxonomy ID	Number of subtrees [D]
Acetobacterium malicum Strain HAAP-1	(65)	*Acetobacterium malicum*	species	52692	0
Acetobacterium paludosum	(66)	*Acetobacterium paludosum*	species	52693	0
Acremonium sp. HAW-OCF3	(67)	*Acremonium sp. HAW-OCF3*	species	311340	0
Anaeromyxobacter dehalogenans Strain K (no ATCC number)	(68)	*Anaeromyxobacter dehalogenans*	species	161493	2
Anaerovibrio lipolyticus	(69)	*Anaerovibrio lipolyticus*	species	82374	3
Bacillus (HPB2)	(70)	*Bacillus sp. HPB-2*	species	259962	0
Bacillus (HPB3)	(70)	*Bacillus sp. HPB-3*	species	259965	0
Bullera unica strain HAW-OCF2	(67)	*Bullera unica*	species	57474	0
Burkholderia sp.BL	(71)	*Burkholderia sp.*	species	36773	0 [E]
Citrobacter freundii	(72)	*Citrobacter freundii*	species	546	18
Clostridium acetobutylicum (ATCC 824)	(73)	*Clostridium acetobutylicum ATCC 824*	strain	272562	0
Clostridium bifermentans [A]	(74)	*Paraclostridium bifermentans*	species	1490	6
Clostridium bifermentans strain HAW-1	(75)	*Paraclostridium bifermentans*	species	1490	6
Clostridium geopurificans MJ1T	(76)	*Clostridium geopurificans*	species	558153	0
Clostridium kluyveri ATCC8527	(21)	*Clostridium kluyveri*	species	1534	2
Clostridium polysaccharolyticum	(69)	*Clostridium polysaccharolyticum*	species	29364	1
Clostridium sp. EDB2	(77)	*Clostridium sp. EDB2*	species	261021	1
Clostridium sp. HAW-E3	(75, 78)	*Clostridium sp.*	species	1506	0 [E]
Clostridium sp. HAW-EB17	(79)	*Clostridium sp.*	species	1506	0 [E]
Clostridium sp. HAW-G3	(78)	*Clostridium sp.*	species	1506	0 [E]
Clostridium sp. HAW-G4	(75, 78)	*Clostridium sp.*	species	1506	0 [E]
Clostridium sp. HAW-HC1	(75, 78)	*Clostridium sp.*	species	1506	0 [E]
Desulfitobacterium chlororespirans Strain Co23	(68)	*Desulfitobacterium chlororespirans*	species	51616	1
Desulfovibrio desulfuricans EFX-DES	(80)	*Desulfovibrio desulfuricans*	species	876	3
Desulfovibrio sp. HAW-EB18	(79)	*Desulfovibrio sp.*	species	885	0 [E]
Desulfovibrio sp. HAW-ES2	(78)	*Desulfovibrio sp.*	species	885	0 [E]
Desulfovibrio spp. desulfuricans A	(81)	*Desulfovibrio desulfuricans*	species	876	3
Desulfovibrio spp. desulfuricans B	(81)	*Desulfovibrio desulfuricans*	species	876	3
Desulfovibrio spp. gigas	(81)	*Desulfovibrio gigas*	species	879	1
Desulfovibrio spp. vulgaris	(81)	*Desulfovibrio vulgaris*	species	881	4
Enterobacter cloacae strain 96-3	(12)	*Enterobacter cloacae*	species	550	43

Supplementary Table 3. 3. (continued)

Geobacter metallireducens	(68)	*Geobacter metallireducens*	species	28232	2
Geobacter sulfurreducens	(68)	*Geobacter sulfurreducens*	species	35554	3
Gordonia sp. KTR9	(56)	*Gordonia sp. KTR9*	species	337191	0
Gordonia sp. YY1	(82)	*Gordonia sp. YY1*	species	396712	0
Halomonas (HAW-OC4)	(83)	*Halomonas sp.*	species	1486246	0 [E]
Klebsiella pneumoniae Strain SCZ-1	(84)	*Klebsiella pneumoniae*	species	573	406
Marinobacter (HAW-OC1)	(83)	*Marinobacter sp.*	species	50741	0 [E]
Methylobacterium extorquens	(85)	*Methylorubrum extorquens*	species	408	5
Methylobacterium organophilum	(85)	*Methylobacterium organophilum*	species	410	0
Methylobacterium rhodesianum	(85)	*Methylorubrum rhodesianum*	species	29427	0
Methylobacterium sp. JS178	(86)	*Methylobacterium sp. JS178*	species	316459	0
Methylobacterium sp. strain BJ001	(85)	*Methylorubrum sp.*	species	2282524	0 [E]
Morganella morganii	(72)	*Morganella morganii*	species	582	10
Morganella morganii strain B2	(12)	*Morganella morganii*	species	582	10
Penicillium sp. HAW-OCF5	(67)	*Penicillium sp. HAW-OCF5*	species	311341	0
Phanerochaete chrysosporium	(87)	*Phanerochaete chrysosporium*	species	5306	1
Prevotella ruminicola	(69)	*Prevotella ruminicola*	species	839	2
Providencia rettgeri	(72)	*Providencia rettgeri*	species	587	3
Pseudoalteromonas (HAW-OC2)	(83)	*Pseudoalteromonas sp.*	species	53249	0 [E]
Pseudoalteromonas (HAW-OC5)	(83)	*Pseudoalteromonas sp.*	species	53249	0 [E]
Pseudomonas (HPB1)	(70)	*Pseudomonas sp.*	species	306	0 [E]
Pseudomonas fluorescens I-C	(22)	*Pseudomonas fluorescens*	species	294	47
Pseudomonas putida II-B	(22)	*Pseudomonas putida*	species	303	39
Pseudomonas sp. HK-6	(88)	*Pseudomonas sp. HK-6*	species	342605	0
Rhizobium rhizogenes BL [A]	(71)	*Agrobacterium rhizogenes*	species	359	1
Rhodococcus sp. Strain A	(57)	*Rhodococcus sp.*	species	1831	0
Rhodococcus sp. strain DN22	(89)	*Rhodococcus sp. DN22*	species	357684	0
Rhodococcus species isolate T7	(17)	*Rhodococcus sp. T7*	species	627444	0
Rhodococcus species isolate T9N	(17)	*Rhodococcus sp. T9N*	species	627445	0
Rhodococcus strain YH1	(54)	*Rhodococcus sp. YH1*	species	89066	0
Rhodococcus rhodochrous strain 11Y	(18)	*Rhodococcus rhodochrous*	species	1829	8
Rhodotorula mucilaginosa strain HAW-OCF1	(67)	*Rhodotorula mucilaginosa*	species	5537	0
Serratia marcescens	(90)	*Serratia marcescens*	species	615	27
Shewanella halifaxensis sp. strain HAW-EB4^T	(91)	*Shewanella halifaxensis HAW-EB4*	strain	458817	0
Shewanella oneidensis Strain MR1	(68)	*Shewanella oneidensis MR-1*	strain	211586	0
Shewanella sediminis sp. strain HAW-EB3^T	(92)	*Shewanella sediminis HAW-EB3*	strain	425104	0
Shewanella sp. HAW EB1	(79)	*Shewanella sp.*	species	50422	0 [E]
Shewanella sp. HAW EB2	(79)	*Shewanella sp.*	species	50422	0 [E]
Shewanella sp. HAW-EB5	(79)	*Shewanella atlantica*	species	271099	0
Stenotrophomonas maltophilia OK-5	(31)	*Stenotrophomonas maltophilia*	species	40324	37
Stenotrophomonas maltophilia PB1	(93)	*Stenotrophomonas maltophilia*	species	40324	37

Supplementary Table 3. 3. (continued)

Streptococcus bovis [A]	(69)	*Streptococcus equinus*	species	1335	7
Variovorax sp. Strain JS1663	(94)	*Variovorax sp. JS1663*	species	1851577	0
Williamsia sp. KTR4	(56)	*Williamsia sp. KTR4*	species	337192	0
Desulfosporosinus [B]	(95)	*Desulfosporosinus*	genus	79206	NC
Fusobacteria isolate HAW-EB21 [B]	(75, 79)	*Fusobacteria*	phylum	32066	NC
Aspergillus niger [C]	(96)	*Aspergillus niger*	species	5061	12
Cladosporium cladosporioides [C]	(71)	*Cladosporium cladosporioides*	species	29917	1

A: The names of those identified species in the paper were revised in NCBI to another name.
B: Rank of the two were higher than species.
C: The two species were fungus, due to too many clades, the species names could not be displayed when analyzed in Megan.
D: The number means the identified microorganism within that species, NC means not checked.
E: The identified strain name from the paper could not be searched in NCBI taxonomy browser, for example: *Burkholderia sp.BL was assigned to Burkholderia sp.* which belonged to unclassified *Burkholderia*.

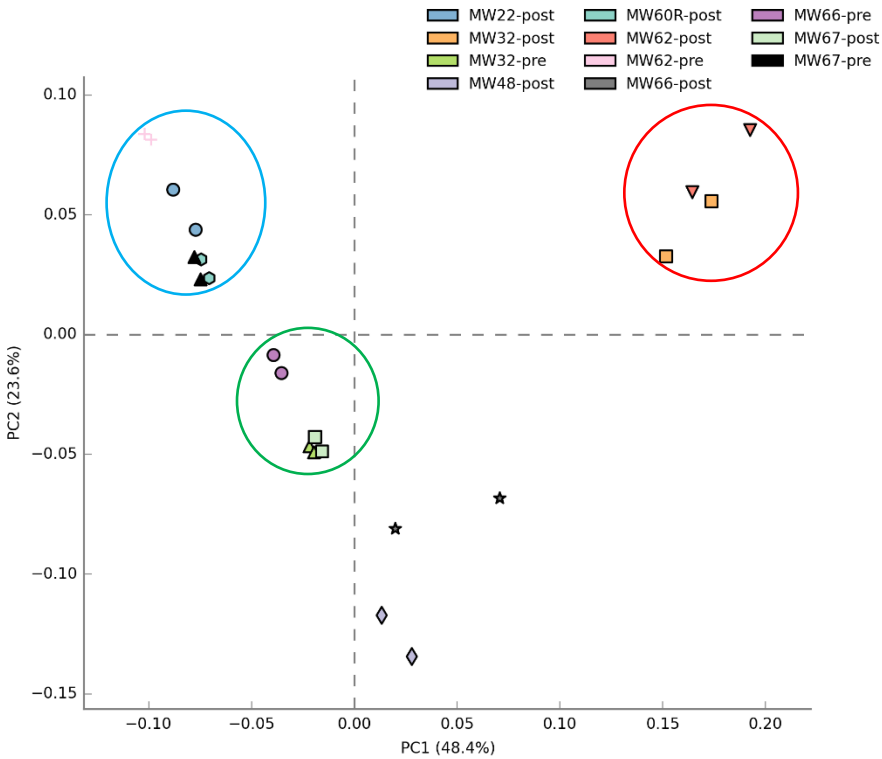
Supplementary Table 3. 4. STAMP analysis parameters for generating Supplementary Figures 3.4 and 3.5. Filters were applied to limit the number of genera or functions shown in each figure. All tests used Welch's t-test (two sided) with the default CI option (Welch's inverted) and default multiple test correction (no correction). For each test, two groups were compared (pre and post biostimulation samples).

A	**Genus MW32**	**Genus MW62**	**Genus MW66**	**Genus MW67**
Parent Level	Entire sample	Entire sample	Entire sample	Entire sample
Profile Level	Genus	Genus	Genus	Genus
Unclassified	Retain unclassified reads	Retain unclassified reads	Retain unclassified reads	Retain unclassified reads
Filtering				
p-value filter >	0.05	0.05	0.05	0.05

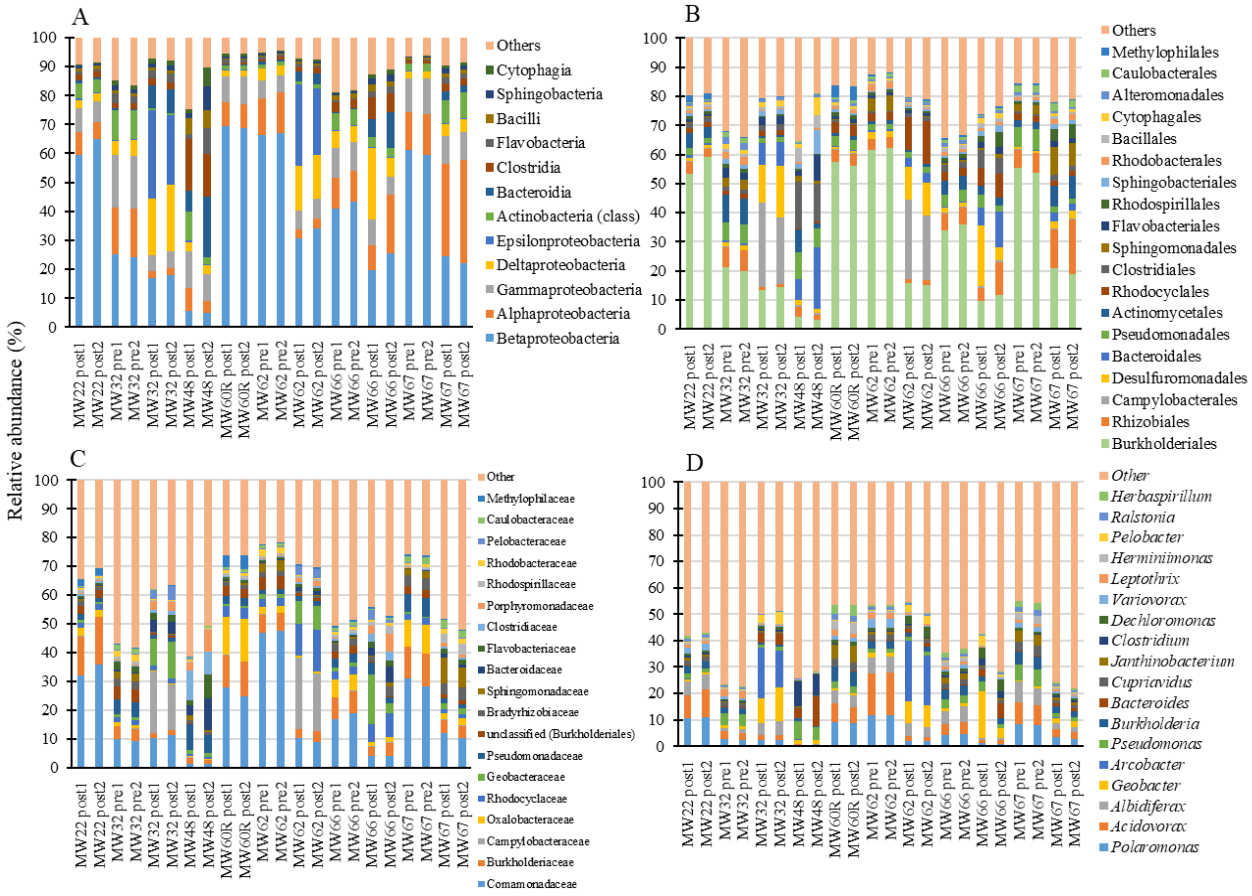
B	**Function 45 most different**
Parent Level	Entire sample
Profile Level	Function
Unclassified	Retain unclassified reads
Filtering	
p-value filter >	0.05
Effect size filter 1 Different between two proportions	
Effect size <	0.09

Supplementary Table 3. 5. Spearman's rank correlation parameters between gene copy number of qPCR and relative abundance of genes associated with RDX degradation.

Gene	S	p value	rho
nfsI	1288.4	0.1826	-0.3296
pnrB	780.89	0.4402	0.19412
xenA	450	0.02197	0.535604
xenB	932.98	0.8836	0.037171
xplA	212.09	0.00013	0.781128

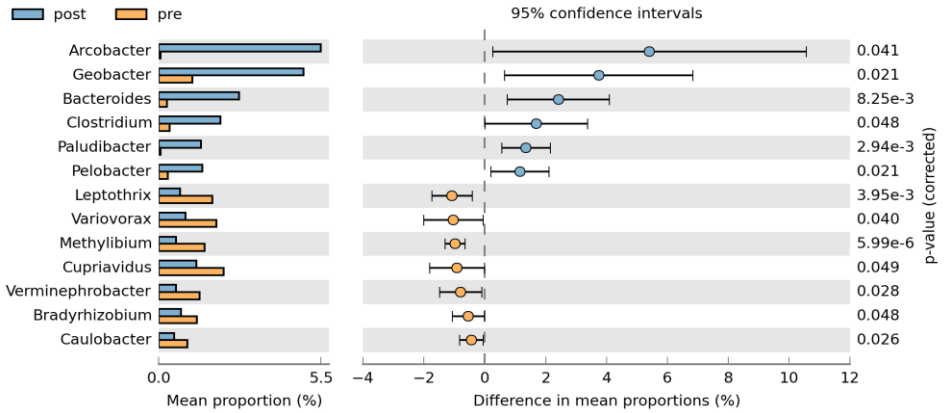


Supplementary Figure 3. 1. Principle component analysis of all samples based on the genus results from MG-RAST. Clustered samples were marked in circles.

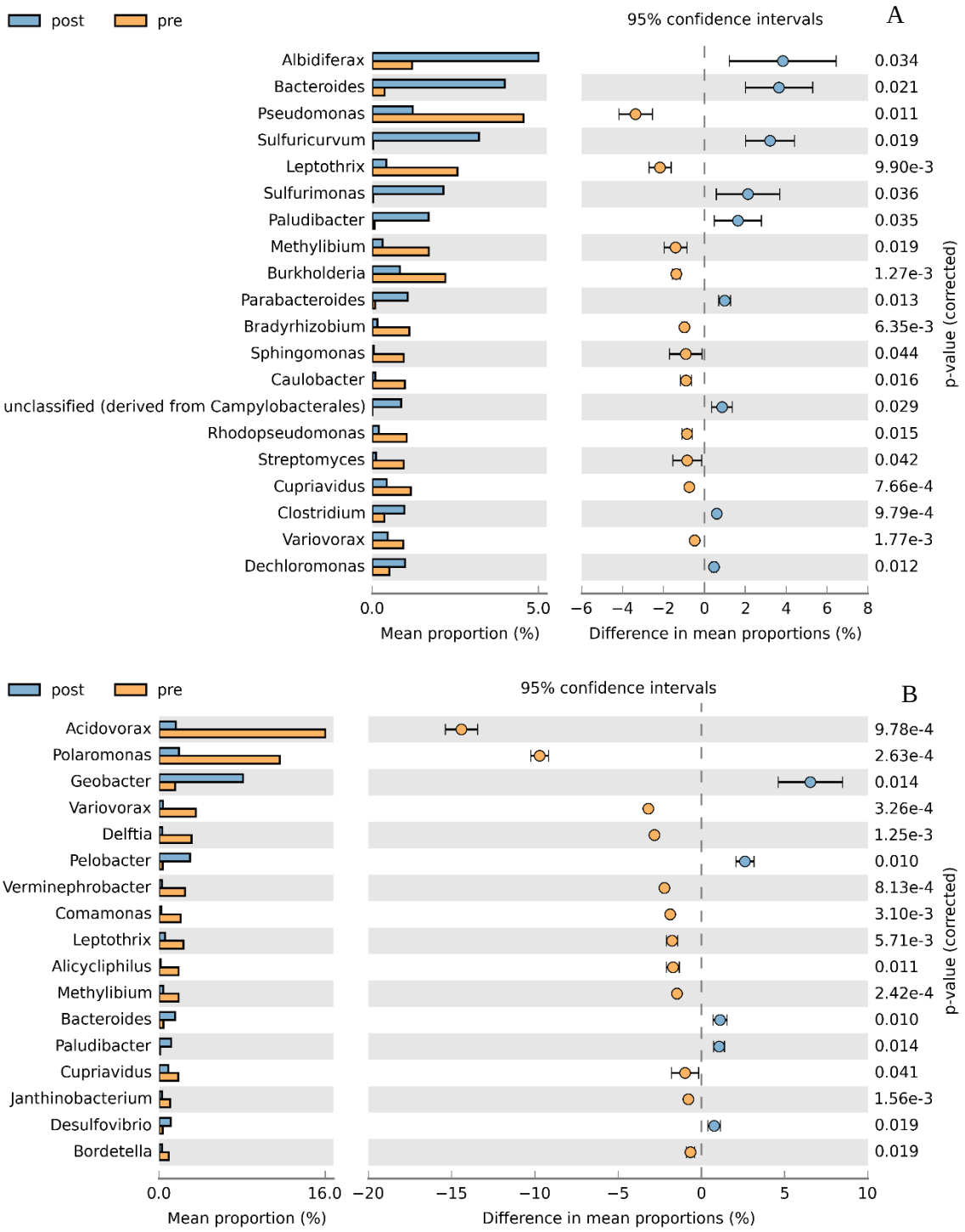


Supplementary Figure 3. 2. The most abundant phylotypes in each sample at the class (A), order (B) family (C), and genus (D) levels. For each classification, phylotypes with an average relative abundance across all samples less than 1% were placed within "other".

Supplementary Figure 3. 3. Relative abundance (%) of the 20 most abundant genera in duplicated samples from each well. Pre and post refer to the pre- and post-biostimulation samples, respectively.

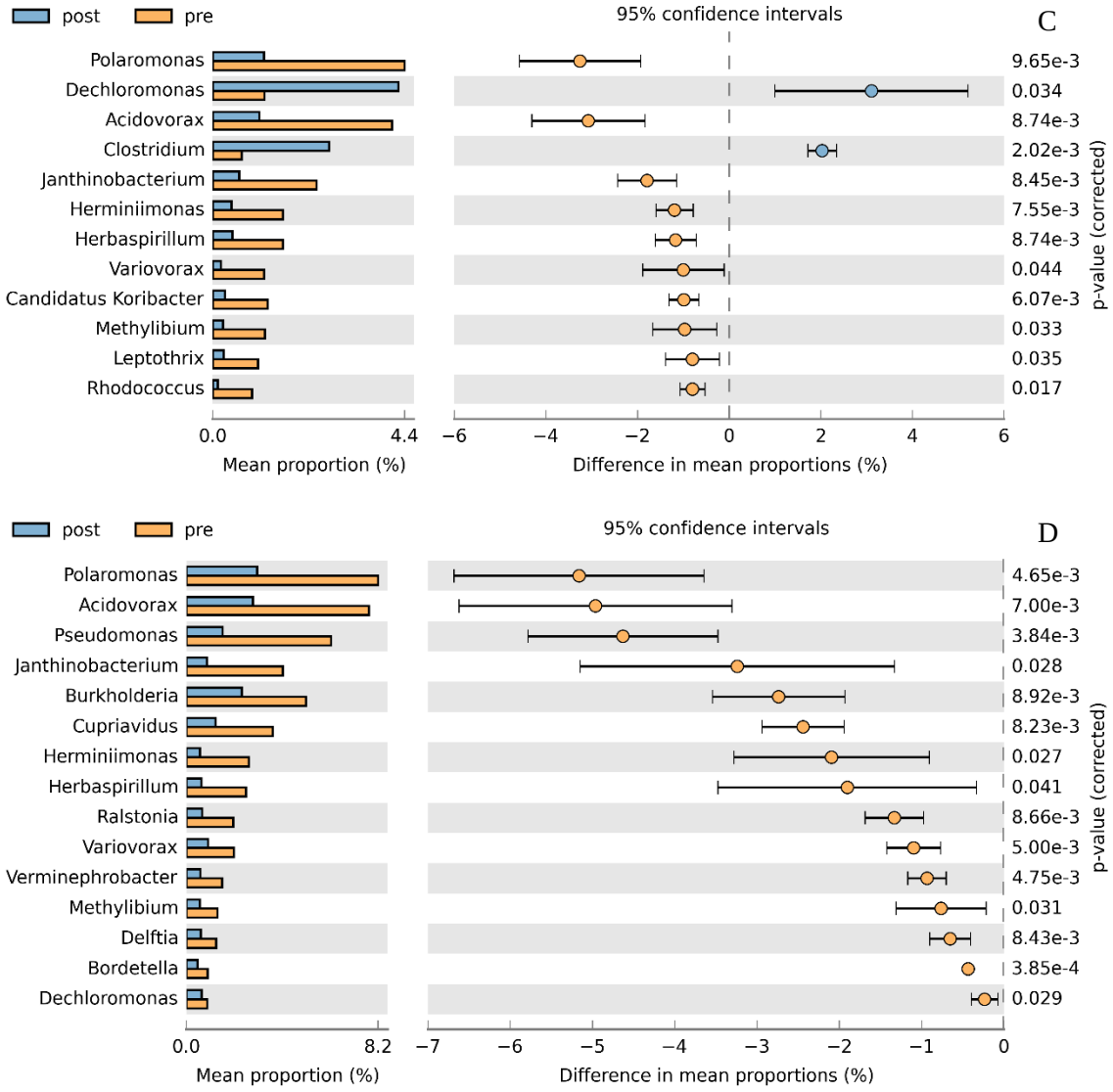


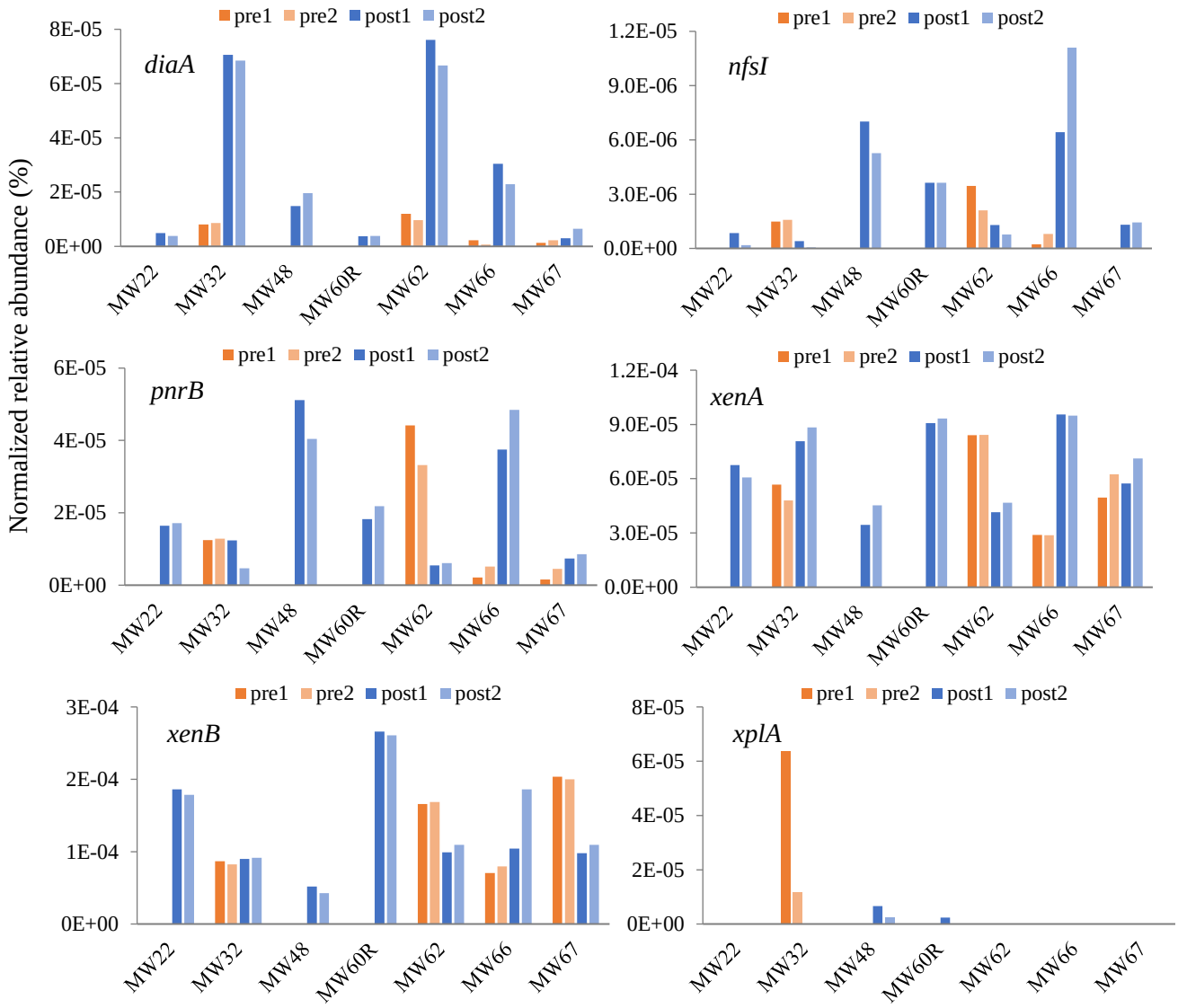
Supplementary Figure 3. 4. A comparison of those significantly different between pre- and post-biostimulation wells from the abundant genera (relative abundance ≥ 1.5) ($p < 0.05$, Welch's two sided t-test).



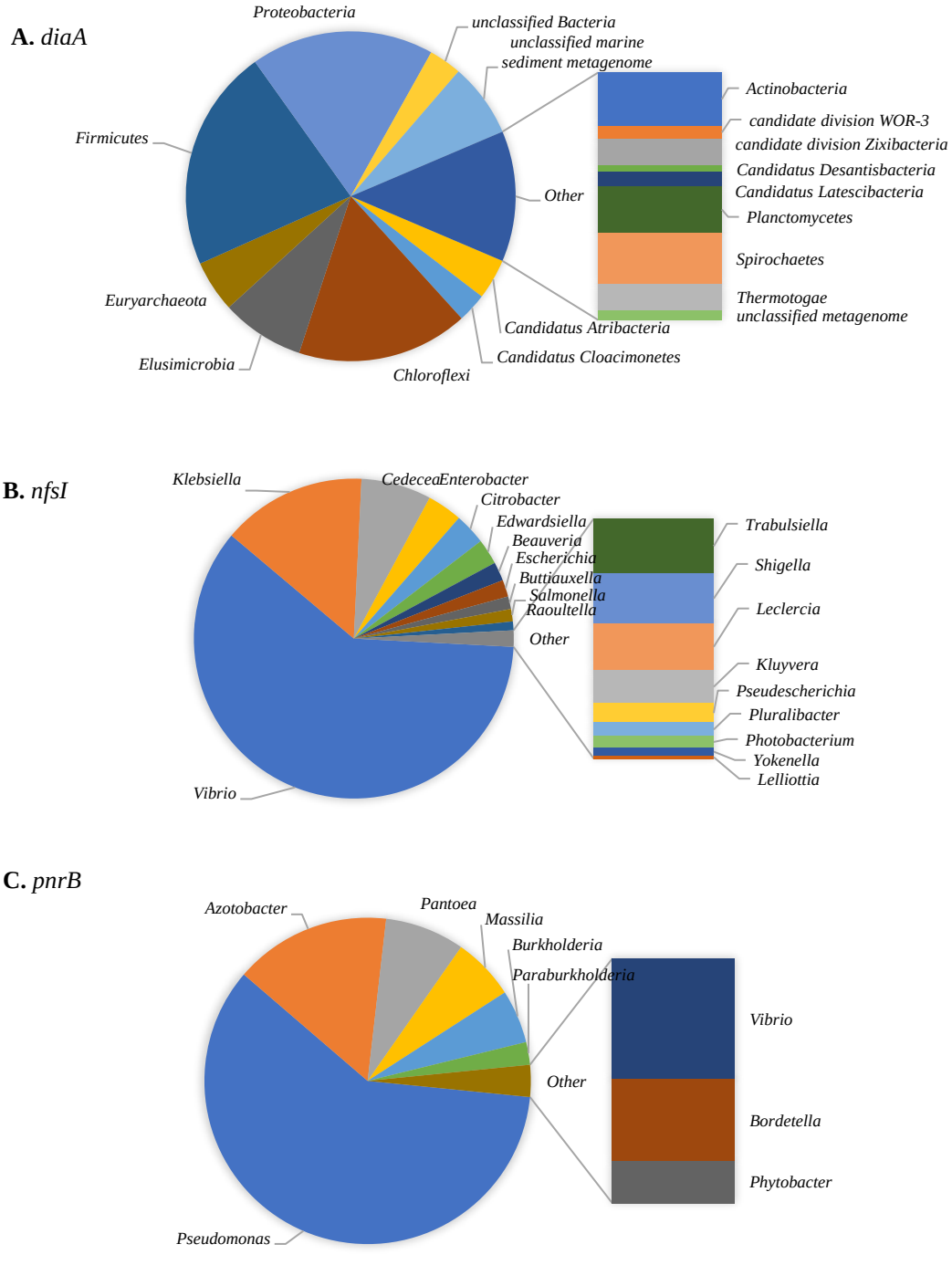
Supplementary Figure 3. 5. Comparison of the most abundant genera pre- and post biostimulation in MW32 (A), MW62 (B), MW66 (C) and MW67 (D) with significant differences ($p < 0.05$, Welch's two sided t-test).

Supplementary Figure 3. 5. (continued)





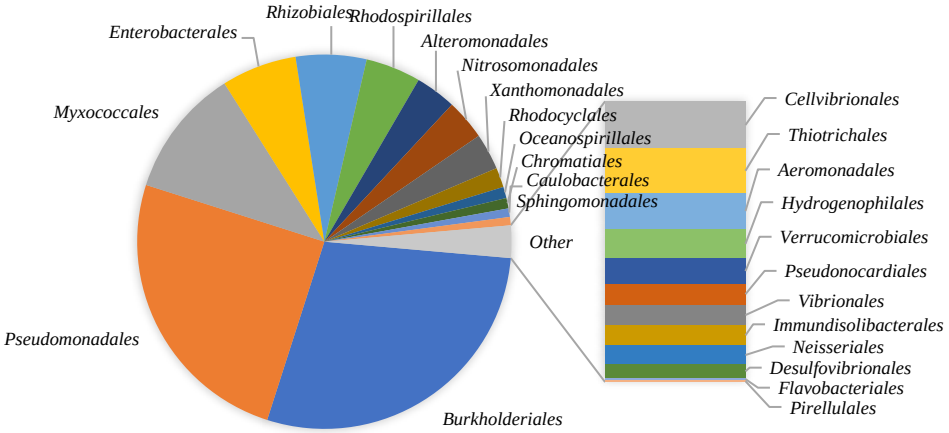
Supplementary Figure 3. 6. Normalized relative abundance (%) of the functional genes relevant to RDX biodegradation across all monitoring wells (MW) in replicate DNA extracts. The legend terms post and pre refer to the post- and pre-biostimulation samples, respectively.



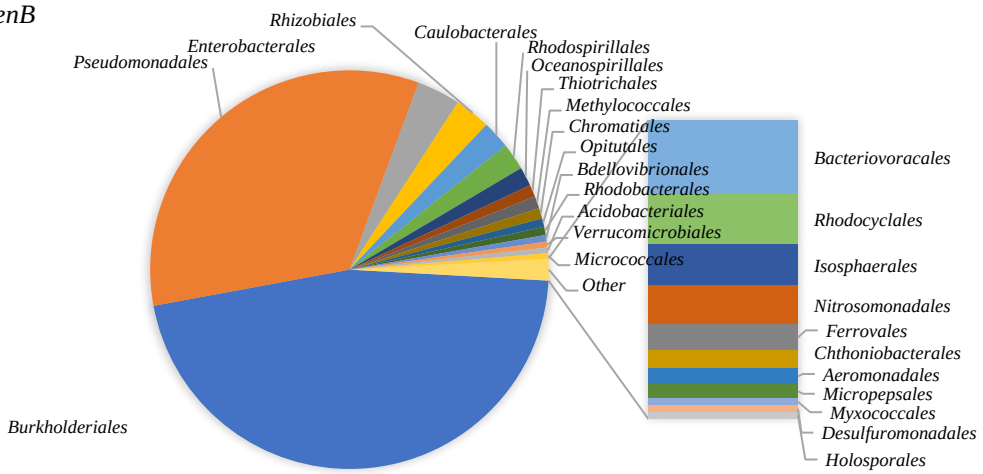
Supplementary Figure 3. 7. Taxonomy of microorganisms associated with aligned references sequences of functional genes: *diaA* (phylum level, A), *nfsI* (genus level, B), *pnrB* (genus level, C), *xenA* (order level, D), *xenB* (order level, E) and *xplA* (genus level, F) sequences across all soils.

Supplementary Figure 3. 7. (continued)

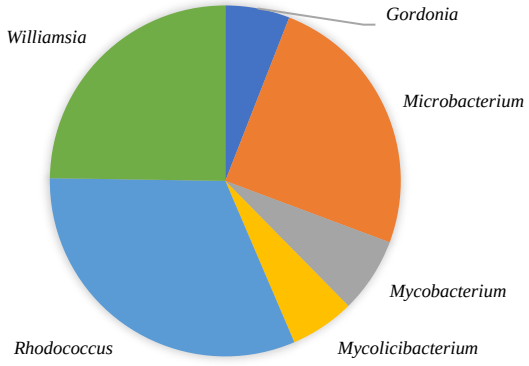
D. *xenA*

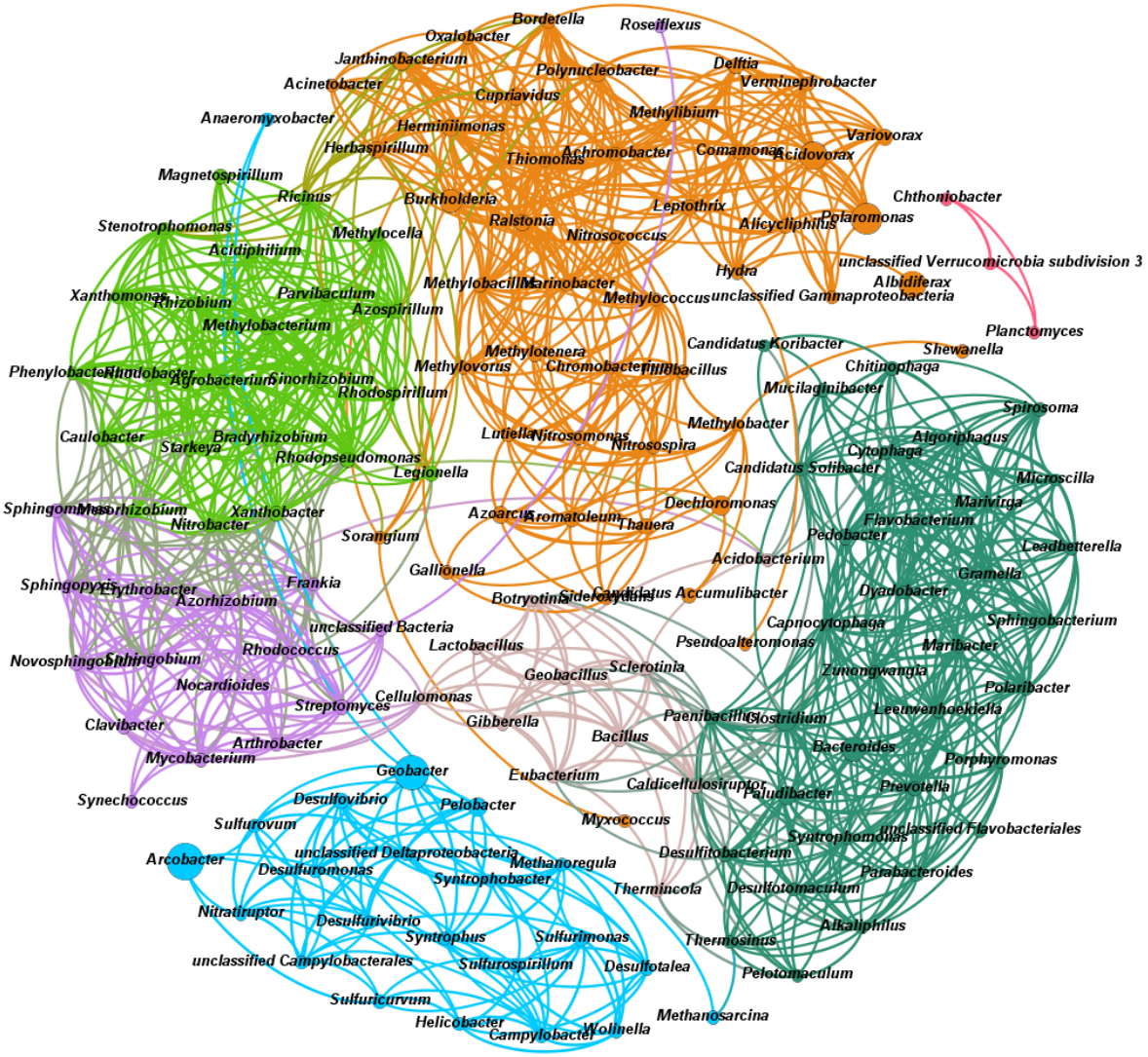


E. *xenB*

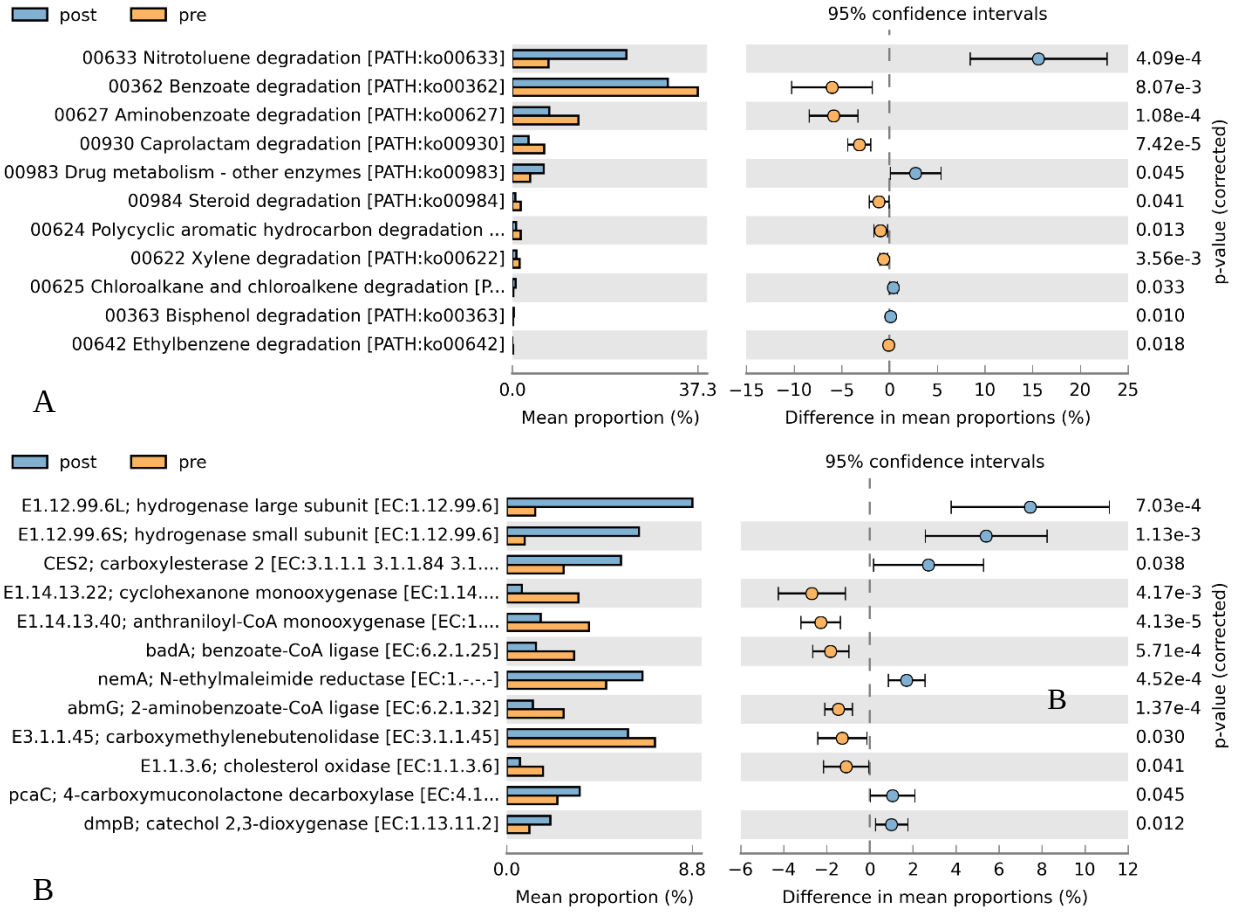


F. *xplA*

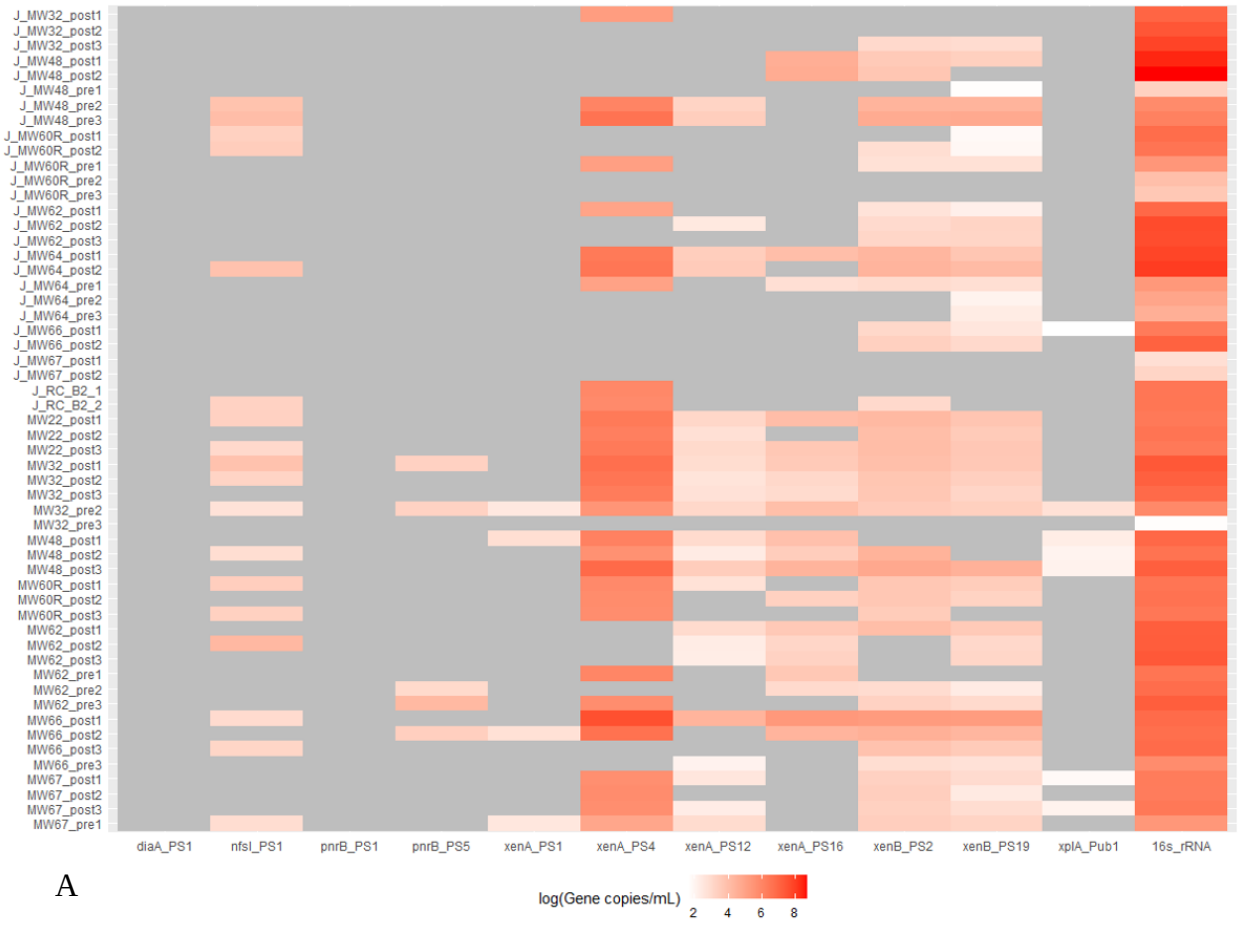




Supplementary Figure 3. 8. Co-occurrence network based on spearman correlation (rho > 0.85 and p-value < 0.01) of main genera found in all samples from post-biostimulated wells. Only genus with an average abundance > 0.1% and present in at least 50% of samples were considered. Node size indicates the relative abundance (0.1% ~ 5.46%). The network was process with Modularity function of Gephi to group nodes colored into 7 different modules with default setting and a resolution of 0.85.



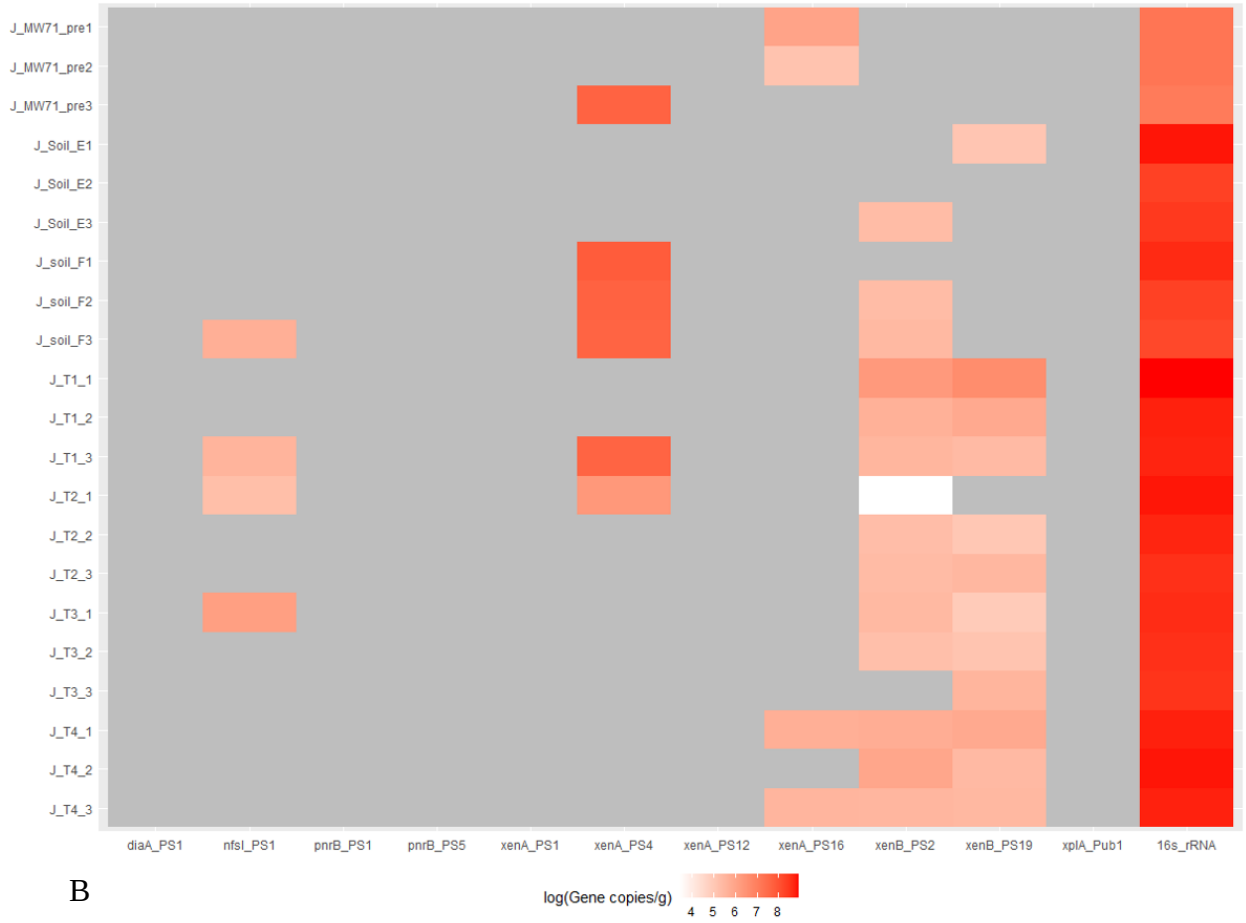
Supplementary Figure 3. 9. Comparison of degradation pathway (A) and functions (B) in category of Xenobiotics biodegradation and metabolism between pre- and post biostimulation wells. For degradation pathway analysis, default options were used for the two groups comparison ($p < 0.05$, Welch's two sided t-test). For functions analysis, an extra filter was added as difference in mean proportions > 1%.



A

Supplementary Figure 3. 10. Heatmap of groundwater and Red Cedar River (RC) $\log_{10}$ gene copies per milliliter (A) and sediment $\log_{10}$ gene copies per gram (B). Grey cells indicate either no amplification or false positive amplification. In the sample name, post and pre refer to the post- and pre-biostimulation samples, J_ refer to the samples from previous work (24).

Supplementary Figure 3. 10. (continued)



B

REFERENCES

REFERENCES

1. EPA. 2017. Technical Fact Sheet – Hexahydro-1,3,5-trinitro1,3,5-triazine (RDX). https://www.epa.gov/sites/production/files/2017-10/documents/ffrro_ecfactsheet_rdx_9-15-17_508.pdf.

2. EPA. 2020. Search Superfund Site Information. https://cumulis.epa.gov/supercpad/cursites/srchsites.cfm.

3. Agency for Toxic Substances and Disease Registry (ATSDR). 2012. Toxicological Profile for RDX.

4. Felt D, Johnson JL, Larson S, Hubbard B, Henry K, Nestler C, Ballard JH. 2013 Evaluation of Treatment Technologies for Wastewater from Insensitive Munitions Production The US Army Engineer Research and Development Center,

5. Wani AH, O'Neal BR, Davis JL, Hansen LD. 2002. Treatability Study for Biologically Active Zone Enhancement (BAZE) for In Situ RDX Degradation in Groundwater. U.S. Army Engineer Research and Development Center, Vicksburg, Mississippi.

6. EPA. 2015. Third Five-Year Review For Cornhusker Army Ammunition Plant Grand Island, Nebraska Environmental Protection Agency Region 7, Lenexa, Kansas.

7. Arcadis. 2004. In-situ enhanced anaerobic explosive treatment field test report, Milan Army Ammunition Plant Milan, Tennessee. Arcadis, Knoxville, Tennessee.

8. Michalsen MM, King AS, Rule RA, Fuller ME, Hatzinger PB, Condee CW, Crocker FH, Indest KJ, Jung CM, Istok JD. 2016. Evaluation of Biostimulation and Bioaugmentation To Stimulate Hexahydro-1,3,5-trinitro-1,3,5,-triazine Degradation in an Aerobic Groundwater Aquifer. Environ Sci Technol 50:7625-32.

9. Michalsen MM, King AS, Istok JD, Crocker FH, Fuller ME, Kucharzyk KH, Gander MJ. 2020. Spatially-distinct redox conditions and degradation rates following field-scale bioaugmentation for RDX-contaminated groundwater remediation. Journal of Hazardous Materials 387:121529.

10. Crocker FH, Indest KJ, Jung CM, Hancock DE, Fuller ME, Hatzinger PB, Vainberg S, Istok JD, Wilson E, Michalsen MM. 2015. Evaluation of microbial transport during aerobic bioaugmentation of an RDX-contaminated aquifer. Biodegradation 26:443-51.

11. Michalsen MM, Weiss R, King A, Gent D, Medina VF, Istok JD. 2013. Push-Pull Tests for Estimating RDX and TNT Degradation Rates in Groundwater. Groundwater Monitoring & Remediation 33:61-68.

12. Kitts CL, Green CE, Otley RA, Alvarez MA, Unkefer PJ. 2000. Type I nitroreductases in soil *Enterobacteria* reduce TNT (2,4,6,-trinitrotoluene) and RDX (hexahydro-1,3,5-trinitro-1,3,5-triazine). Can J Microbiol 46:278-82.

13. Peterson FJ, Mason RP, Hovsepian J, Holtzman JL. 1979. Oxygen-sensitive and -insensitive nitroreduction by *Escherichia coli* and rat hepatic microsomes. J Biol Chem 254:4009-14.

14. Li WZ, Yin P, Yang YX. 1987. [Properties of TNT-degrading enzymes in intact cells of *Citrobacter freundii*]. Wei Sheng Wu Xue Bao 27:257-63.

15. Caballero A, Lazaro JJ, Ramos JL, Esteve-Nunez A. 2005. PnrA, a new nitroreductase-family enzyme in the TNT-degrading strain *Pseudomonas putida* JLR11. Environ Microbiol 7:1211-9.

16. Indest KJ, Crocker FH, Athow R. 2007. A TaqMan polymerase chain reaction method for monitoring RDX-degrading bacteria based on the *xplA* functional gene. J Microbiol Methods 68:267-74.

17. Bernstein A, Adar E, Nejidat A, Ronen Z. 2011. Isolation and characterization of RDX-degrading *Rhodococcus* species from a contaminated aquifer. Biodegradation 22:997-1005.

18. Seth-Smith HM, Rosser SJ, Basran A, Travis ER, Dabbs ER, Nicklin S, Bruce NC. 2002. Cloning, sequencing, and characterization of the hexahydro-1,3,5-Trinitro-1,3,5-triazine degradation gene cluster from Rhodococcus rhodochrous. Appl Environ Microbiol 68:4764-71.

19. Andeer PF, Stahl DA, Bruce NC, Strand SE. 2009. Lateral transfer of genes for hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) degradation. Appl Environ Microbiol 75:3258-62.

20. Bhushan B, Halasz A, Spain JC, Hawari J. 2002. Diaphorase catalyzed biotransformation of RDX via N-denitration mechanism. Biochem Biophys Res Commun 296:779-84.

21. Chakraborty S, Sakka M, Kimura T, Sakka K. 2008. Cloning and expression of a *Clostridium kluyveri* gene responsible for diaphorase activity. Biosci Biotechnol Biochem 72:735-41.

22. Fuller ME, McClay K, Hawari J, Paquet L, Malone TE, Fox BG, Steffan RJ. 2009. Transformation of RDX and other energetic compounds by xenobiotic reductases *XenA* and *XenB*. Appl Microbiol Biotechnol 84:535-44.

23. Fuller ME, McClay K, Higham M, Hatzinger PB, Steffan RJ. 2010. Hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) Bioremediation in Groundwater: Are Known RDX-Degrading Bacteria the Dominant Players? Bioremediation Journal 14:121-134.

24. Collier JM, Chai B, Cole JR, Michalsen MM, Cupples AM. 2019. High throughput quantification of the functional genes associated with RDX biodegradation using the SmartChip real-time PCR system. Appl Microbiol Biotechnol 103:7161-7175.

25. Li RW, Giarrizzo JG, Wu S, Li W, Duringer JM, Craig AM. 2014. Metagenomic Insights into the RDX-Degrading Potential of the Ovine Rumen Microbiome. PLOS ONE 9:e110505.

26. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards RA. 2008. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics 9:386.

27. Cox MP, Peterson DA, Biggs PJ. 2010. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics 11:485.

28. Rho M, Tang H, Ye Y. 2010. FragGeneScan: predicting genes in short and error-prone reads. Nucleic Acids Res 38:e191.

29. Pruitt KD, Tatusova T, Maglott DR. 2005. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 33:D501-4.

30. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44:D457-62.

31. Lee BU, Choi MS, Kim DM, Oh KH. 2017. Genome Shuffling of *Stenotrophomonas maltophilia* OK-5 for Improving the Degradation of Explosive RDX (Hexahydro-1,3,5-trinitro-1,3,5-triazine). Curr Microbiol 74:268-276.

32. Fish JA, Chai B, Wang Q, Sun Y, Brown CT, Tiedje JM, Cole JR. 2013. FunGene: the functional gene pipeline and repository. Front Microbiol 4:291.

33. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nature Methods 12:59-60.

34. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-20.

35. Papadopoulos JS, Agarwala R. 2007. COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics 23:1073-9.

36. Katoh K, Rozewicki J, Yamada KD. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160-1166.

37. Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47:W256-W259.

38. Sherrill-Mix S. 2019. taxonomizr: Functions to Work with NCBI Accessions and Taxonomy. , vR package version 0.5.3. . https://CRAN.R-project.org/package=taxonomizr.

39. RStudio Team. 2020. RStudio: Integrated Development for R., RStudio, PBC. Boston, MA, http://www.rstudio.com/.

40. R Core Team. 2018. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. , https://www.R-project.org/.

41. Harrell FE. 2020. Hmisc: Harrell Miscellaneous, https://cran.r-project.org/web/packages/Hmisc/index.html.

42. Bates D, Maechler M, Davis TA, Oehlschlägel J, Riedy J. 2019. Matrix: Sparse and Dense Matrix Classes and Methods, https://cran.r-project.org/web/packages/Matrix/index.html.

43. Csardi G, Nepusz T. 2006. The igraph software package for complex network research. InterJournal, Complex Systems 1695.

44. Bastian M, Heymann S, Jacomy M. 2006. Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media.

45. Aronesty E. 2013. Comparison of Sequencing Utility Programs. The Open Bioinformatics Journal 7:1-8.

46. Huson DH, Beier S, Flade I, Gorska A, El-Hadidi M, Mitra S, Ruscheweyh HJ, Tappu R. 2016. MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. PLoS Comput Biol 12:e1004957.

47. Parks DH, Tyson GW, Hugenholtz P, Beiko RG. 2014. STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics 30:3123-4.

48. Ritalahti KM, Amos BK, Sung Y, Wu Q, Koenigsberg SS, Loffler FE. 2006. Quantitative PCR targeting 16S rRNA and reductive dehalogenase genes simultaneously monitors multiple Dehalococcoides strains. Appl Environ Microbiol 72:2765-74.

49. Stedtfeld RD, Williams MR, Fakher U, Johnson TA, Stedtfeld TM, Wang F, Khalife WT, Hughes M, Etchebarne BE, Tiedje JM, Hashsham SA. 2016. Antimicrobial resistance dashboard application for mapping environmental occurrence and resistant pathogens. FEMS Microbiol Ecol 92.

50. Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis., https://ggplot2.tidyverse.org.

51. Kahng H-Y, Lee B-U, Cho Y-S, Oh K-H. 2007. Purification and characterization of the NAD(P)H-nitroreductase for the catabolism of 2,4,6-trinitrotoluene (TNT) in *Pseudomonas* sp. HK-6. Biotechnology and Bioprocess Engineering 12:433.

52. Rylott EL, Jackson RG, Sabbadin F, Seth-Smith HMB, Edwards J, Chong CS, Strand SE, Grogan G, Bruce NC. 2011. The explosive-degrading cytochrome P450 XplA: Biochemistry, structural features and prospects for bioremediation. Biochimica Et Biophysica Acta-Proteins and Proteomics 1814:230-236.

53. Coleman NV, Nelson DR, Duxbury T. 1998. Aerobic biodegradation of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) as a nitrogen source by a *Rhodococcus* sp., strain DN22. Soil Biology & Biochemistry 30:1159-1167.

54. Nejidat A, Kafka L, Tekoah Y, Ronen Z. 2008. Effect of organic and inorganic nitrogenous compounds on RDX degradation and cytochrome P-450 expression in Rhodococcus strain YH1. Biodegradation 19:313-320.

55. Seth-Smith HMB, Edwards J, Rosser SJ, Rathbone DA, Bruce NC. 2008. The explosive-degrading cytochrome P450 system is highly conserved among strains of *Rhodococcus* spp. Applied and Environmental Microbiology 74:4550-4552.

56. Thompson KT, Crocker FH, Fredrickson HL. 2005. Mineralization of the cyclic nitramine explosive hexahydro-1,3,5-trinitro-1,3,5-triazine by *Gordonia* and *Williamsia* spp. Applied and Environmental Microbiology 71:8265-8272.

57. Fuller ME, Perreault N, Hawari J. 2010. Microaerophilic degradation of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) by three *Rhodococcus* strains. Letters in Applied Microbiology 51:313-318.

58. Wilson FP, Cupples AM. 2016. Microbial community characterization and functional gene quantification in RDX-degrading microcosms derived from sediment and groundwater at two naval sites. Applied Microbiology and Biotechnology 100:7297-7309.

59. Bryant C, DeLuca M. 1991. Purification and characterization of an oxygen-insensitive NAD(P)H nitroreductase from *Enterobacter cloacae*. J Biol Chem 266:4119-25.

60. Bryant C, Hubbard L, McElroy WD. 1991. Cloning, nucleotide sequence, and expression of the nitroreductase gene from *Enterobacter cloacae*. J Biol Chem 266:4126-30.

61. Kwon MJ, O'Loughlin EJ, Antonopoulos DA, Finneran KT. 2011. Geochemical and microbiological processes contributing to the transformation of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) in contaminated aquifer material. Chemosphere 84:1223-30.

62. Livermore JA, Jin YO, Arnseth RW, Lepuil M, Mattes TE. 2013. Microbial community dynamics during acetate biostimulation of RDX-contaminated groundwater. Environ Sci Technol 47:7672-8.

63. Wang D, Boukhalfa H, Marina O, Ware DS, Goering TJ, Sun F, Daligault HE, Lo CC, Vuyisich M, Starkenburg SR. 2017. Biostimulation and microbial community profiling reveal insights on RDX transformation in groundwater. Microbiologyopen 6.

64. Jugnia L-B, Manno D, Dodard S, Greer CW, Hendry M. 2019. Manipulating redox conditions to enhance in situ bioremediation of RDX in groundwater at a contaminated site. Science of The Total Environment 676:368-377.

65. Adrian NR, Arnett CM. 2004. Anaerobic biodegradation of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) by *Acetobacterium malicum* strain HAAP-1 isolated from a methanogenic mixed culture. Curr Microbiol 48:332-40.

66. Sherburne LA, Shrout JD, Alvarez PJ. 2005. Hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) degradation by *Acetobacterium paludosum*. Biodegradation 16:539-47.

67. Bhatt M, Zhao JS, Halasz A, Hawari J. 2006. Biodegradation of hexahydro-1,3,5-trinitro-1,3,5-triazine by novel fungi isolated from unexploded ordnance contaminated marine sediment. J Ind Microbiol Biotechnol 33:850-8.

68. Kwon MJ, Finneran KT. 2008. Hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) and Octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (HMX) Biodegradation Kinetics Amongst Several Fe(III)-Reducing Genera. Soil and Sediment Contamination: An International Journal 17:189-203.

69. Eaton HL, Duringer JM, Murty LD, Craig AM. 2013. Anaerobic bioremediation of RDX by ovine whole rumen fluid and pure culture isolates. Appl Microbiol Biotechnol 97:3699-710.

70. Singh R, Soni P, Kumar P, Purohit S, Singh A. 2009. Biodegradation of high explosive production effluent containing RDX and HMX by denitrifying bacteria. World Journal of Microbiology and Biotechnology 25:269-275.

71. Lee S-Y, Brodman BW. 2004. Biodegradation of 1,3,5-Trinitro-1,3,5-triazine (RDX). Journal of Environmental Science and Health, Part A 39:61-75.

72. Kitts CL, Cunningham DP, Unkefer PJ. 1994. Isolation of three hexahydro-1,3,5-trinitro-1,3,5-triazine-degrading species of the family *Enterobacteriaceae* from nitramine explosive-contaminated soil. Applied and Environmental Microbiology 60:4608.

73. Zhang C, Hughes JB. 2003. Biodegradation pathways of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) by *Clostridium acetobutylicum* cell-free extract. Chemosphere 50:665-671.

74. Regan KM, Crawford RL. 1994. Characterization of *Clostridium bifermentans* and its biotransformation of 2,4,6-trinitrotoluene (TNT) and 1,3,5-triaza-1,3,5-trinitrocyclohexane (RDX). Biotechnology Letters 16:1081-1086.

75. Zhao J-S, Paquet L, Halasz A, Manno D, Hawari J. 2004. Metabolism of octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine by *Clostridium bifermentans* strain HAW-1 and several other H2-producing fermentative anaerobic bacteria. FEMS Microbiology Letters 237:65-72.

76. Kwon MJ, Wei N, Millerick K, Popovic J, Finneran K. 2014. *Clostridium geopurificans* Strain MJ1 sp. nov., A Strictly Anaerobic Bacterium that Grows via Fermentation and Reduces the Cyclic Nitramine Explosive Hexahydro-1,3,5-Trinitro-1,3,5-Triazine (RDX). Current Microbiology 68:743-750.

77. Bhushan B, Halasz A, Thiboutot S, Ampleman G, Hawari J. 2004. Chemotaxis-mediated biodegradation of cyclic nitramine explosives RDX, HMX, and CL-20 by *Clostridium* sp. EDB2. Biochemical and Biophysical Research Communications 316:816-821.

78. Zhao J-S, Spain J, Hawari J. 2003. Phylogenetic and metabolic diversity of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX)-transforming bacteria in strictly anaerobic mixed cultures enriched on RDX as nitrogen source. FEMS Microbiology Ecology 46:189-196.

79. Zhao J-S, Spain J, Thiboutot S, Ampleman G, Greer C, Hawari J. 2004. Phylogeny of cyclic nitramine-degrading psychrophilic bacteria in marine sediment and their potential role in the natural attenuation of explosives. FEMS Microbiology Ecology 49:349-357.

80. Arnett CM, Adrian NR. 2009. Cosubstrate independent mineralization of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) by a *Desulfovibrio* species under anaerobic conditions. Biodegradation 20:15-26.

81. Boopathy R, Gurgas M, Ullian J, Manning JF. 1998. Metabolism of Explosive Compounds by Sulfate-Reducing Bacteria. Current Microbiology 37:127-131.

82. Ronen Z, Yanovich Y, Goldin R, Adar E. 2008. Metabolism of the explosive hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) in a contaminated vadose zone. Chemosphere 73:1492-1498.

83. Bhatt M, Zhao J-S, Monteil-Rivera F, Hawari J. 2005. Biodegradation of cyclic nitramines by tropical marine sediment bacteria. Journal of Industrial Microbiology and Biotechnology 32:261-267.

84. Zhao J-S, Halasz A, Paquet L, Beaulieu C, Hawari J. 2002. Biodegradation of Hexahydro-1,3,5-Trinitro-1,3,5-Triazine and Its Mononitroso Derivative Hexahydro-1-Nitroso-3,5-Dinitro-1,3,5-Triazine by *Klebsiella pneumoniae* Strain SCZ-1 Isolated from an Anaerobic Sludge. Applied and Environmental Microbiology 68:5336.

85. Van Aken B, Yoon JM, Schnoor JL. 2004. Biodegradation of Nitro-Substituted Explosives 2,4,6-Trinitrotoluene, Hexahydro-1,3,5-Trinitro-1,3,5-Triazine, and Octahydro-1,3,5,7-Tetranitro-1,3,5-Tetrazocine by a Phytosymbiotic *Methylobacterium* sp. Associated with Poplar Tissues (*Populus deltoides x nigra* DN34). Applied and Environmental Microbiology 70:508.

86. Fournier D, Trott S, Hawari J, Spain J. 2005. Metabolism of the Aliphatic Nitramine 4-Nitro-2,4-Diazabutanal by *Methylobacterium* sp. Strain JS178. Applied and Environmental Microbiology 71:4199.

87. Fournier D, Halasz A, Spain J, Spanggord RJ, Bottaro JC, Hawari J. 2004. Biodegradation of the Hexahydro-1,3,5-Trinitro-1,3,5-Triazine Ring Cleavage Product 4-Nitro-2,4-Diazabutanal by *Phanerochaete chrysosporium*. Applied and Environmental Microbiology 70:1123.

88. Chang HW, Kahng HY, Kim SI, Chun JW, Oh KH. 2004. Characterization of *Pseudomonas* sp. HK-6 cells responding to explosive RDX (hexahydro-1,3,5-trinitro-1,3,5-triazine). Applied Microbiology and Biotechnology 65:323-329.

89. Coleman NV, Spain JC, Duxbury T. 2002. Evidence that RDX biodegradation by *Rhodococcus* strain DN22 is plasmid-borne and involves a cytochrome p-450. Journal of Applied Microbiology 93:463-472.

90. Young DM, Unkefer PJ, Ogden KL. 1997. Biotransformation of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) by a prospective consortium and its most effective isolate *Serratia marcescens*. Biotechnology and Bioengineering 53:515-522.

91. Zhao J-S, Manno D, Leggiadro C, apos, Neil D, Hawari J. 2006. *Shewanella halifaxensis* sp. nov., a novel obligately respiratory and denitrifying psychrophile. International Journal of Systematic and Evolutionary Microbiology 56:205-212.

92. Zhao J-S, Manno D, Beaulieu C, Paquet L, Hawari J. 2005. *Shewanella sediminis* sp. nov., a novel Na+-requiring and hexahydro-1,3,5-trinitro-1,3,5-triazine-degrading bacterium from marine sediment. International Journal of Systematic and Evolutionary Microbiology 55:1511-1520.

93. Binks PR, Nicklin S, Bruce NC. 1995. Degradation of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) by *Stenotrophomonas maltophilia* PB1. Applied and Environmental Microbiology 61:1318.

94. Mahan KM, Zheng H, Fida TT, Parry RJ, Graham DE, Spain JC. 2017. Iron-Dependent Enzyme Catalyzes the Initial Step in Biodegradation of Nitroglycine by *Variovorax* sp. Strain JS1663. Applied and Environmental Microbiology 83:e00457-17.

95. Cho K-C, Lee DG, Fuller ME, Hatzinger PB, Condee CW, Chu K-H. 2015. Application of 13C and 15N stable isotope probing to characterize RDX degrading microbial communities under different electron-accepting conditions. Journal of Hazardous Materials 297:42-51.

96. Bhushan B, Halasz A, Spain J, Thiboutot S, Ampleman G, Hawari J. 2002. Biotransformation of Hexahydro-1,3,5-trinitro-1,3,5-triazine Catalyzed by a NAD(P)H: Nitrate Oxidoreductase from *Aspergillus niger*. Environmental Science & Technology 36:3104-3108.

CHAPTER 4 Identification of the Phylotypes and Predicted Functional Genes Involved in *cis*-Dichloroethene and 1,4-Dioxane Aerobic Biodegradation in Soil Microcosms

This chapter is being prepared for submission to a peer reviewed journal: Hongyu Dang and Alison M. Cupples. Identification of the Phylotypes and Predicted Functional Genes Involved in *cis*-Dichloroethene and 1,4-Dioxane Aerobic Biodegradation in Soil Microcosms.

Abstract

Co-contamination with chlorinated compounds and 1,4-dioxane has been reported at many sites. Recently, there has been an increased interest in aerobic bioremediation because of the potential to degrade multiple contaminants concurrently. However, the likelihood of implementing a successful bioremediation approach is dependent on the microorganisms present. Towards improving bioremediation efficacy, the current study examined laboratory microcosms (inoculated separately with two soils) to determine the phylotypes and functional genes associated with the biodegradation of two common co-contaminants (*cis*-dichloroethene [cDCE] and 1,4-dioxane). The impact of amending microcosms with lactate on cDCE and 1,4-dioxane biodegradation was also investigated. In one soil, when all three substrates were present (1,4-dioxane, cDCE, lactate), 1,4-dioxane removal was slower compared to when only one additional substrate was added (cDCE or lactate). In contrast, in microcosms amended with another soil, 1,4-dioxane removal trends for all three treatments were similar, indicating the present of either lactate or cDCE or both did not impact 1,4-dioxane biodegradation. Lactate appeared to improve the biological removal of cDCE in microcosms inoculated with either soil. Stable isotope probing (SIP) was then used to determine which phylotypes were actively involved in carbon uptake from cDCE and 1,4-dioxane in both soil communities. The most enriched phylotypes for ^{13}C assimilation from 1,4-dioxane included *Rhodopseudomonas* and *Rhodanobacter*. Propane

monooxygenase was predicted (by PICRUSt2) to be dominant in the 1,4-dioxane amended microbial communities and propane monooxygenase gene abundance values correlated with other enriched (but less abundant) phylotypes for ^{13}C-1,4-dioxane assimilation. The dominant enriched phylotypes for ^{13}C assimilation from cDCE included *Bacteriovorax*, *Pseudomonas* and *Sphingomonas*. In the cDCE amended soil microcosms, PICRUSt2 predicted the presence of DNA encoding glutathione S-transferase (a known cDCE upregulated enzyme). Overall, the work demonstrated concurrent removal of cDCE and 1,4-dioxane by indigenous soil microbial communities and the enhancement of cDCE removal by lactate. The data generated on the phylotypes responsible for carbon uptake (as determined by SIP) could be incorporated into diagnostic molecular methods for site characterization. The results suggest aerobic concurrent biodegradation of cDCE and 1,4-dioxane should be considered for chlorinated solvent site remediation.

1. Introduction

The clean-up of sites with mixed contamination poses a significant challenge to the remediation community. Developing synergistic approaches that could reduce the concentrations of multiple contaminants has the potential to result in considerable cost savings. From the list of co-contaminants found in soil and groundwater, the chlorinated solvents and their metabolites (tetrachloroethene [PCE], trichloroethene [TCE], *cis*-dichloroethene [cDCE], vinyl chloride [VC]) are particularly prevalent (found at > 3,000 Department of Defense sites) and problematic due to their tendency to form large, dissolved-phase plumes, their recalcitrant nature and the subsequent risk to human health. Remediation efforts have frequently involved biostimulation, through the addition of carbon sources, or bioaugmentation, which involves in the injection of mixed microbial cultures containing *Dehalococcoides mccartyi* (1). With the expansion of this

remedial practice over the last decade, the number of sites in the US now numbers well over 2,300, and bioaugmentation has been performed in at least 11 other countries (*P Hatzinger, personal communication*).

Although clearly a very successful approach, bioremediation with *D. mccartyi* involves several significant limitations and thus may not be appropriate for all chlorinated solvent contaminated sites. Specifically, it is unlikely to be employed at large oxic sites because of the requirement for highly reducing conditions for *D. mccartyi* and the associated cost of driving such large sites anaerobic. Secondly, the approach will be less desirable at sites with multiple contaminants if those co-contaminants can be degraded more easily under aerobic conditions (e.g. benzene, toluene, 1,4-dioxane). Further, the accumulation of the known human carcinogen, VC, from the dechlorination process represents a significant risk if complete dechlorination does not occur. Additionally, driving sites anaerobic can result in long-term secondary groundwater impacts such as hydrogen sulfide formation, acidification, mobilization of reduced metals and methane accumulation. In contrast, aerobic approaches have the advantage that the geochemistry of the site is not significantly impacted.

cDCE is a major degradation product of TCE by both abiotic and biotic degradation (2). For example, at a TCE contaminated site (Dover Air Force Base, DE) dechlorination by the indigenous microbial community only transformed TCE to cDCE (3). In laboratory batch and column tests for enhanced biological dissolution of PCE, cDCE was the main product of PCE dehalogenation and accumulated when PCE and TCE were present at high concentrations (4, 5). In fact, “cDCE stall” is a well-recognized term in the remediation community for the

accumulation of cDCE at chlorinated solvent sites. Given the common occurrence of cDCE, identifying the potential for cDCE transformation remains an important issue.

1,4-Dioxane, a probable human carcinogen and common chlorinated solvent stabilizer, has been found at numerous contaminated sites across the U.S. (6, 7). In an examination from 49 remediation installations at U.S. Air Force sites, 1,4-dioxane was detected in 781 groundwater wells, and 64% of wells that contained 1,4-D also contained TCE (8). In an evaluation of >2000 sites in California, the chlorinated solvents were found in 94% of the sites with detections of 1,4-dioxane (9).

Many bacteria have been identified to metabolically or co-metabolically degrade 1,4-dioxane under aerobic conditions (10, 11). However, aerobic 1,4-dioxane degradation can be impacted by the presence of chlorinated compounds. For *Pseudonocardia dioxanivorans* strain CB1190, researchers reported 1,1,1-trichloroethane (1,1,1-TCA) and 1,1-dichloroethene (1,1-DCE) illustrated similar inhibitory effects on 1,4-dioxane degradation (12). In the same study, 1,1-DCE was a slightly more potent inhibitor for 1,4-dioxane degradation than 1,1,1-TCA for *Pseudomonas mendocina* strain KR1, while 1,1,1-TCA was a much more potent inhibitor for 1,4-dioxane degradation than 1,1-DCE for a *Escherichia coli* recombinant strain expressing toluene-4-monooxygenase from strain KR1 (12). A later study with *P. dioxanivorans* CB1190 indicated chlorinated compounds inhibited 1,4-dioxane biodegradation in the following order: 1,1-DCE > cDCE > TCE > 1,1,1-TCA (13).

To address the problem of co-contamination, efforts have focused specifically on the removal of both chlorinated compounds and 1,4-dioxane. To degrade TCE, DCEs with 1,4-dioxane, *P. dioxanivorans* CB1190 was combined with hydrogen peroxide and tungstated zirconia, which partially removed those contaminants with the remainder being degraded by *P.*

dioxanivorans CB1190 (14, 15). In another study, *P. dioxanivorans* CB1190 was combined with the anaerobic bioaugmentation culture KB-1 (a chloroethene degrading consortium) resulting in TCE transformation to cDCE, as well as cDCE and 1,4-dioxane degradation by *P. dioxanivorans* CB1190 (16). Another strain *Azoarcus* sp. DD4, was found to degrade 1,4-dioxane with 1,1-DCE using propane as the main substrate (17). More recently, *Azoarcus* sp. DD4 was sequentially used with SDC-9 (another chloroethene degrading consortium) to achieve transformation of TCE to cDCE and VC by SDC-9 and co-metabolic removal of VC, cDCE and 1,4-dioxane by *Azoarcus* sp. DD4 with the addition of propane (18). These studies suggest mixed microbial communities will likely be needed to facilitate co-contamination remediation. An interesting question arising from these trends concerns the biodegrading abilities of indigenous mixed communities and their potential contribution to site remediation.

Towards understanding the potential of natural mixed communities, the current work builds on previous research documenting aerobic 1,4-dioxane biodegradation in soil microcosms (19). In the current work, stable isotope probing (SIP) is utilized to identify which microorganisms are involved in carbon uptake from cDCE and 1,4-dioxane. SIP is a cultivation independent method to link identity with function (20) such as contaminant biodegradation (21-25). As aerobic contaminant biodegradation often relies on co-metabolism, the impact of an additional substrate (lactate) was also investigated. Lactate was selected because it is commonly used in biostimulation (to drive sites anaerobic) (3, 26) and would therefore already be acceptable to many regulatory agencies. The objectives were to 1) examine removal rates of the co-contaminants cDCE and 1,4-dioxane, with and without lactate addition, with indigenous mixed microbial communities 2) identify the microorganisms responsible for the uptake of ^{13}C from cDCE as well as from 1,4-dioxane during biodegradation and 3) predict the functional

genes present and correlate their presence to specific phylotypes. The overall rationale behind the current project is to provide knowledge to enhance the aerobic remediation of two important groundwater contaminants (cDCE, 1,4-dioxane) for oxic sites.

2. Methods

2.1 Chemicals and Soil Inocula

Unlabeled 1,4-dioxane (99.8%) and cDCE were purchased from Sigma-Aldrich (MO, USA). Labeled 1,4-dioxane [$(^{13}C)_4H_8O_2$] was purchased from Santa Cruz Biotechnology (TX, USA) with 99.2% isotopic purity and 98% purity, and labeled cDCE [$^{13}C_2H_2Cl_2$] was purchased from Sigma-Aldrich (MO, USA) with 99% isotopic purity and 97% purity. Two soils were collected from 5 sampling stations in 6 replicate plots within Treatments 1 and 2 at the Michigan State University (MSU) Main Cropping System Experiment at Kellogg Biological Station Long-Term Ecological Research (KBS LTER) (42°24′N, 85°23′W). Both soils received conventional levels of chemical inputs, however, Treatment 1 is chisel plowed and Treatment 2 is under no-till management. For additional information see https://lter-kbs-msu-edu.proxy1.cl.msu.edu/research/site-description-and-maps/. All samples for each treatment were mixed, then stored at 4 °C in the dark. These soils were selected because the analysis of shotgun sequencing data generated from a previous study (27) indicated the presence of numerous microorganisms previously associated with 1,4-dioxane and cDCE biodegradation (as discussed in the results section).

2.2 Microcosms Setup

For each set of amendments, microcosms were established in 160 mL serum bottles (wrapped with aluminum foil) with 10 g of soil and 20 mL of media. For each soil, triplicate microcosms were amended with one of the following four sets of amendments: 1) cDCE, 1,4-dioxane and

lactate, 2) cDCE and 1,4-dioxane, without lactate, 3) 1,4-dioxane with lactate and 4) cDCE with lactate. All microcosms were closed with a rubber seal and aluminum crimp. For each soil, all four treatments included triplicate abiotic autoclaved controls. All bottles were incubated at room temperature (~21 °C) on a shaker (110 rpm). The media was based on that used to enrich the 1,4-dioxane degrader *Pseudonocardia dioxanivorans* CB1190 (except nitrilotriacetic acid was removed) (28). One liter of the media contained 100 mL of a buffer stock [K_2HPO_4 (32.4 g/L), KH_2PO_4 (10 g/L), NH_4Cl (20 g/L)] and 100 mL of a trace metal stock [$MgSO_4.7H_2O$ (2 g/L), $FeSO_4.7H_2O$ (0.12 g/L), $MnSO_4.H_2O$ (0.03 g/L), $ZnSO_4.7H_2O$ (0.03 g/L), and $CoCl_2.6H_2O$ (0.01 g/L)]. The initial liquid concentrations of 1,4-dioxane and cDCE were ~ 6 mg/L and ~ 4 mg/L. The liquid concentration of cDCE was calculated based on Henry's law (29). The treatments with sodium lactate were amended at 0.56 g/L (~5 mM).

Additional microcosms were established for each soil (160 mL bottles, 10 g soil 1 or 2, same media) for the SIP experiments. For each soil, triplicate abiotic control microcosms (sterilized by autoclaving) and six microcosms were amended with unlabeled 1,4-dioxane or cDCE (similar concentrations as above). Another six microcosms were amended with ^{13}C labeled 1,4-dioxane or ^{13}C labeled cDCE. As the above experiments indicated cDCE was improved by the addition of lactate, the cDCE bottles were also amended with 5 mM of lactate (and closed with a rubber seal and aluminum crimp). The 1,4-dioxane amended bottles were not amended with lactate and were opened for 6 hours every three days for aeration.

2.3 Analytic Methods

Liquid samples (0.1 mL) were withdrawn (with sterilized disposable needles and a 1 mL syringe), then filtered (with a 0.22 µm, 4 mm nylon syringe filter, Thomas Scientific, NJ) for 1,4-dioxane analysis. The filtered samples were injected into a GC-FID (Hewlett Packard 5890)

equipped with a column (Restek, Stabilwax-DB, 30m, 0.53 mmID, 1µm) using a similar method to that previously described (30). The injector temperature was maintained at 220 °C and the detector temperature was set at 250 °C. The oven temperature was programmed to initiate at 80 °C for 1 min, then increased to 140 °C with a ramp of 20 °C /min. The gas phase concentration of cDCE was determined (1 mL gastight syringe, 0.2 mL of the gaseous sample) with a GC-FID (Hewlett Packard 5890) equipped with a capillary column (Alltech, AT-624, 30m × 0.53mm ID × 3.0µm) using a similar method described in a previous study (31). The injector temperature was maintained at 180 °C and the detector at 240 °C. The oven temperature was programmed to initiate at 45 °C for 4 min, then increased to 165 °C with a ramp of 20 °C /min, held at 165 °C for 1 min.

2.4 DNA extraction, Fractionation and Sequencing

Duplicate soil 1 and soil 2 inoculated microcosms amended with either labeled or unlabeled chemicals (16 bottles, 2 chemicals, 2 with unlabeled amendment X 2 with labeled amendment X 2 soils) were sacrificed for DNA extraction at ~50% cDCE or 1,4-dioxane removal using QIAGEN PowerSoil DNA extraction kit as per manufacturer's protocol. DNA extracts (approx. 10 µg) were loaded into Quick-Seal polyallomer tubes (13 by 51 mm, 5.1 mL; Beckman Coulter (Brea, CA) along with a Tris-EDTA (10 mM Tris, 1 mM EDTA, pH 8)-CsCl solution for ultracentrifugation. The density of the mixture inside the tube was determined with a model AR200 digital refractometer (Leica Microsystems Inc., Buffalo Grove, IL), and it was adjusted to a final value of 1.730 g/mL by adding small volumes of CsCl solution or TE buffer until the tube could be sealed. The sealed tubes were then ultracentrifuged at 178,000×g (20 °C) for 46 h in a StepSaver 70 V6 vertical titanium rotor (8 by 5.1 mL capacity) within a Sorvall WX 80 Ultra Series centrifuge (Thermo Scientific, Waltham, MA). Following ultracentrifugation, the

tubes were placed onto a fraction system (Beckman Coulter) and fractions (~26, 200 μL) were collected. The buoyant density of each fraction was measured, and CsCl was removed by glycogen-assisted ethanol precipitation. The DNA concentration in each fraction was quantified using the Quant-iT™ dsDNA High-Sensitivity Assay Kit.

For each chemical (labeled and unlabeled) and soil, triplicate DNA extracts of three fractions with higher buoyant density (1.73-1.75 g/mL) and one fraction with lighter buoyant density (~1.70 g/mL) were submitted for 16S rRNA gene amplicon sequencing at Research Technology Support Facility (RTSF) at MSU. The heavy buoyant density fractions were selected based on the minimum concentration of DNA required for sequencing. In total, two 96 well plates were submitted for sequencing (2 chemicals, 2 soils, 4 fractions, 3 replicate fractions, 2 microcosm replicates, 2 isotopes). The V4 hypervariable region of the 16S rRNA gene was amplified using dual indexed Illumina compatible primers 515f/806r as described by James Kozich (32). PCR products were batch normalized using Invitrogen SequalPrep DNA Normalization plates and the products recovered from the plates pooled. The pool was cleaned up and concentrated using AmpureXP magnetic beads; it was QC'd and quantified using a combination of Qubit dsDNA HS, Agilent 4200 TapeStation HS DNA1000 and Kapa Illumina Library Quantification qPCR assays. The pool was loaded onto an Illumina MiSeq v2 standard flow cell and sequencing was performed in a 2x250 bp paired end format using a MiSeq v2 500 cycle reagent cartridge. Custom sequencing and index primers were added to appropriate wells of the reagent cartridge. Base calling was performed by Illumina Real Time Analysis (RTA) v1.18.54 and output of RTA was demultiplexed and converted to FastQ format with Illumina Bcl2fastq v2.19.1. The sequencing data for 1,4-dioxane and cDCE SIP was submitted to NCBI

under Bioproject PRJNA719874 (accession numbers SAMN18623434 to SAMN18623529) and PRJNA719920 (accession numbers SAMN18624005 to SAMN18624100), respectively.

2.5 Analysis of Sequencing Data

The amplicon sequencing data in the fastq file format was analyzed with Mothur (version 1.44.2) (33) using the MiSeq Standard Operating Procedure (32). The procedure included trimming the raw sequences and quality control. The database used for alignment was SILVA bacteria database (Release 138) for the V4 region (34). Chimeras, mitochondrial and chloroplast lineage sequences were removed, then the sequences were classified into OTUs. The downstream analysis was conducted using microbiome (35) (version 1.10.0), phyloseq (36) (version 1.32.0), ampvis2 (37) (version 2.6.5), ggplot2 (38) (version 3.3.2), Hmisc (39) (version 4.4-1), Matrix (40) (version 1.2-18), igraph (41) (version 1.2.6), ggpubr (42) (version 0.4.0) in R (43) (version 4.0.4) with R studio (44) (version 1.1.456). Additionally, the software Statistical Analysis of Taxonomic and Functional Profiles (STAMP) (45) (version 2.1.3) was utilized to statistically analyze the results.

The package microbiome was used to combine the OTUs shared file, taxonomy and metadata and it was also used to transform the counts of reads for OTUs into relative abundance. Phyloseq and ggplot2 were used for the Non-metric Multi-dimensional Scaling (NMDS) plots, alpha diversity measurements plots, the bar plot for the classification of the microbial community at phylum level for different soils and treatments and for exporting a subset of OTUs information based on the variables in metadata. Ampvis2 was used to generate the rarefaction curves and heatmaps of average abundance at the genus level of the sample replicates. OTUs with at least 0.06% average relative abundance and 50% occurrence were selected for building the correlation network. The packages Hmisc and Matrix were used to calculate the correlation of OTUs with

Spearman correlation. Strong correlations (correlation coefficient ≥ 0.7) and Benjamini-Hochberg method adjusted p value ($p < 0.01$) were set to filter the correlation result. The filtered correlation result were used to build occurrence network with the package igraph and these were visualized in Gephi (46).

The OTUs enriched in the heavy fractions of the ^{13}C labeled cDCE or 1,4-dioxane amended microcosms were determined using STAMP. Specifically, OTU relative abundance was compared between the heavy fractions of the microcosms amended with the labeled chemical and the heavy fractions of the microcosms amended with the unlabeled chemical (Welch's two-sided t-test, $p < 0.05$, with default settings). STAMP was also used to investigate which OTUs were enriched in the light fractions of the labeled amended microcosms compared to the light fractions of the unlabeled amended microcosms to eliminate false positives in the heavy fraction analysis. In addition, the enriched OTUs were subject to further statistical analysis in RStudio (Wilcoxon Rank Test, ggplot2 and ggpubr).

2.6 Function Prediction and Correlation

The microbial functions from KEGG orthologs (KO) (47) were predicted from the sequencing data using the PICRUSt2 pipeline (48). The functions related to 1,4-dioxane and cDCE degradation identified in previous research were manually picked (10, 11, 49-53) to generate the heatmaps of relative abundance across all sample replicates with the R package ComplexHeatmap (54) (version 2.4.3). In addition, OTUs from 1,4-dioxane and cDCE SIP experiments with an average relative abundance ≥ 0.05% were collected and pooled together with functions associated with 1,4-dioxane and cDCE degradation for running Spearman correlations. OTUs correlated with at least 4 and 2 functions for 1,4-dioxane and cDCE

degradation, respectively, with absolute values of correlation coefficients higher than 0.6 were chosen for plotting the heatmaps with the same R package.

2.7 Species Associated with 1,4-dioxane and DCE Degradation

Previously, our group submitted DNA extracts from the same soils (Treatments 1 and 2 from KBS) for shotgun sequencing (27). In the current work, the shotgun sequences (processed by Trimmomatic (55)) were assembled with Megahit (56) (version 1.2.4) using the pair end plus single end option. Minimum and maximum kmer sizes were 27 and 127 with the kmer size step of 10. Previously identified 1,4-dioxane and cDCE degraders were searched for in the National Center for Biotechnology Information (NCBI) taxonomy browser to find their lowest ranks (primarily rank of species) and taxonomy IDs (Supplementary Table 4.1). The assembled reads were then aligned to the NCBI nucleotide database (nt) with the taxids option in BLASTN (57) (version 2.10.0-Linux_x86_64). The results were restricted to evalue $\leq 1\times 10^{-5}$ (with output format 6) and identity ≥ 60 % and the resulting files were then imported into Megan (58) (community edition version 6.19.7). Each BLASTN output was processed to map the Megan genomic DNA accession database for generating the phylogenetic trees of the species associated with 1,4-dioxane or DCE degradation.

3. Results

3.1 Degradation of 1,4-Dioxane and cDCE With or Without Lactate

1,4-Dioxane and cDCE concentrations were determined in microcosms with inocula from two soils, with or without the additional amendment(s) of lactate/1,4-dioxane/cDCE (Figures 1 and 2). For soil 1 inoculated microcosms, when all three substrates were added together, the differentiation between removal trends for 1,4-dioxane between the samples and abiotic controls was not strong, with 95% confidence intervals (CIs) for the regression lines overlapping the

entire course of the incubation (Figure 4.1A, part A). However, when only lactate was added, the 1,4-dioxane regression lines 95% CIs between the abiotic controls and samples separated before day 20 (Figure 4.1A, part B). Similarly, when only cDCE was added, the 1,4-dioxane regression lines 95% CIs between the samples and abiotic controls separated at ~ day 20 (Figure 4.1, part C). Based on these trends, for this particular microbial community, one hypothesis is that when all three substrates are present (Figure 4.1A, part A), 1,4-dioxane removal is slower compared to when only one additional substrate is added (Figure 4.1A, parts B and C). Compared to soil 1 microcosms, the impact of additional chemicals on 1,4-dioxane removal was different for soil 2 microcosms. That is, 1,4-dioxane biodegradation was similar for all three treatments, indicating the present of either lactate or cDCE or both did not impact 1,4-dioxane removal in this soil community (Figure 4.2A, parts A-C).

The most notable trend for cDCE biodegradation in both soil 1 and soil 2 microcosms concerns the addition of lactate. In soil 1 microcosms, without the addition of lactate, cDCE removal was slower, as indicated by the slope of the regression line (0.015) (Figure 4.1B, part C) compared to regression line slopes (0.022 and 0.021) from the treatments with lactate (Figure 4.1B, parts A and B). In soil 2 microcosms, although the cDCE regression line slope (0.026) was greater when lactate was not added (Figure 4.2B, part C), the overlap between the regression lines 95% CIs remained until ~ day 40, compared to ~ day 10 for both of the lactate amended treatments (Figure 4.2C, parts A and B). The trends for both soils support the hypothesis that the presence of lactate improves the biological removal of cDCE. Based on these results, lactate was added to the cDCE SIP experiments but not to the 1,4-dioxane SIP experiments.

For soil 1 microcosms, the cDCE regression line slopes were similar (0.022 vs. 0.021) when either lactate was added with 1,4-dioxane (Figure 4.1B, part A), or when only lactate was

added (Figure 4.1B, part B), suggesting 1,4-dioxane does not impact cDCE removal in this soil community. For soil 2 microcosms, the cDCE regression line slopes differed (0.022 vs. 0.017) when either lactate was added with 1,4-dioxane (Figure 4.2B, part A), or when lactate was added by itself (Figure 4.2B, part B), again suggesting 1,4-dioxane likely does not impact cDCE removal (when lactate is present).

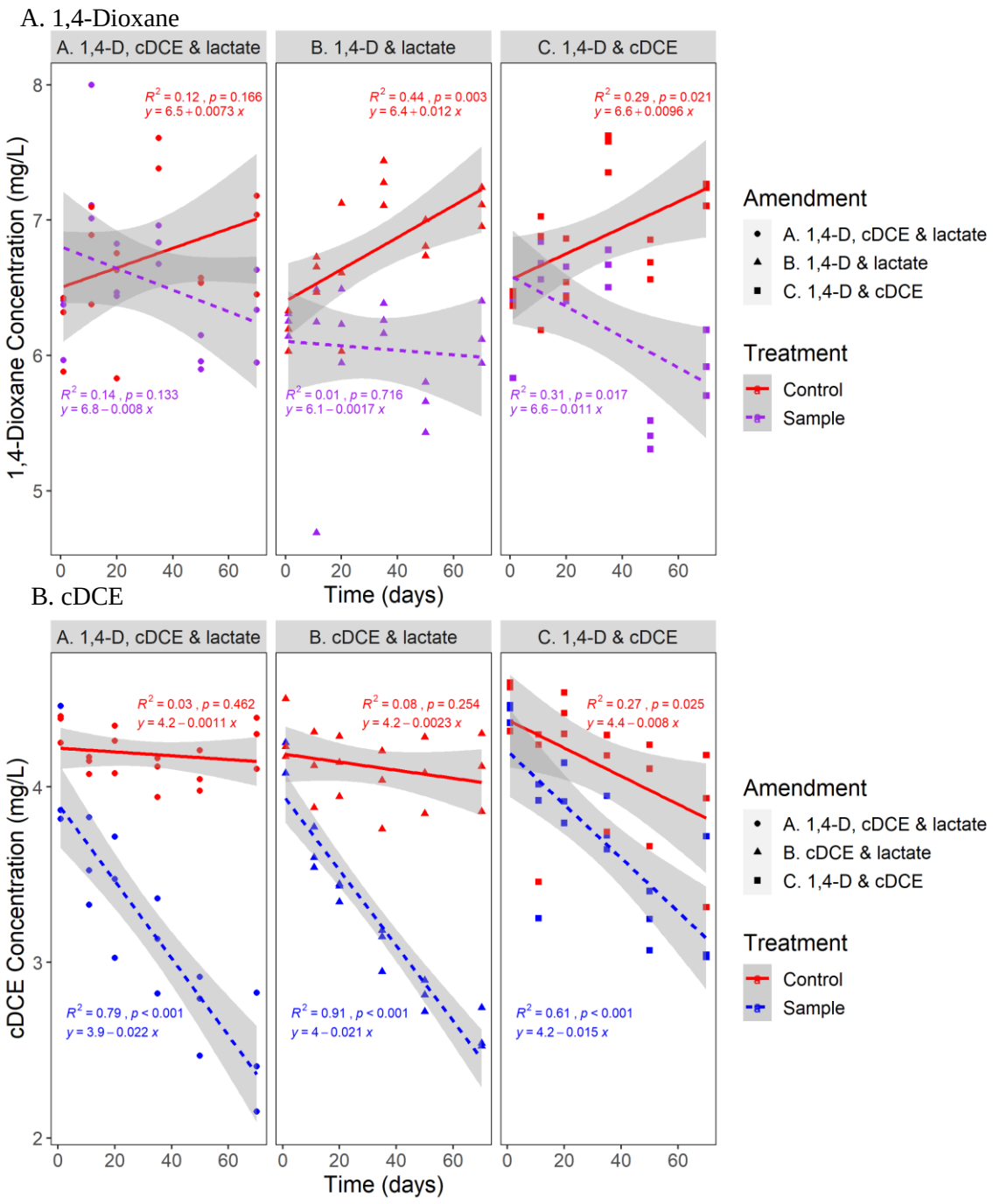


Figure 4. 1. 1,4-Dioxane (A) and cDCE (B) concentrations in triplicate sample microcosms (purple [A] and blue [B]) and triplicate abiotic controls (red) inoculated with soil 1 and different amendments. The shaded areas indicate 95% confidence intervals along the linear regression model.

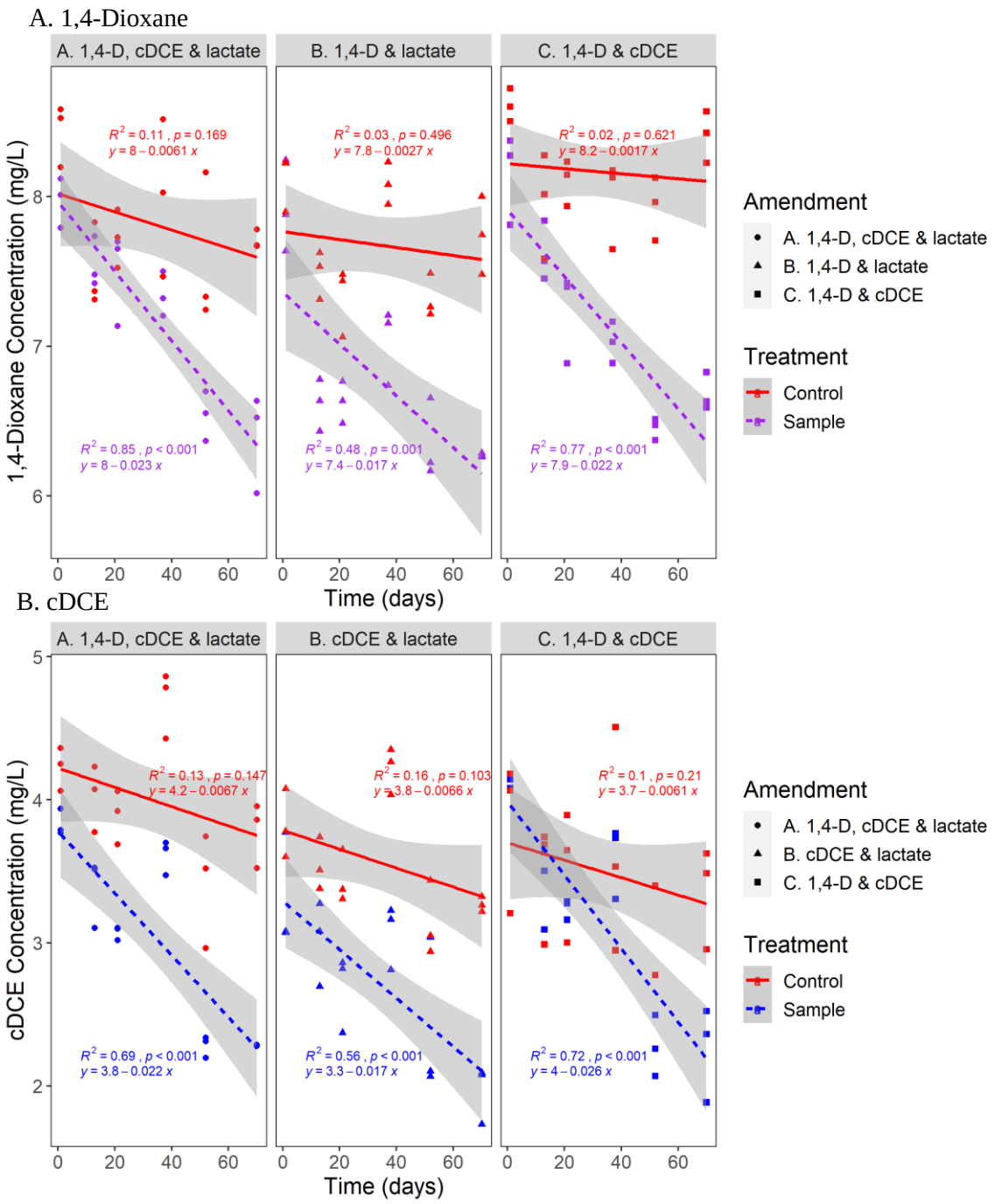


Figure 4. 2. 1,4-Dioxane (A) and cDCE (B) concentrations in triplicate sample microcosms (purple [A] and blue [B]) and triplicate abiotic controls (red) inoculated with soil 2 and different amendments. The shaded areas indicate 95% confidence intervals along the linear regression model.

In microcosms amended with all three substrates, decreases in cDCE concentrations occurred earlier than decreases for 1,4-dioxane in both soils 1 and 2 (as shown by an earlier

separation of the regression line 95% CIs between the samples and controls, Figure 4.1A, B, part A, Figure 4.2A, B part A). The pattern suggests there is a sequential removal for cDCE and 1,4-dioxane with the addition of lactate. In comparison, when the microcosms were amended with only 1,4-dioxane and cDCE (Figure 4.1A, B, part C, Figure 4.2A, B part C) the trend was less clear. In soil 2 microcosms, 1,4-dioxane removal started before cDCE removal while in soil 1 microcosms, the removal for 1,4-dioxane and cDCE started at a similar time.

The SIP experiments involved the addition of cDCE and 1,4-dioxane separately to microcosms inoculated with each soil (Supplementary Figure 4.1). Triplicate samples for each were sacrificed at 44 days for DNA extraction (~50% removal of 1,4-dioxane or cDCE). The concentration of cDCE in abiotic controls declined towards the end of the study, likely a result of gas phase leakage through the previously punctured rubber septa (Supplementary Figure 4.1 C and D).

3.2 Microbial Community Analysis

The rarefaction curves of the SIP fractions plateaued, indicating the majority of the species were sequenced (Supplementary Figure 4.2). A larger number of species were found in microcosms amended with soil 2 from all fractions (Supplementary Figure 4.2). The NMDS analysis suggested the community composition was different between the light and heavy fractions in both the cDCE (Figure 4.3A) and 1,4-dioxane (Figure 4.3B) amended microcosms, indicating a successful fractionation process. While a clear separation between the two soils was visible in the fractions originating from the cDCE amendments (Figure 4.3A), the separation was less pronounced in the 1,4-dioxane fractions (Figure 4.3B), suggesting a greater similarity in the latter samples. Alpha diversity and richness indices (Figure 4.3C and D) indicated a greater distinction between the light and heavy fractions of the cDCE amended samples compared to

those amended with 1,4-dioxane. NMDS analysis also provided clear distinctions between the communities based on the amended substrate (cDCE or 1,4-dioxane) (Supplementary Figure 4.3). The diversity and richness indices were higher in the samples amended with cDCE compared to those amended with 1,4-dioxane (Supplementary Figure 4.3).

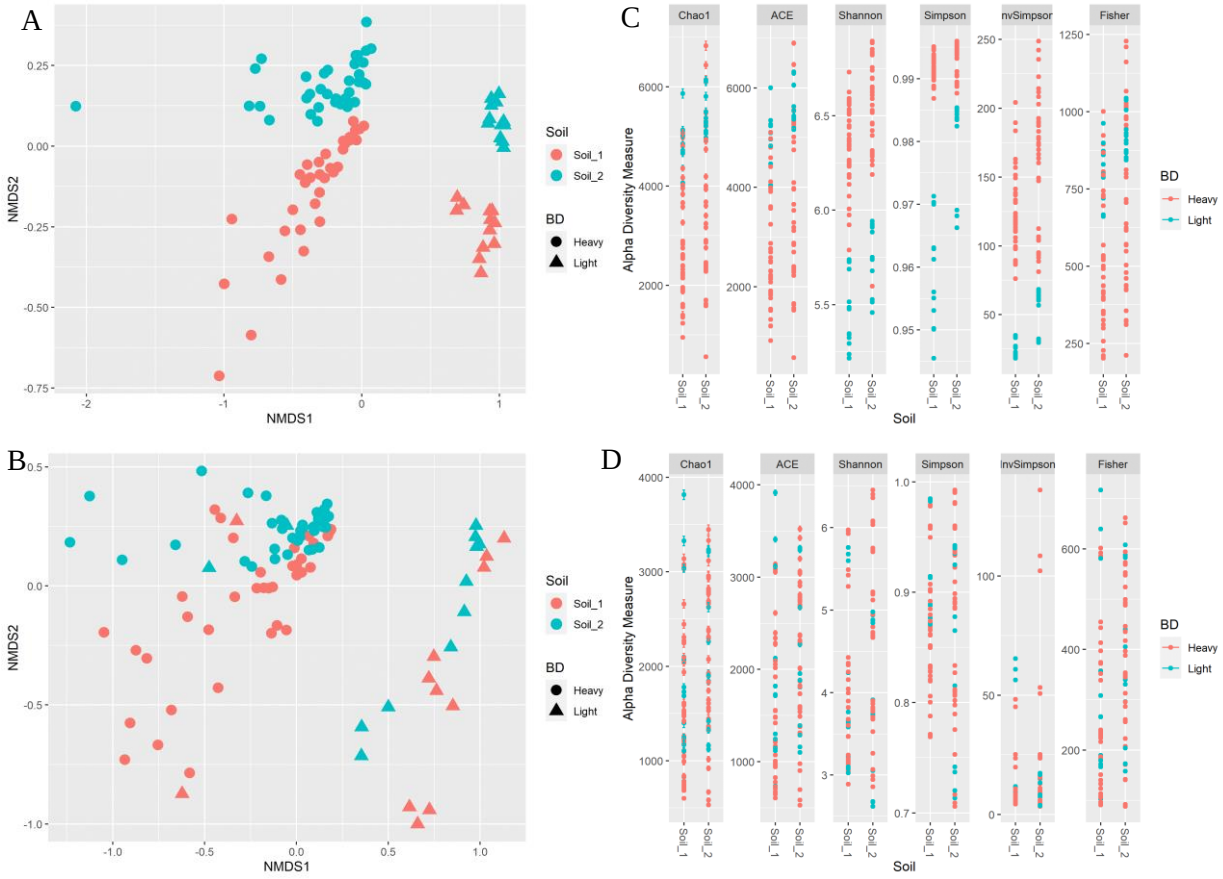


Figure 4. 3. Non-metric Multi-dimensional Scaling (NMDS) plots (A, B) and alpha diversity measurements (C, D) for the cDCE (A, C) and 1,4-dioxane (B, D) SIP experiments with soil 1 and 2.

Using the rarefied even depth of 95% of the minimum sum of OTU counts, 32 phyla were identified (Figure 4.4A). Major phyla included *Proteobacteria* and *Actinobacteria*, and these were more abundant in all heavy fractions compared to the light fractions, while *Bacteroidetes* was dominant in the light fractions. *Gemmatimonadetes* was more abundant in the heavy fractions only in the cDCE amended samples. Other major phyla included *Chloroflexi,*

Firmicutes and *Verrucomicrobia*. In many cases, a clear distinction is visible between the ^{12}C and ^{13}C amended fractions, indicating differences between communities, as discussed below.

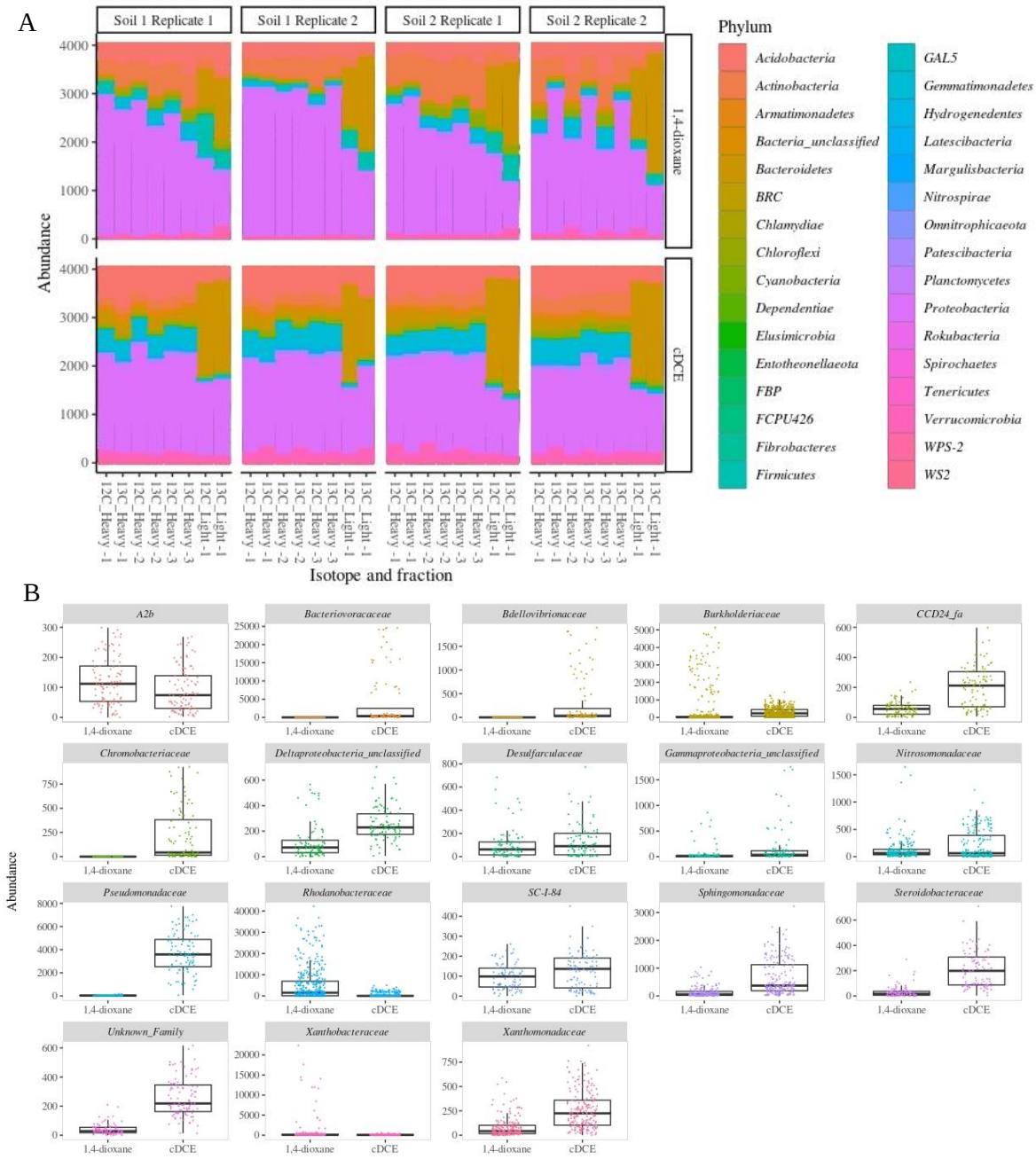


Figure 4. 4. Classification to the phylum level for both replicates and soils amended with 1,4-dioxane (upper plot) or cDCE (lower plot) with a rarefied even depth of 95% of the minimum sum of OTU counts, each column represents cumulative values for three fractions (A). The classification (family level) of the top 30 OTUs (across all samples) within the most dominant phylum (*Proteobacteria*) (B) without rarefaction.

As *Proteobacteria* represented the phylum with the greatest number of sequences, the most abundant families (top 30 OTUs) within this phylum were determined (Figure 4.4B). Those that illustrated a higher abundance in the cDCE amended samples included*: Bacteriovoracaceae, Bdellovibrionaceae, CCD24_fa, Chromobacteriaceae, Deltaproteobacteria*_unclassified, *Pseudomonadaceae, Sphingomonadaceae* and *Steroidobacteraceae*. In the 1,4-dioxane amended samples, *Rhodanobacteraceae* was more abundant and several samples illustrated a high abundance of *Burkholderiaceae* and *Xanthobacteraceae.* At the genus level for all phyla, the most abundant genera in the 1,4-dioxane amended samples classified as *Rhodanobacter*, *Chujaibacter* (both *Proteobacteria*) and an uncharacterized genus within *Bacteroidetes* (Supplementary Figure 4.4). While in the cDCE amended samples, the most abundant genera included unclassified *Bacteria*, *Pseudomonas* (*Proteobacteria*) and *Gp6* (*Actinobacteria*) (Supplementary Figure 4.4).

The current work also involved the analysis of shotgun sequencing data from the same two soils from a previous study (27). Here, multiple species previously associated with 1,4-dioxane and cDCE degradation were identified, including *Pseudonocardia dioxanivorans* (1,4-dioxane degrader) and *Polaromonas* sp. JS666 (cDCE degrader) (Supplementary Figure 4.5).

3.3 Phylotypes Responsible for ^{13}C Uptake from cDCE and 1,4-Dioxane

Phylotypes enriched in the heavy fractions of the ^{13}C cDCE or ^{13}C 1,4-dioxane amended samples compared to the controls (heavy fractions from ^{12}C cDCE or ^{12}C 1,4-dioxane amended samples) were determined using Welch's two sided t-test (within STAMP, $p < 0.05$) (Supplementary Figure 4.6 and 4.7). The dominant enriched genera for ^{13}C uptake from 1,4-dioxane included *Rhodopseudomonas* and *Rhodanobacter* (Supplementary Figure 4.6). Whereas the dominant enriched genera for ^{13}C uptake from cDCE included *Bacteriovorax*, *Pseudomonas* and

Sphingomonas (Supplementary Figure 4.7). An additional statistical test (Wilcoxon Rank, $p < 0.05$) confirmed the enrichment of *Rhodopseudomonas* and *Rhodanobacter* in one or both replicates of both soils (Figure 4.5). For cDCE, multiple genera were enriched in replicates of soil 1 and 2 (Figure 4.6). Enriched genera in soil 1 included: *Bacteriovorax*, *Bradyrhizobium* and two unclassified genera from *Blastocatellaceae*. Enriched genera in soil 2 included: *Bradyrhizobium*, *Caulobacter*, an uncultured genus within *Vicinamibacterales*, *Pseudomonas* and *Sphingomonas*. The greater diversity of dominant enriched OTUs in cDCE microcosms between soils, compared to 1,4-dioxane microcosms between soils, is consistent with the NMDS analysis indicating clear distinctions between cDCE communities between soils compared to 1,4-dioxane communities between soils (Figure 4.3).

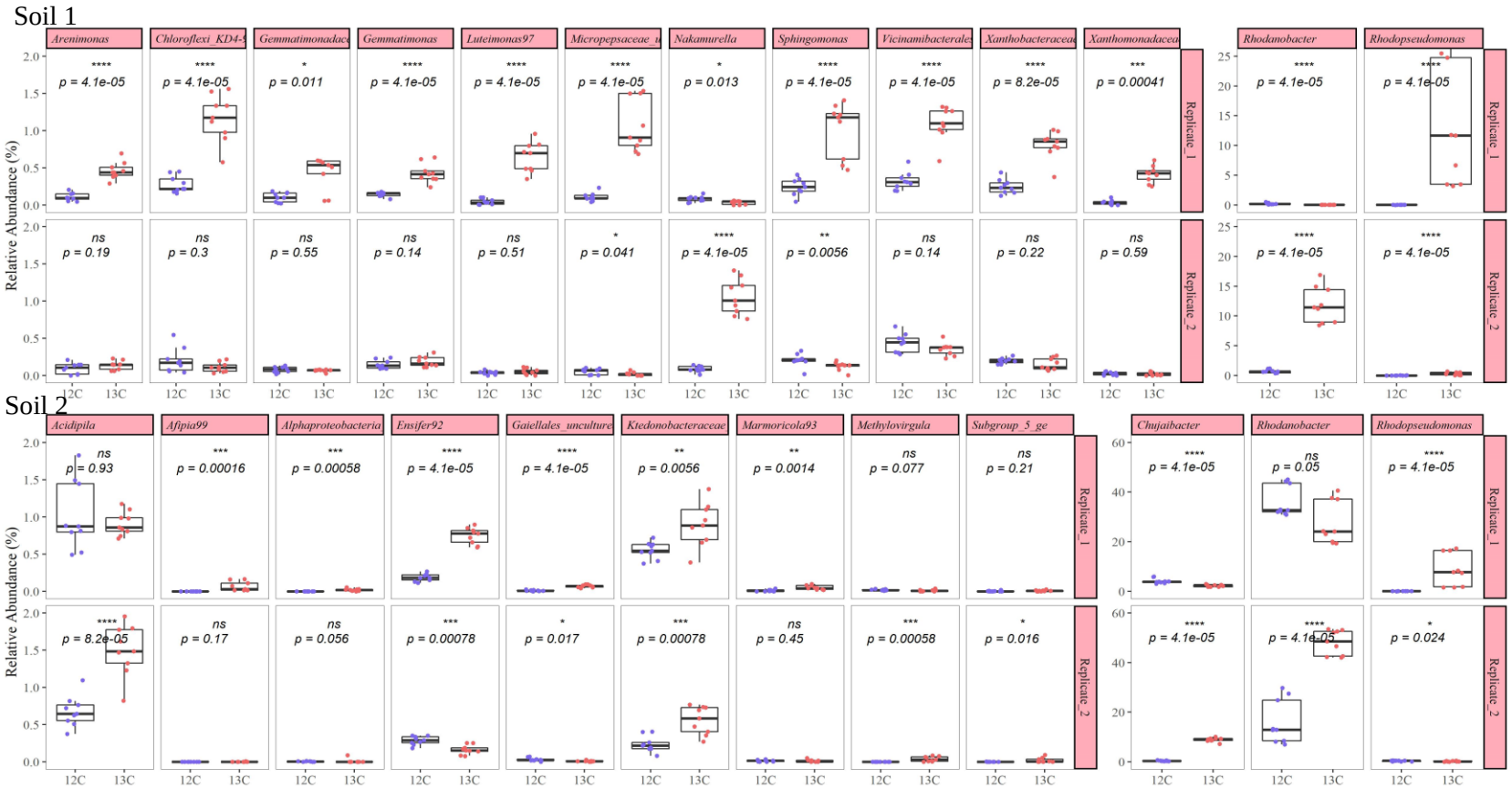


Figure 4. 5. Boxplots with Wilcoxon Rank test results between phylotypes enriched (as determined by STAMP) in 13C amended heavy fractions (red dots) compared to the 12C amended heavy fractions (purple dots) by soil (upper is Soil 1 and lower is Soil 2) and by replicate of 1,4-dioxane amended samples. The graphs on the right have a different y-axis. P values of 0.0001, 0.001, 0.01, 0.05, >0.05 are presented by ****, ***, **, *, ns.

Soil 1

Soil 2

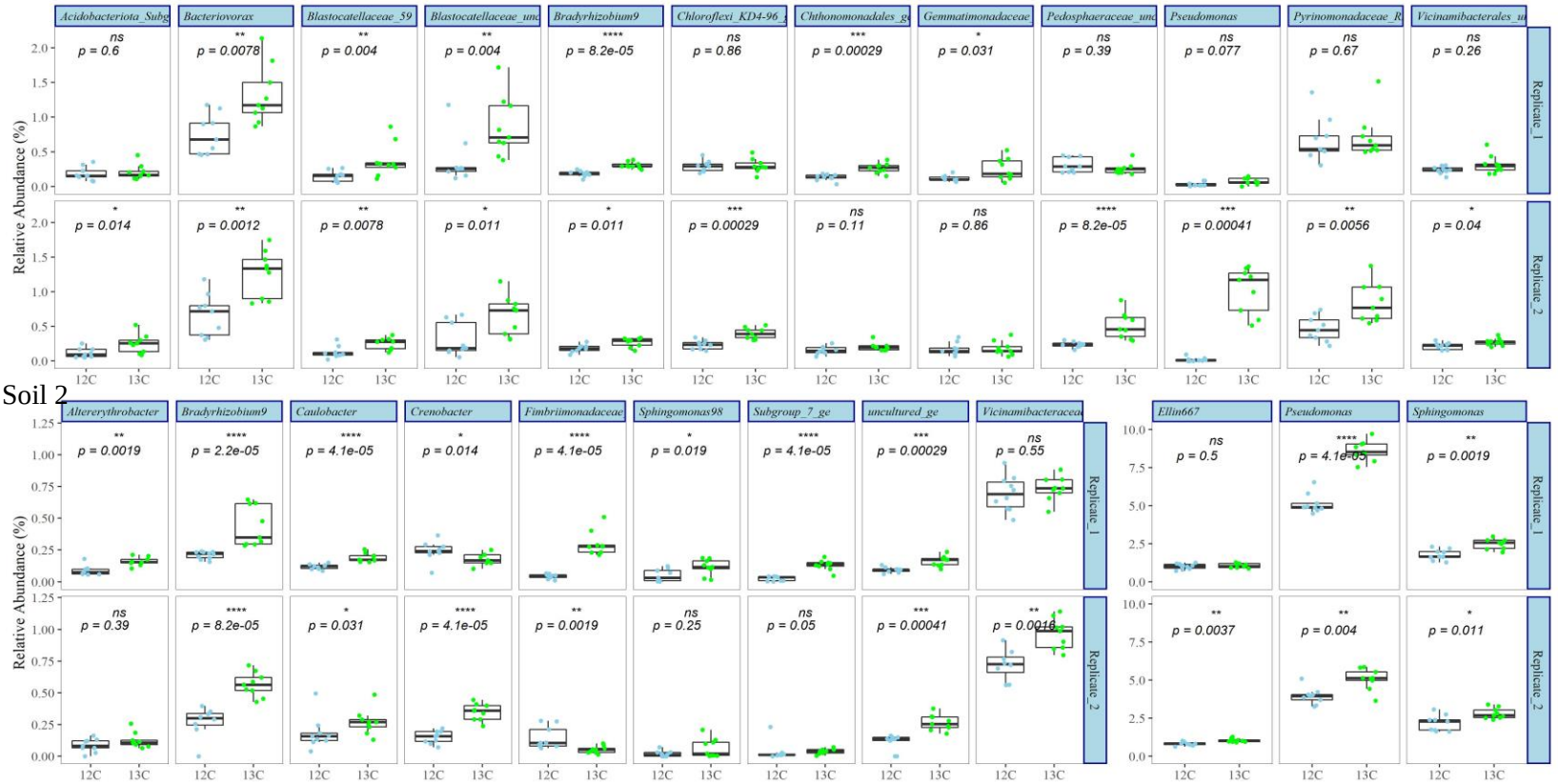


Figure 4. 6. Boxplots with Wilcoxon Rank test results between phylotypes enriched (as determined by STAMP) in 13C amended heavy fractions (green dots) compared to the 12C amended heavy fractions (blue dots) by soil (upper is Soil 1 and lower is Soil 2) and by replicate of cDCE amended samples. The graphs for Soil 2 have a different y-axis. P values of 0.0001, 0.001, 0.01, 0.05, >0.05 are presented by ****, ***, **, *, ns.

3.4 Co-Occurrence Networks

Co-occurrence networks were generated to illustrate the differences between soil 1 and soil 2 microbial communities involved in 1,4-dioxane and cDCE degradation (those enriched in ^{13}C heavy fractions, STAMP analysis) (Supplementary Figure 4.8). The OTUs with a correlation coefficient > 0.7 were connected with lines. The main genera were represented by 166 and 172 nodes (analyzed as OTUs present in at least 50% of the samples and with the abundance ≥ 0.06%) for the degradation of 1,4-dioxane and cDCE, respectively.

For the microbial communities associated with 1,4-dioxane biodegradation, 67 OTUs showed a significant difference between soil 1 and 2 (26 and 41 were more abundant in soil 1 and soil 2, respectively) (Supplementary Figure 4.8 A). In contrast, more OTUs (99) illustrated a significant difference between the soil 1 and soil 2 microbial communities for those involved in cDCE biodegradation (44 and 45 OTUs were more abundant in soil 1 and 2, and these OTUs were clearly separated) (Supplementary Figure 4.8 B).

The network also illustrated the relationships between the enriched and other abundant OTUs. The enriched OTUs (by STAMP analysis) displayed on the networks are summarized (Supplementary Table 4.2). In the microbial communities associated with 1,4-dioxane biodegradation, the majority of OTUs classified as *Actinobacteria*, *Proteobacteria* and *Acidobacteria* (Supplementary Figure 4.9A). *Actinobacteria* and *Proteobacteria* were related to each other. A total of 58 and 8 enriched OTUs were displayed in soil 1 and 2, respectively. Most of the enriched OTUs were *Actinobacteria* and *Proteobacteria* (Supplementary Figure 4.9B).

In the microbial communities associated with cDCE biodegradation, the majority of the OTUs were *Proteobacteria*, *Bacteroidetes*, *Acidobacteria* and *Gemmatimonadetes* (Supplementary Figure 4.10A). *Proteobacteria* connected more with *Gemmatimonadetes.* A total

of 26 and 16 enriched OTUs were displayed in soil 1 and 2, respectively. The enriched OTUs were *Bacteroidetes, Acidobacteria* in soil 1 while the enriched OTUs were *Proteobacteria* in soil 2 (Supplementary Figure 4.10B).

3.5 Predicted Functions and Correlations with OTUs

In the current work, PICRUSt2 predicted KO functions formerly associated with 1,4-dioxane biodegradation in the 1,4-dioxane amended microcosms included toluene monooxygenase, propane monooxygenase (most abundant) and methane monooxygenase (10, 11) (Supplementary Figure 4.11). In the cDCE amended microcosms, the abundant function associated with cDCE included glutathione S-transferase (51, 59). For 1,4-dioxane, correlations between gene and phylotype abundance indicated propane monooxygenase positively correlated with *Rokubacteriales*, *KD4-96* (*Chloroflexi*), Gitt-GS-36 (*Chloroflexi*) and uncultured genera from *Vicinamibacterales* and *Gemmatimonadaceae* (Figure 4.7A). For cDCE, glutathione S-transferase was positively correlated with *BIrii4* and an unclassified genus from *Xanthomonadales* (Figure 4.7B).

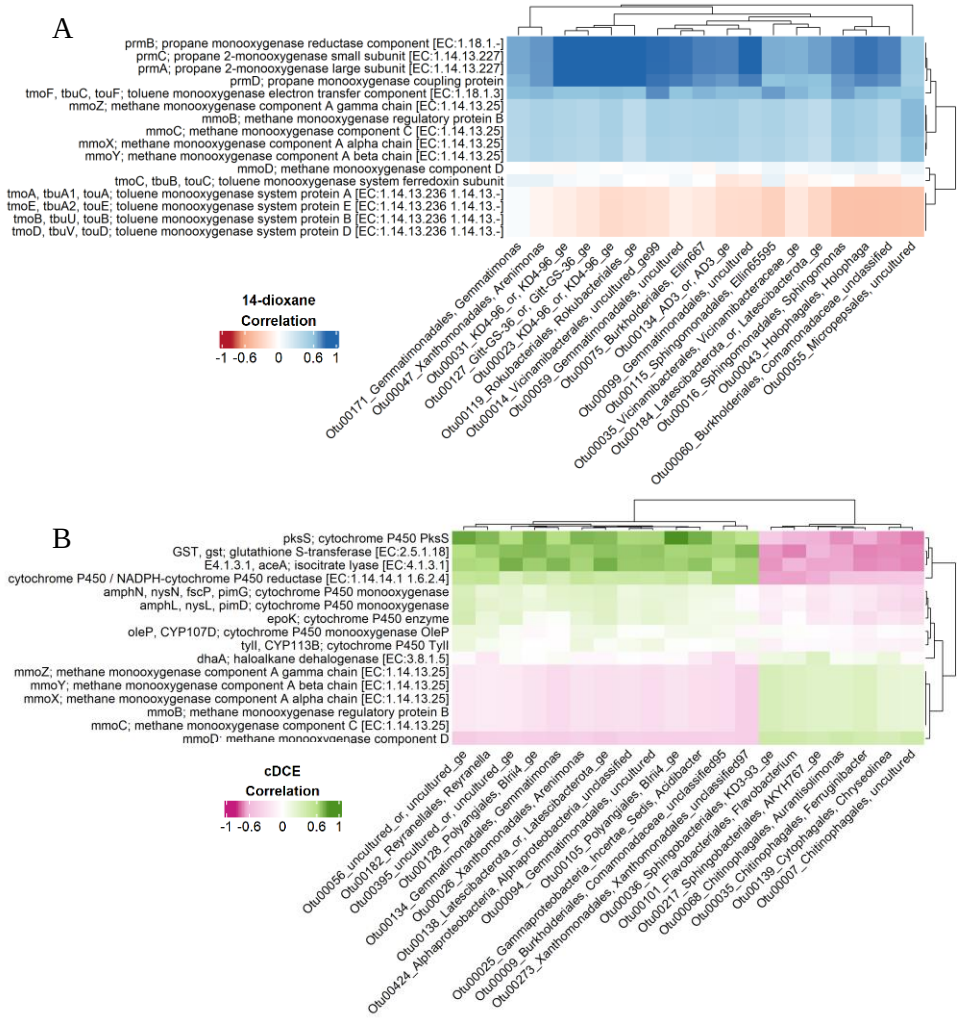


Figure 4. 7. Correlation of KO functions associated with degradation and OTUs with average abundance higher than 0.05% from 1,4-dioxane (A) and cDCE in SIP tests. 18 OTUs had an absolute correlation coefficient high than 0.59 with at least 4 of function in 1,4-dioxane SIP. 20 OTUs had an absolute correlation coefficient high than 0.6 with at least 2 of function in 1,4-dioxane SIP.

4. Discussion

There have been many reports of the common occurrence of TCE and 1,4-dioxane at contaminated sites (8, 9, 18). As TCE is reduced to cDCE by both indigenous microbial communities or dechlorinating cultures (3, 16, 18, 26) and "cDCE-stall" is common at contaminated sites the removal of this metabolite is also of concern. Previously, we reported 1,4-dioxane biodegradation in the two soils examined here (19). Here, we build on that research by

investigating the potential for the concurrent biodegradation of both cDCE and 1,4-dioxane. Further, the microorganisms involved in the uptake of carbon from each chemical were identified using DNA-based SIP. Additionally, the functional genes involved in the degradation of cDCE and 1,4-dioxane were predicted using PICRUSt2 (48) and their abundance was correlated to OTUs present.

The impact of additional treatments on 1,4-dioxane biodegradation differed between the two soils. For soil 1, when all three substrates were present (lactate, cDCE and 1,4-dioxane), 1,4-dioxane removal was slower. However, in soil 2, 1,4-dioxane biodegradation was similar for all three treatments, indicating the presence of either lactate or cDCE or both did not impact 1,4-dioxane removal in this soil community. Inhibition of 1,4-dioxane biodegradation by additional substrates has been noted by others. For example, when propane was added to *Azoarcus* sp. DD4, co-metabolism of 1,4-dioxane was delayed and followed the co-metabolism of chloroethenes (1,1-dichlorothene, VC and cDCE) (17, 18). Further, research with *Pseudonocardia dioxanivorans* CB1190 indicated from four common co-contaminants (1,1-DCE, 1,1,1-TCA, cDCE, TCE), cDCE was the second most inhibitory chemical to aerobic 1,4-dioxane degradation (13).

The most notable trend for cDCE removal was the positive impact of lactate. Also, when lactate was present, 1,4-dioxane did not impact cDCE removal and decreases in cDCE concentrations occurred earlier than 1,4-dioxane decreases in both soil microcosms. In comparison, when the microcosms were amended with only 1,4-dioxane and cDCE the trend was less clear. In soil 2 microcosms, 1,4-dioxane removal started before cDCE removal while in soil 1 microcosms, the removal for 1,4-dioxane and cDCE started at a similar time.

Unlike previous studies which have primarily involved the biodegradation of co-contaminants by isolates or commercially available mixed communities (16-18), this work examined contaminant biodegradation by indigenous microbial communities. The NMDS analysis indicated cDCE produced a clear difference between the microbial communities of the two soils. The difference between the two microbial communities was less distinct for the soil microcosms amended with 1,4-dioxane. These trends could suggest cDCE is more important for impacting microbial community structure, perhaps through inhibition or as a beneficial substrate. Interestingly, the microbial richness and diversity levels were also higher in the cDCE amended samples. Also, the enriched phylotypes illustrated greater differences between soils in the cDCE amended microcosms, compared to the 1,4-dioxane amended microcosms.

To date, many aerobic 1,4-dioxane and cDCE degraders have been identified (Supplementary Table 4.1). In the current study, multiple 1,4-dioxane and cDCE degraders were detected in shotgun sequencing data from samples inoculated with both soils. To determine if these species were actively involved in biodegradation, SIP was employed to determine which microorganisms were responsible for carbon uptake from each chemical. From the twelve 1,4-dioxane degrading phylotypes identified by shotgun sequencing, only two were associated with carbon uptake from 1,4-dioxane. Specifically, the previously reported 1,4-dioxane degraders *Rhodanobacter* sp. and *Xanthobacter flavus* were detected via shotgun sequencing and OTUs classifying as *Rhodanobacter* and the family *Xanthobacteraceae* were detected via SIP. From the nineteen cDCE degrading phylotypes detected via shotgun sequencing only one genus (*Pseudomonas*) was detected via SIP. These results provide support for the importance of SIP, over sequencing alone, for connecting identity with function.

In the current work, many genera were enriched during the SIP experiments, suggesting a wide range of microorganisms were assimilating carbon from the biodegradation of 1,4-dioxane or cDCE. Significantly enriched genera from the biodegradation of ^{13}C-1,4-dioxane included *Rhodopseudomonas* and *Rhodanobacter*. Consistent with the current study, *Rhodopseudomonas* was previously associated with the incorporation of ^{13}C from 1,4-dioxane in aerobic experiments with activated sludge, and its abundance increased with the degradation of 1,4-dioxane in both batch tests and a full-scale treatment system (22). *Rhodanobacter* was also reported as a metabolizer for 1,4-dioxane, with the addition of tetrahydrofuran accelerating 1,4-dioxane degradation (60). Combined with the results from the current work, these studies indicate the importance of both genera for 1,4-dioxane biodegradation and future work should examine their occurrence and activity at 1,4-dioxane contaminated sites.

The other SIP identified genera during 1,4-dioxane degradation illustrated lower relative abundance levels compared to *Rhodopseudomonas* and *Rhodanobacter.* An OTU classifying as *Afipia* was enriched in microcosms inoculated with soil 2. Similarly, others have linked this genus (*Afipia* sp. D1) to the assimilation of carbon from 1,4-dioxane (61). *Afipia* was also abundant in uncontaminated soil microcosms during 1,4-dioxane degradation (62, 63). Here, an unclassified genus from *Xanthobacteraceae* was associated with carbon uptake from 1,4-dioxane in soil 1 microcosms. This family includes the 1,4-dioxane degraders *Xanthobacter* sp. YN2 (64) and *Xanthobacter flavus* DT8 (65). *Xanthobacteraceae* significantly increased in 1,4-dioxane degradation tests with domestic wastewater activated sludge, and a novel 1,4-dioxane-hydroxylating monooxygenase was identified from *Xanthobacter* strains (66). Another enriched OTU from the 1,4-dioxane SIP study classified within the family *Xanthomonadaceae*. This family was previously linked to 1,4-dioxane degradation in activated sludge from a full-scale

bioreactor for landfill leachate treatment (67, 68). In other 1,4-dioxane degradation studies, dominant or enriched genera included *Chryseobacterium*, *Dokdonella*, *Pseudonocardia*, *Bradyrhizobium*, *Mycobacterium*, *Nocardioides*, and *Kribbella* (19, 69), however these genera were not identified via SIP in the current work.

Dominant genera significantly enriched in the biodegradation of ^{13}C-cDCE in either or both soil microcosms and replicates, included *Bacteriovorax*, *Pseudomonas* and *Sphingomonas*. The dominance of *Pseudomonas* is consistent with previous studies associating this genus with cDCE degradation (Supplementary Table 4.1). Although isolates from the genus *Bacteriovorax* have not been previously linked with cDCE biodegradation, this genus has previously been associated with hydrocarbon biodegradation (70, 71).

In the current work, two enriched genera (*Sphingomonas* and *Bradyrhizobium*) during ^{13}C-cDCE degradation were abundant or enriched in previous 1,4-dioxane degradation studies (14, 15, 68). *Sphingomonas* was dominant during 1,4-dioxane degradation by *P. dioxanivorans* CB1190 when residuals of chlorinated volatile organic compounds (including cDCE) were present (14, 15). However, in the current study, these genera were not associated with ^{13}C uptake from 1,4-dioxane. Interestingly, there were multiple novel genera, with no previous links to cDCE or 1,4-dioxane, identified in the current study as carbon assimilators.

Multiple functions for 1,4-dioxane and cDCE biodegradation were predicted in the soil microcosms using PICRUSt2 (48). The most abundant function for 1,4-dioxane biodegradation was propane monooxygenase. Many OTUs positively correlated with propane monooxygenase, including for example *KD4-96* (*Chloroflexi*), an uncultured genus from the class of *Vicinamibacterales* and from the family of *Gemmatimonadaceae*. The high abundance of propane monooxygenase is consistent with previous work describing the dominance of propane

monooxygenase from *Rhodococcus jostii* RHA1 and *Rhodococcus* sp. RR1 in 1,4-dioxane degrading microcosms inoculated with these soils (19). Several other previously identified 1,4-dioxane degrading enzymes were predicted to be present in the soil microcosms, however, additional research is needed to confirm if these enzymes are active.

The functional genes associated with aerobic cDCE degradation include cytochrome P450 monooxygenase and glutathione S-transferase *Polaromonas* sp. strain JS666 (50, 51, 59). In the current study, both biomarkers correlated with a number of OTUs, but not *Polaromonas* strain JS666. However, these OTUs were not enriched in the SIP experiments, suggesting other enzymes may be involved or other methods (beyond the predictions provided by PICRUSt) are needed to obtain such information.

In summary, this work demonstrated the concurrent removal of cDCE and 1,4-dioxane by indigenous soil microbial communities and the enhancement of cDCE removal by lactate. Through the use of SIP, multiple genera, both previously identified and not previously identified degraders, were enriched and benefited from the degradation of 1,4-dioxane and cDCE. In addition, a wide range of genes involved in the degradation were predicted to be associated with contaminant removal. These genera and genes were more diverse than previously reported. The extraction of DNA at only one time point during biodegradation is a potential limitation of the current study. Further, it is unknown if the enriched genera participated in carbon uptake from 1,4-dioxane and cDCE, or from their metabolites. Combining the current research with more quantitative approaches (e.g. qPCR, RNA-Seq) would enhance the information gained from the functional gene analysis. The data generated in the current study has the potential to be incorporated into diagnostic tests for assessing biodegradation potential at contaminated sites, for example, quantification of *Rhodopseudomonas* and *Rhodanobacter* at 1,4-dioxane contaminated

sites. Although the results suggest aerobic concurrent biodegradation of cDCE and 1,4-dioxane should be considered for chlorinated solvent site remediation, additional research is needed to determine if appropriately low contaminant concentrations can be reached.

Acknowledgements

Thanks to Stacey VanderWulp from MSU for providing the soil samples from KBS LTER. This work was partially supported by NSF grant number 1902250.

APPENDIX

APPENDIX

Supplementary Table 4. 1. Identified 1,4-dioxane and DCE degraders with the lowest rank name and taxonomy ID from NCBI.

Strains/species for 1,4-dioxane degradation	Species/strain name in NCBI	NCBI rank	NCBI taxID	Number of subtrees	Reference
Pseudonocardia dioxanivorans CB1190	*Pseudonocardia dioxanivorans CB1190*	strain	675635	0	(10, 11, 28)
Rhodococcus ruber 219	*Rhodococcus ruber*	species	1830	4	(72)
Pseudonocardia benzennivorans B5	*Pseudonocardia benzenivorans*	species	228005	0	(10, 73)
Mycobacterium sp. PH-06	*Mycobacterium dioxanotrophicus*	species	482462	0	(11, 74)
Afipia sp. D1	*Afipia sp. D1*	species	882658	0	(61)
Mycobacterium sp. D6	*Mycobacterium sp. D6*	species	882659	0	(61)
Mycobacterium sp. D11	*Mycobacterium sp. D11*	species	882660	0	(61)
Pseudonocardia sp. D17	*Pseudonocardia sp. D17*	species	882661	0	(61)
Acinetobacter baumannii DD1	*Acinetobacter baumannii*	species	470	1003	(75)
Rhodanobacter AYS5	*Rhodanobacter sp.*	species	1883446	0	(60)
Xanthobacter flavus DT8	*Xanthobacter flavus*	species	281	0	(65)
Rhodococcus aetherivorans JCM 14343	*Rhodococcus aetherivorans*	species	191292	1	(76)
Pseudonocardia tetrahydrofuranoxydans sp. K1	*Pseudonocardia tetrahydrofuranoxydans*	species	102884	1	(11, 77)
Pseudonocardia sp. ENV478	*Pseudonocardia sp. ENV478*	species	377619	0	(11, 78)
Rhodococcus ruber T1	*Rhodococcus ruber*	species	1830	4	(61)
Rhodococcus ruber T5	*Rhodococcus ruber*	species	1830	4	(61)
Rhodococcus ruber ENV 425	*Rhodococcus ruber*	species	1830	4	(79)
Rhodococcus RR1	*Rhodococcus sp. RR1*	species	402393	0	(10, 11, 80)
Flavobacterium sp.	*Flavobacterium sp.*	species	239	0 [A]	(81)
Mycobacterium vaccae	*Mycolicibacterium vaccae*	species	1810	3	(82)
Mycobacterium sp. ENV 421	*Mycobacterium sp. ENV421*	species	1213407	0	(11, 83)
Pseudomonas mendocina KR1	*Pseudomonas mendocina*	species	300	7	(10, 11, 84)
Ralstonia pickettii PKO1	*Ralstonia pickettii*	species	329	5	(10, 85)
Burkholderia cepacia G4	*Burkholderia cepacia*	species	292	5	(10, 11, 86)
Methylosinus trichosporium OB3b	*Methylosinus trichosporium OB3b*	strain	595536	0	(10, 11, 87)
Pseudonocardia acacia JCM	*Pseudonocardia acaciae*	species	551276	1	(76)
Pseudonocardia asaccharolytica JCM	*Pseudonocardia asaccharolytica*	species	54010	1	(76)
Pseudomonas pickettii PKO1	*Ralstonia pickettii*	species	329	5	(85)
Rhodococcus sp. YYL	*Rhodococcus sp. YYL*	species	423618	0	(11, 88)
Rhodococcus josti RHA1	*Rhodococcus jostii RHA1*	strain	101510	0	(11, 89)
Pseudonocardia K1	*Pseudonocardia sp.*	species	60912	0 [A]	(10, 90)
Mycobacterium vaccae JOB5	*Mycolicibacterium vaccae*	species	1810	3	(10, 91)
Rhodococcus rhodochrous ATCC 21198	*Rhodococcus rhodochrous ATCC 21198*	strain	1429046	0	(92)

Strains/species for aerobic DCE degradation	Species/strain name in NCBI	NCBI rank	NCBI taxID	Number of subtrees	Reference
Methylosinus trichosporium OB3b	*Methylosinus trichosporium OB3b*	strain	595536	0	(93)
Rhodococcus rhodochrous ATCC 21198	*Rhodococcus rhodochrous ATCC 21198*	strain	1429046	0	(92)
Xanthobacter autotrophicus	*Xanthobacter autotrophicus*	species	280	1	(94)
Ralstonia pickettii PKO1	*Ralstonia pickettii*	species	329	5	(94)
Cupriavidus necator JMP134	*Cupriavidus pinatubonensis JMP134*	strain	264198	0	(94)
Burkholderia vietnamenisis G4	*Burkholderia vietnamiensis G4*	strain	269482	0	(94)
Polaromonas chloroethenica JS666	*Polaromonas sp. JS666*	species	296591	0	(94)
Methylococcus capsulatus Bath	*Methylococcus capsulatus str. Bath*	strain	243233	0	(94)
Pseduomonas stutzeri OX1	*Pseudomonas stutzeri*	species	316	19	(94)
Nocardioides CF8	*Nocardioides sp. CF8*	species	110319	0	(94)
Rhodococcus globerulus AD45	*Rhodococcus globerulus*	species	33008	1	(94)
Gordonia rubripertinca B-276	*Gordonia rubripertincta*	species	36822	1	(94)
Mycobacterium chubuense NBB4	*Mycolicibacterium chubuense NBB4*	strain	710421	0	(94)
Rhodococcus sp.	*Rhodococcus sp.*	species	1831	0 [A]	(95)
Ralstonia sp.	*Ralstonia sp.*	species	54061	0 [A]	(95)
Variovorax sp.	*Variovorax sp.*	species	1871043	0 [A]	(95)
Comamonas testosteroni RF2	*Comamonas testosteroni*	species	285	12	(96)
Bacillus sp.	*Bacillus sp.*	species	1409	0 [A]	(97)
Pseudomonas sp. OX1	*Pseudomonas sp. OX1*	species	320855	0	(98)
Ralstonia sp. TRW-1	*Ralstonia sp.*	species	54061	0 [A]	(99)
Pseudomonas sp. YKD221	*Pseudomonas sp.*	species	306	0 [A]	(100)
Rhodococcus sp. Strain AD45	*Rhodococcus sp. AD45*	species	103808	0	(101)
Pseudomonas plecoglossicida	*Pseudomonas plecoglossicida*	species	70775	2	(102)
Methylocystis sp. strain M	*Methylocystis sp. M*	species	51782	0	(103)
Mycobacterium sp. strain TRW-2	*Mycobacterium sp.*	species	1785	0	(103)
Mycobacterium vaccae strain JOB5	*Mycolicibacterium vaccae*	species	1810	3	(103)
Pseudomonas sp. strain JR1	*Pseudomonas sp. JR1*	species	47160	0	(103)
Pseudomonas butanavora	*Thauera butanivorans*	species	86174	1	(103)
Pseudomonas putida strain F1	*Pseudomonas putida F1*	strain	351746	0	(103)
Rhodococcus sp. strain PB1	*Rhodococcus sp.*	species	1831	0 [A]	(103)

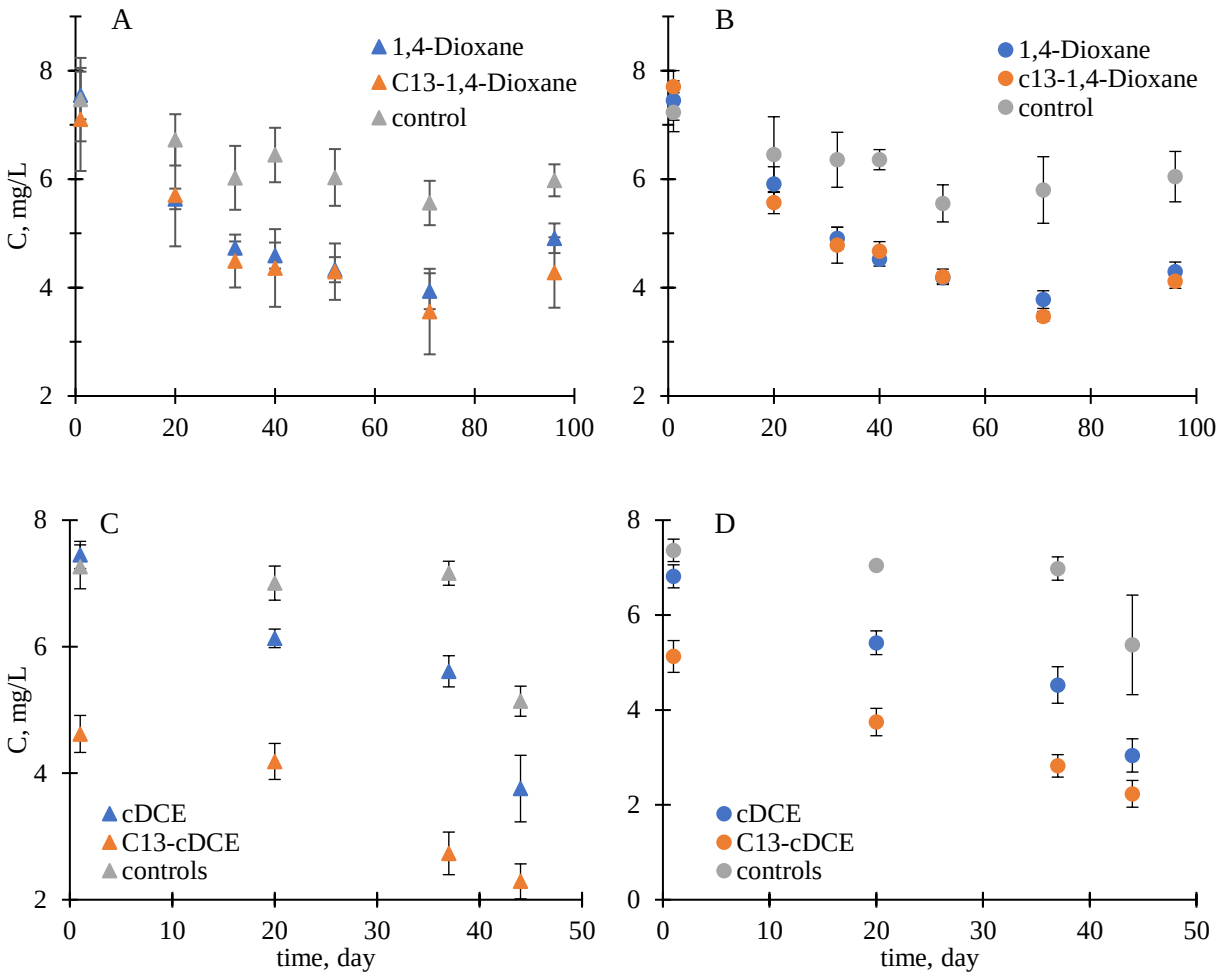
Supplementary Table 4. 1.(continued)

Rhodococcus erythropolis strain BD1	*Rhodococcus erythropolis*	species	1833	10	(103)
Xanthobacter sp. strain Py2	*Xanthobacter autotrophicus Py2*	strain	78245	0	(103)

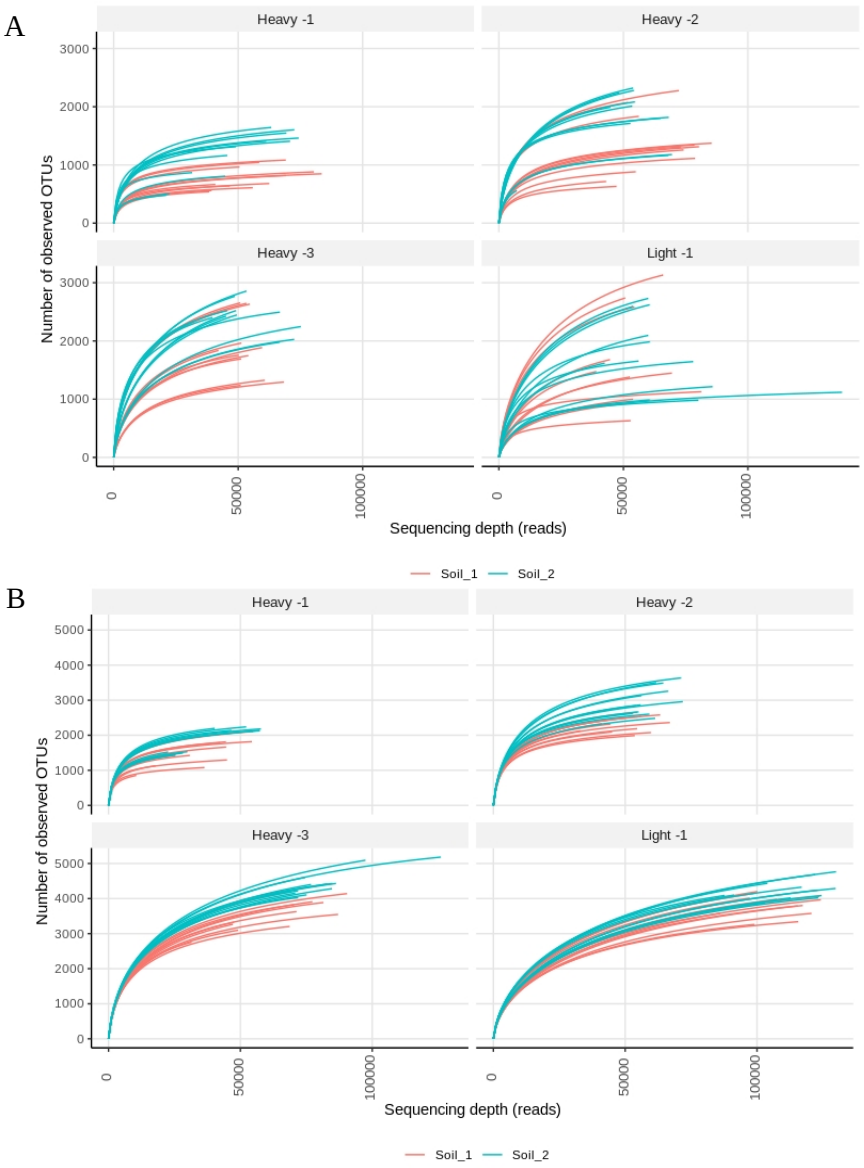
A: The identified strain name from the paper could not be searched in NCBI taxonomy browser, for example: *Pseudonocardia K1 was assigned to Pseudonocardia sp.* which belonged to unclassified *Pseudonocardia*.

Supplementary Table 4. 2. Enriched OTUs captured by the co-occurrence network. These OTUs were enriched in heavy fractions of ^{13}C labeled chemicals amended samples determined by STAMP.

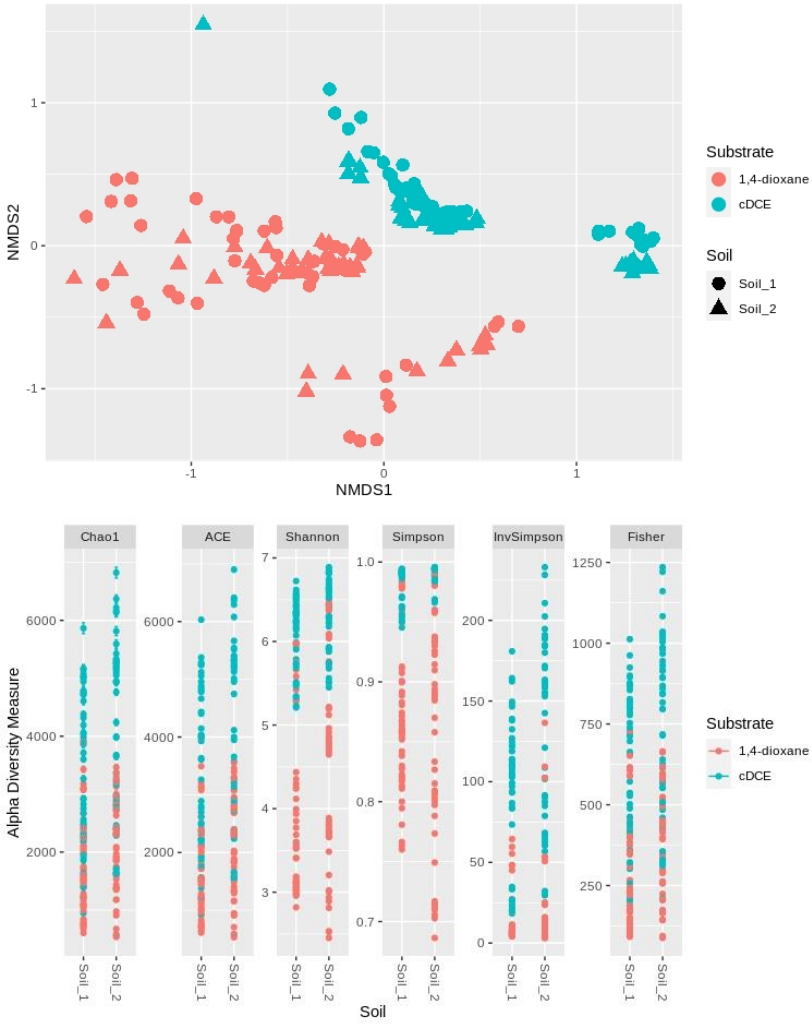
1,4-dioxane degradation soil 1	**1,4-dioxane degradation soil 2**	**cDCE degradation soil 1**	**cDCE degradation soil 1**
Otu00004_Rhodopseudomonas	*Otu00001_Rhodanobacter*	*Otu00001_Bacteriovorax*	*Otu00002_Pseudomonas*
Otu00010_Rhodanobacter	*Otu00004_Rhodopseudomonas*	*Otu00012_RB4*	*Otu00005_Sphingomonas*
Otu00014_uncultured_ge99	*Otu00003_Chujaibacter*	*Otu00013_Blastocatellaceae_unclassified*	*Otu00014_Ellin667*
Otu00016_Sphingomonas	*Otu00006_Chujaibacter*	*Otu00027_uncultured*	*Otu00020_Vicinamibacteraceae_ge*
Otu00019_Xanthobacteraceae_unclassified7	*Otu00008_Acidipila*	*Otu00031_Flavobacterium*	*Otu00022_Luteimonas98*
	Otu00013_Ktedonobacteraceae_unclassified		
Otu00022_Nitrosospira	*Otu00026_Ensifer92*	*Otu00041_37-3_ge*	*Otu00057_Bradyrhizobium9*
Otu00023_KD4-96_ge	*Otu00126_Chujaibacter*	*Otu00045_KD4-96_ge*	*Otu00070_Adhaeribacter*
Otu00025_Candidatus_Udaeobacter		*Otu00057_Bradyrhizobium9*	*Otu00081_SC-I-84_ge*
Otu00031_KD4-96_ge		*Otu00060_Blastocatellaceae_unclassified6*	*Otu00095_Crenobacter*
Otu00035_Vicinamibacteraceae_ge		*Otu00061_Blastocatellaceae_unclassified59*	*Otu00114_Fimbriimonadaceae_ge*
Otu00036_Pedobacter		*Otu00074_uncultured*	*Otu00136_Dechloromonas86*
		Otu00075_Chitinophagaceae_unclassified79	
Otu00039_Bacteria_unclassified			*Otu00137_Polaromonas63*
Otu00041_uncultured		*Otu00076_Vicinamibacteraceae_ge*	*Otu00149_Caulobacter*
Otu00043_Holophaga		*Otu00096_uncultured_ge*	*Otu00169_uncultured*
Otu00044_Blastocatellaceae_unclassified		*Otu00116_uncultured95*	*Otu00185_Novosphingobium97*
Otu00045_uncultured		*Otu00122_Subgroup__ge*	*Otu00199_Altererythrobacter*
			Otu00200_Sphingobacteriaceae_unclassified
Otu00046_uncultured_ge		*Otu00140_uncultured_ge*	
Otu00047_Arenimonas		*Otu00142_NS-2_marine_group_ge*	
Otu00049_KD4-96_ge		*Otu00143_Blastocatellaceae_unclassified8*	
Otu00050_uncultured		*Otu00148_Pseudomonas*	
Otu00052_Subgroup_7_ge		*Otu00160_MND*	
Otu00055_uncultured		*Otu00161_Chthonomonadales_ge*	
Otu00058_Nakamurella		*Otu00174_Chitinophaga*	
		Otu00179_Nitrosomonadaceae_unclassified98	
Otu00059_uncultured			
Otu00061_KD4-96_ge		*Otu00186_Sphingomonas98*	
Otu00063_Gemmatimonas		*Otu00191_Gemmatimonas*	
Otu00068_uncultured		*Otu00197_Ellin57*	
Otu00070_Alicyclobacillus		*Otu00233_uncultured99*	
Otu00075_Ellin667			
Otu00082_Pseudolabrys96			
Otu00088_Microbacteriaceae_unclassified9			
Otu00089_Luteimonas97			
Otu00095_Gammaproteobacteria_unclassified			
Otu00097_uncultured			
Otu00099_uncultured			
Otu00115_Ellin65595			
Otu00116_MND			
Otu00119_Rokubacteriales_ge			
Otu00122_uncultured			
Otu00124_Haliangium			
Otu00125_Luedemannella6			
Otu00129_uncultured_ge			
Otu00130_Candidatus_Solibacter			
Otu00132_Nitrospira			
Otu00134_AD3_ge			
Otu00138_Phenylobacterium			
Otu00140_Subgroup_7_ge			
Otu00141_Acidibacter			
Otu00142_uncultured_ge99			
Otu00148_uncultured_ge			
Otu00149_uncultured_ge54			
Otu00150_WPS-2_ge			
Otu00151_MB-A2-8_ge			
Otu00152_uncultured			
Otu00153_Xanthomonadaceae_unclassified			
Otu00154_Elsterales_unclassified			
Otu00164_uncultured_ge			
Otu00169_Alcaligenaceae_unclassified			
Otu00171_Gemmatimonas			
Otu00173_MB-A2-8_ge			



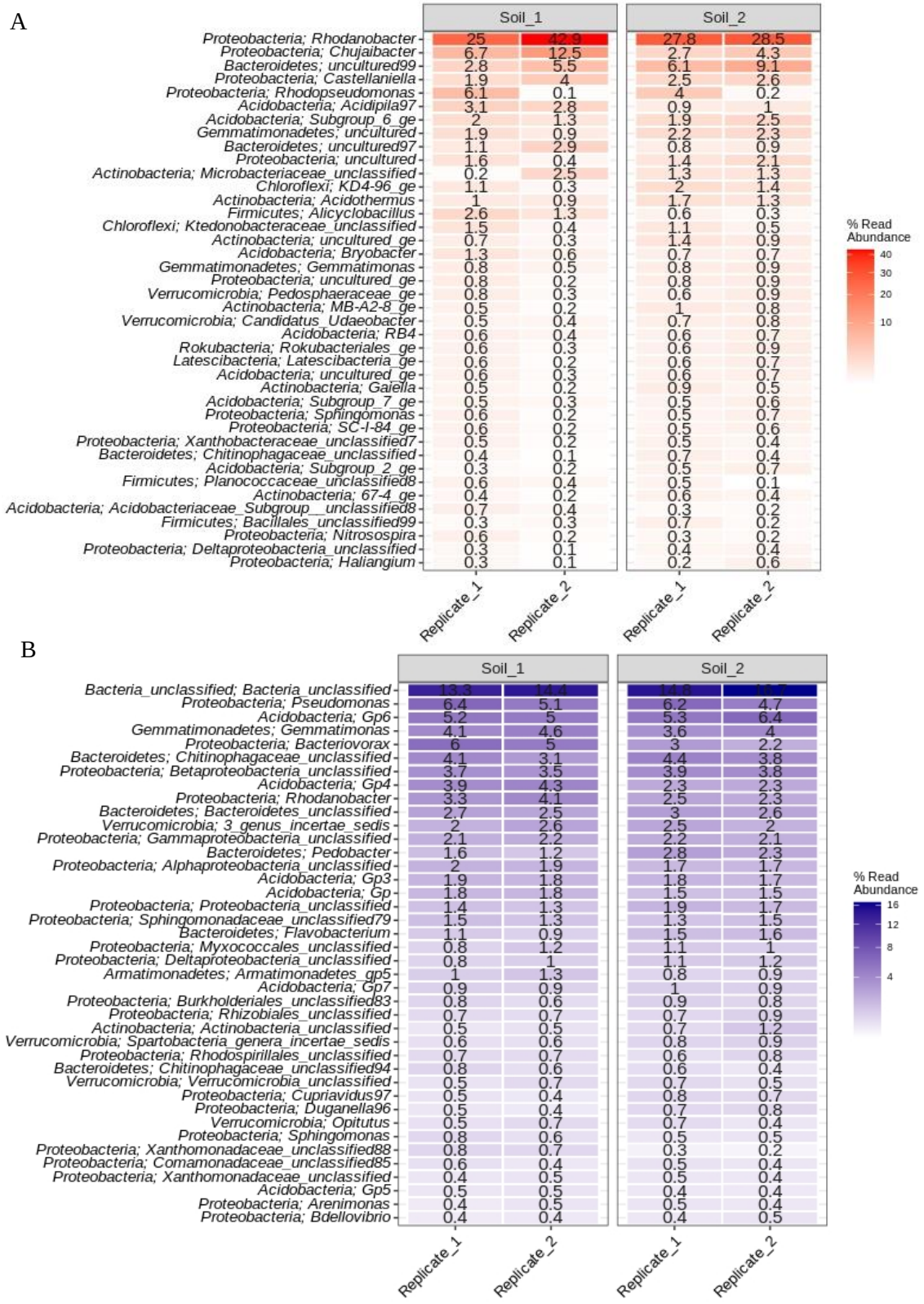
Supplementary Figure 4. 1. Average concentration of labeled and unlabeled 1,4-dioxane (A, B), and cDCE (C, D) in triplicate sample microcosms and triplicate abiotic controls inoculated with soil 1 (A, C) and 2 (B, D).



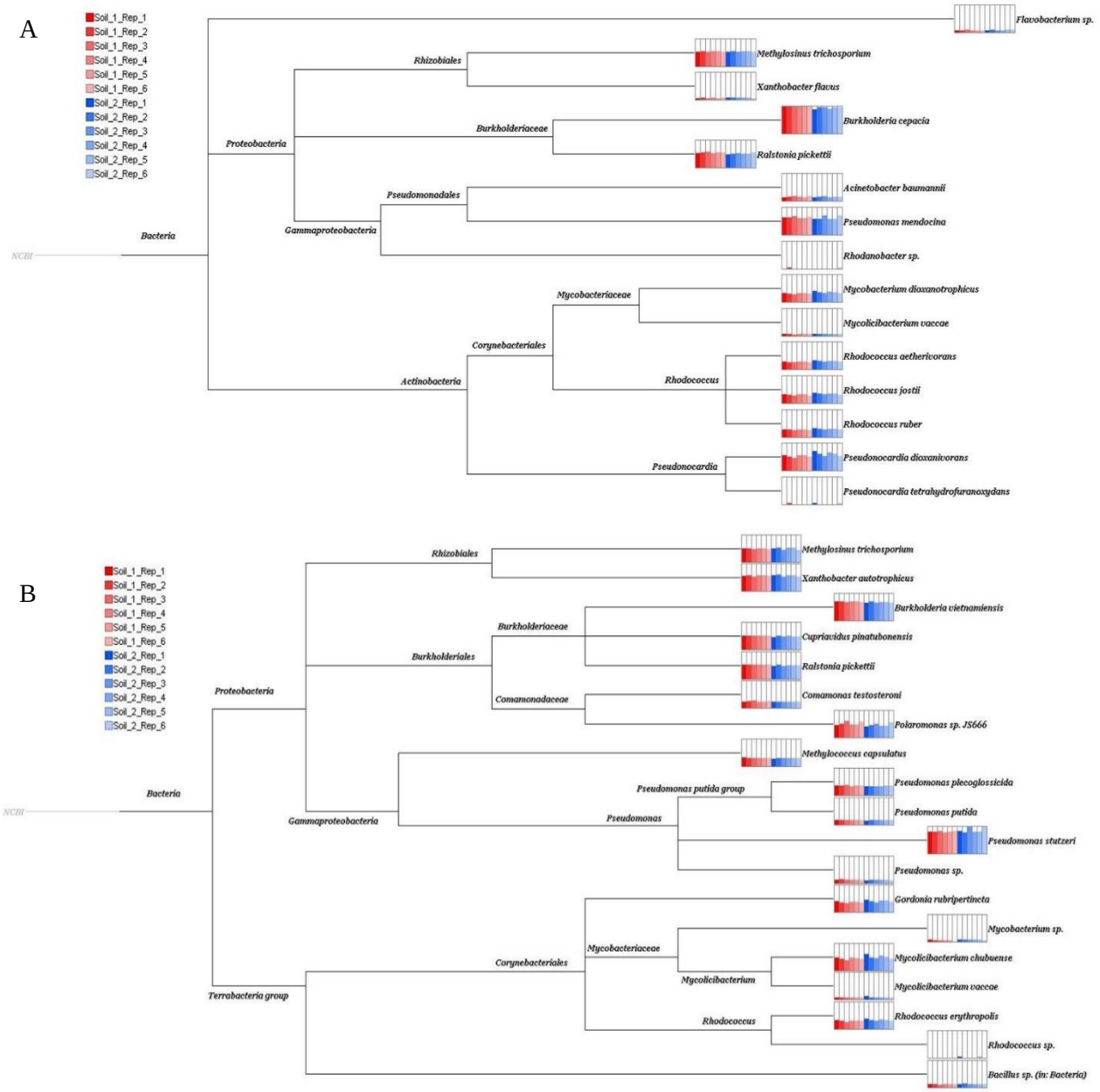
Supplementary Figure 4. 2. Rarefaction curves for DNA extracts in heavy and light fractions from 1,4-dioxane (A) and cDCE (B) SIP experiments in microcosms amended with soil 1. and 2.



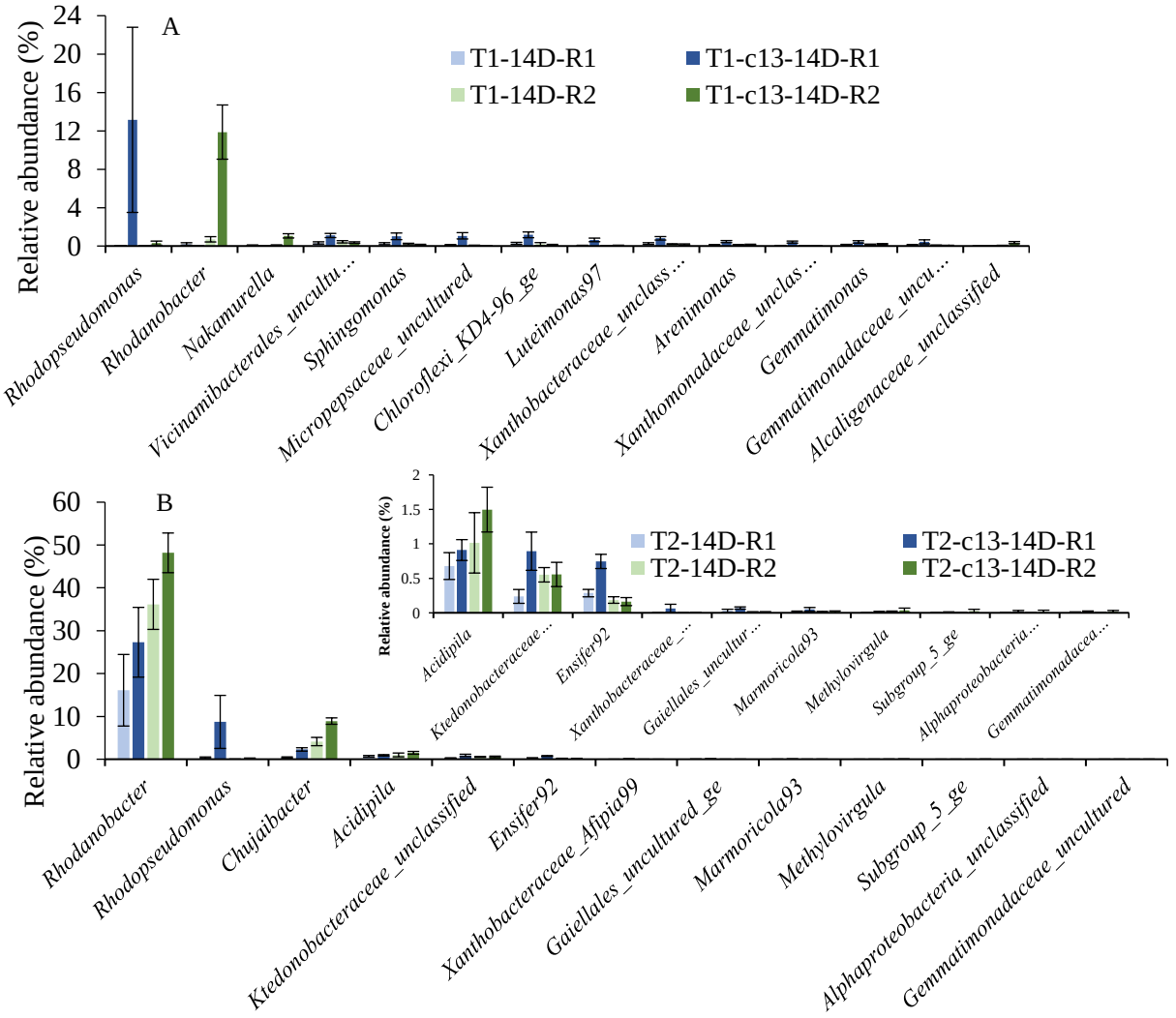
Supplementary Figure 4. 3. Non-metric Multi-dimensional Scaling (NMDS) plot and alpha diversity measurements for sequencing results of 1,4-dioxane and cDCE SIP tests by KBS soil 1 and 2.



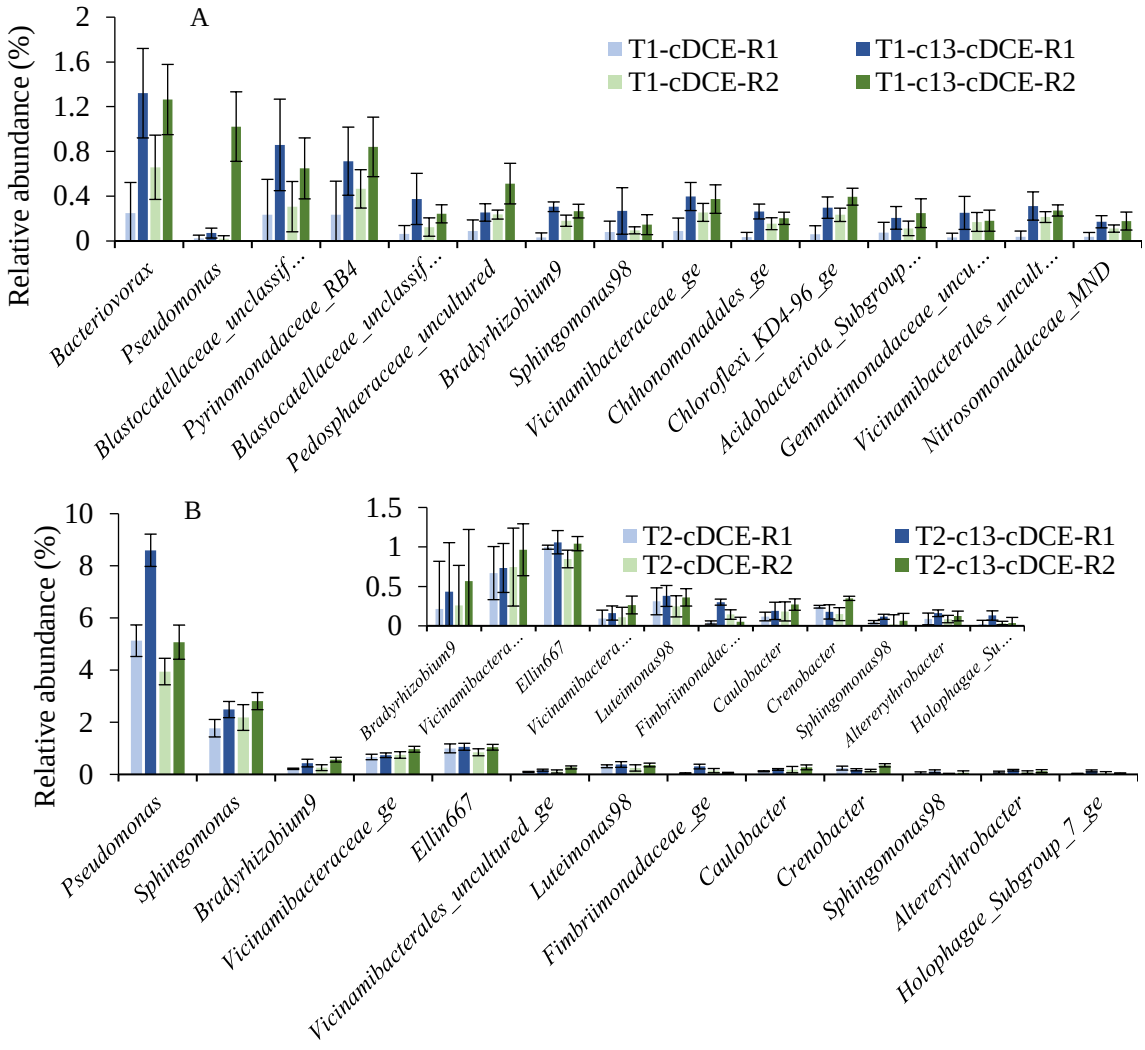
Supplementary Figure 4. 4. The most abundant (top 40) genera (by mean, with phylum) in all SIP fractions from 1,4-dioxane (A) and cDCE (B) amended microcosms inoculated with soil 1 or 2.



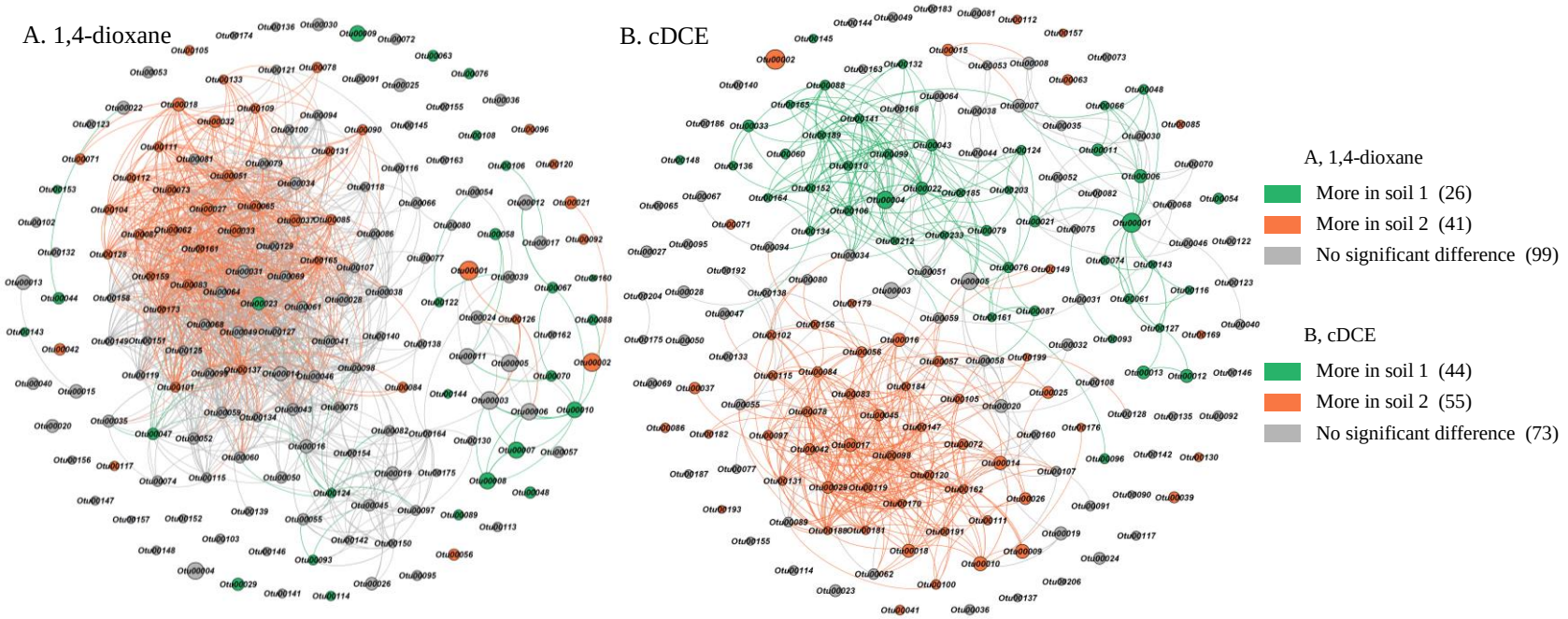
Supplementary Figure 4. 5. Species or strains previously associated with 1,4-dioxane (A) or cDCE (B) biodegradation present in KBS soil 1 (red) and soil 2 (blue) from shotgun sequencing data.



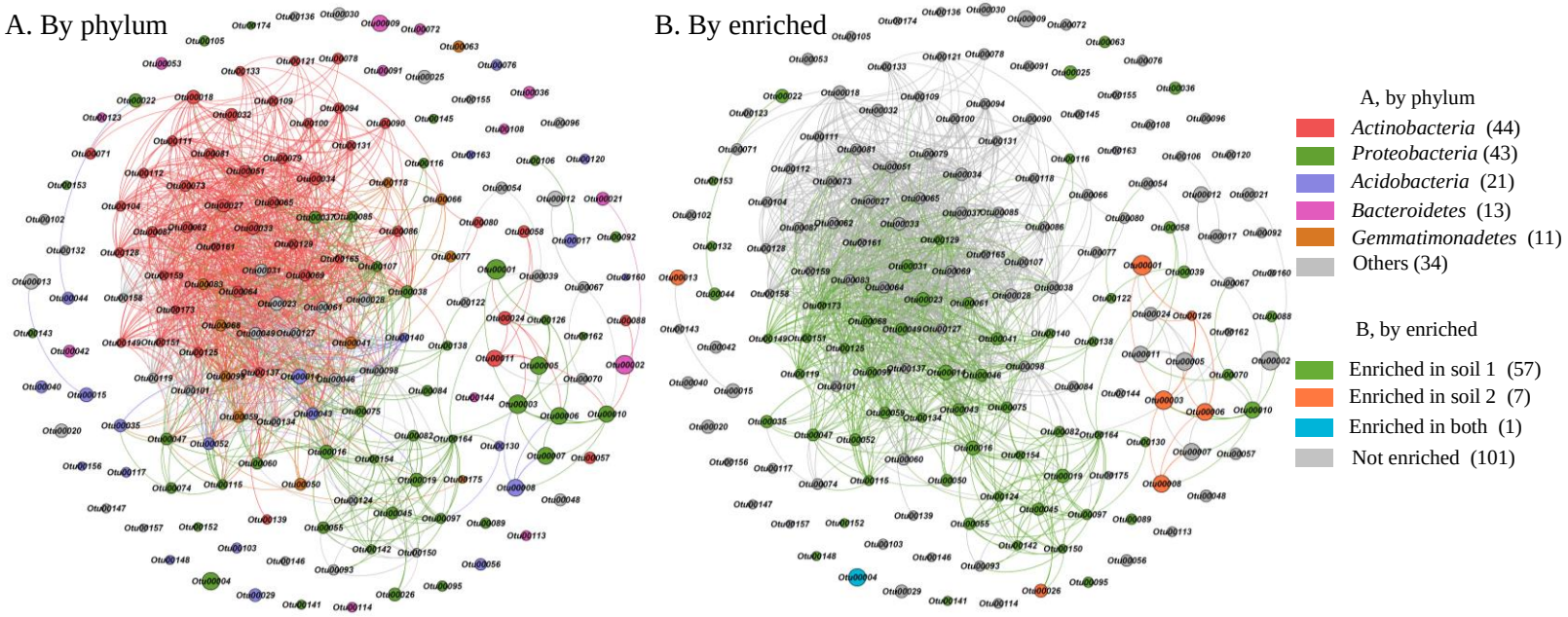
Supplementary Figure 4. 6. Phylotypes statistically enriched (Welch's two sided t-test, $p<0.05$) in heavy fractions (1.730-1.747 g/mL) of ^{13}C 1,4-dioxane amended samples compared to fractions of comparable buoyant density (1.730-1.748 g/mL) in ^{12}C 1,4-dioxane amended samples in soil 1 (A) and 2 (B). Values and error bars represent averages and standard deviations from three fractions each (with each fraction being sequenced and quantified in triplicate). After removing the background phylotypes that were also enriched in light fractions, a total of 282 and 28 phylotypes were enriched 1,4-dioxane amended samples for soil 1 and 2, respectively. The figure only displayed phylotypes with a difference of average abundance from ^{13}C 1,4-dioxane and amended ^{12}C 1,4-dioxane samples higher than 0.15% (A) and 0.01% (B) for soil 1 and soil 2. The insert was in a smaller scale.



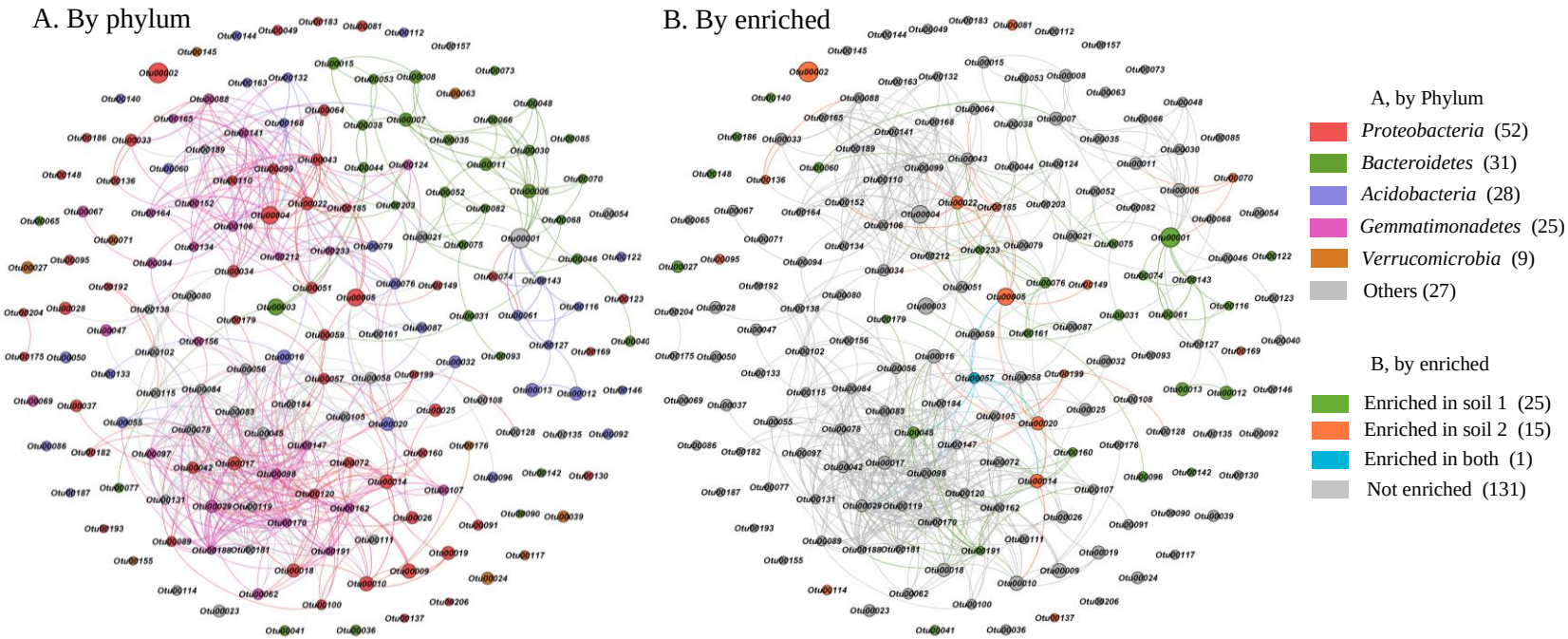
Supplementary Figure 4. 7. Phylotypes statistically enriched (Welch's two sided t-test, $p<0.05$) in heavy fractions (1.733-1.744 g/mL) of ^{13}C cDCE amended samples compared to fractions of comparable buoyant density (1.733-1.745 g/mL) in ^{12}C DCE amended samples in soil 1 (A) and 2 (B). Values and error bars represent averages and standard deviations from three fractions each (with each fraction being sequenced and quantified in triplicate). After removing the background phylotypes that were also enriched in light fractions, a total of 30 and 25 phylotypes were enriched in DCE amended samples for soil 1 and soil 2, respectively . The figure only displayed phylotypes with a difference of average abundance from ^{13}C DCE and amended ^{12}C DCE samples higher than 0.1% (A) and 0.05% (B) for soil 1 and soil 2. The insert was in a smaller scale.



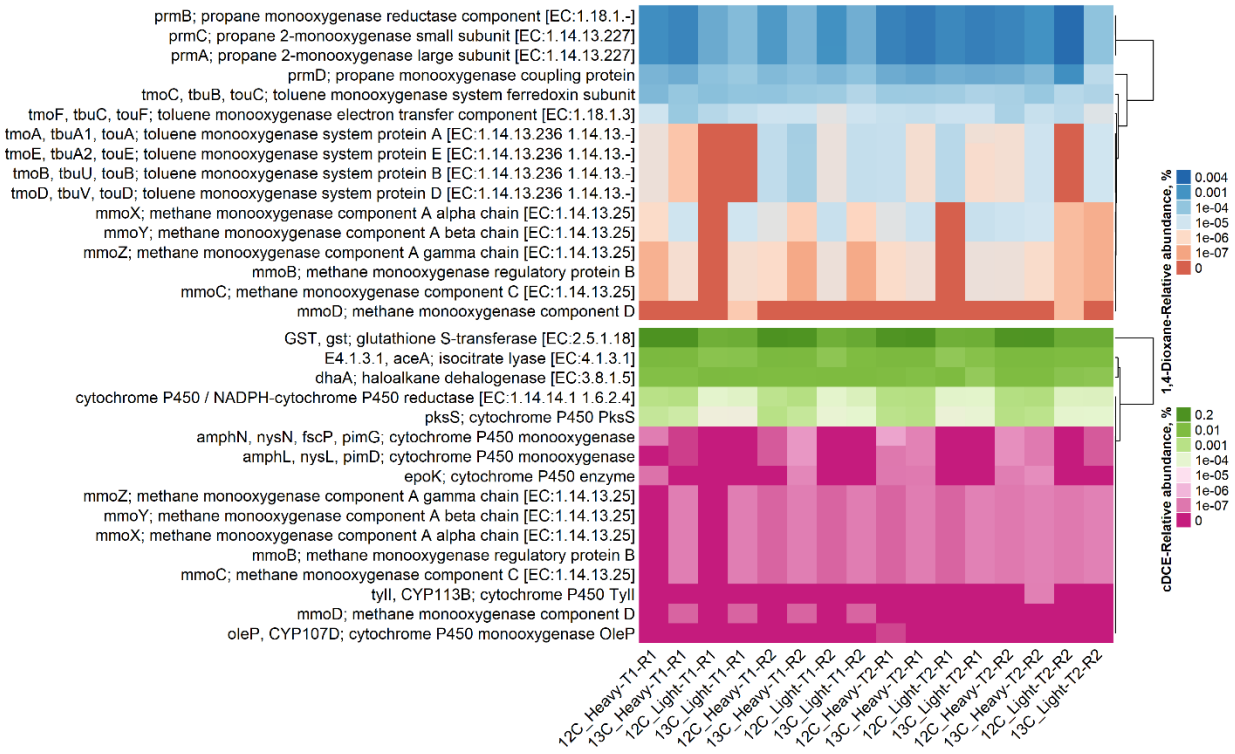
Supplementary Figure 4. 8. Co-occurrence networked based on spearman correlation (rho > 0.70 and p-value < 0.01) for the main OTUs from microbial communities for 1,4-dioxane (A) and cDCE (B) degradation. Connected nodes with lines had a rho > 0.7. Filters for main OTUs: present in at least 50% of samples, average abundance > 0.06% (A) and > 0.1% (B). There were 166 (A) and 172 (B) nodes met the filters. The networks were colored with OTUs show significant difference (p<0.05) of samples from heavy fraction of soil 1 and 2 amended with ^{13}C labeled 1,4-dioxane or cDCE. Number of nodes belonging to that group was in the parentheses.



Supplementary Figure 4. 9. Co-occurrence networked based on spearman correlation (rho > 0.70 and p-value < 0.01) for the main OTUs from microbial communities for 1,4-dioxane degradation. Connected nodes with lines had a rho > 0.7. Filters for main OTUs: present in at least 50% of samples, average abundance > 0.06%. There were 166 nodes met the filters. A: OTUs were colored by phylum, B: OTUs were colored by if its abundance is significantly higher in DNA with C13 isotope high BD value fractions from soil 1 or soil 2. Number of nodes belonging to that group was in the parentheses.



Supplementary Figure 4. 10. Co-occurrence networked based on spearman correlation (rho > 0.70 and p-value < 0.01) for the main OTUs from microbial communities for cDCE degradation. Connected nodes with lines had a rho > 0.7. Filters for main OTUs: present in at least 50% of samples, average abundance > 0.1%. There were 172 nodes met the filters. A: OTUs were colored by phylum, B: OTUs were colored by if its abundance is significantly higher in DNA with C13 isotope high BD value fractions from soil 1 or soil 2. Number of nodes belonging to that group was in the parentheses.



Supplementary Figure 4. 11. KO functions associated with 1,4-dioxane and cDCE in SIP fractions obtained from microcosm replicates (R1 and R2) in both soil 1 (T1) and 2 (T2).

REFERENCES

REFERENCES

1. Steffan RJ, Vainberg S. 2013. Production and handling of *Dehalococcoides* bioaugmentation cultures, p 89-115. *In* Stroo HF, Leeson A, Ward CH (ed), Bioaugmentation for Groundwater Remediation. Springer, New York.

2. Brown RA, Mueller JG, Seech AG, Henderson JK, Wilson JT. 2009. Interactions between biological and abiotic pathways in the reduction of chlorinated solvents. Remediation Journal 20:9-20.

3. Ellis DE, Lutz EJ, Odom JM, Buchanan RJ, Bartlett CL, Lee MD, Harkness MR, DeWeerd KA. 2000. Bioaugmentation for Accelerated In Situ Anaerobic Bioremediation. Environmental Science & Technology 34:2254-2260.

4. Yang Y, McCarty PL. 2002. Comparison between donor substrates for biologically enhanced tetrachloroethene DNAPL dissolution. Environ Sci Technol 36:3400-4.

5. Yang Y, McCarty PL. 2000. Biologically Enhanced Dissolution of Tetrachloroethene DNAPL. Environmental Science & Technology 34:2979-2984.

6. EPA. 2017. Technical Fact Sheet - 1,4-Dioxane. https://www.epa.gov/sites/production/files/2014-03/documents/ffrro_factsheet_contaminant_14-dioxane_january2014_final.pdf.

7. Zenker MJ, Borden RC, Barlaz MA. 2003. Occurrence and Treatment of 1,4-Dioxane in Aqueous Environments. Environmental Engineering Science 20:423-432.

8. Anderson RH, Anderson JK, Bower PA. 2012. Co-occurrence of 1,4-dioxane with trichloroethylene in chlorinated solvent groundwater plumes at US Air Force installations: Fact or fiction. Integrated Environmental Assessment and Management 8:731-737.

9. Adamson DT, Mahendra S, Walker KL, Rauch SR, Sengupta S, Newell CJ. 2014. A Multisite Survey To Identify the Scale of the 1,4-Dioxane Problem at Contaminated Groundwater Sites. Environmental Science & Technology Letters 1:254-258.

10. Mahendra S, Alvarez-Cohen L. 2006. Kinetics of 1,4-dioxane biodegradation by monooxygenase-expressing bacteria. Environ Sci Technol 40:5435-42.

11. He Y, Mathieu J, Yang Y, Yu P, da Silva MLB, Alvarez PJJ. 2017. 1,4-Dioxane Biodegradation by *Mycobacterium dioxanotrophicus* PH-06 Is Associated with a Group-6 Soluble Di-Iron Monooxygenase. Environmental Science & Technology Letters 4:494-499.

12. Mahendra S, Grostern A, Alvarez-Cohen L. 2013. The impact of chlorinated solvent co-contaminants on the biodegradation kinetics of 1,4-dioxane. Chemosphere 91:88-92.

13. Zhang S, Gedalanga PB, Mahendra S. 2016. Biodegradation Kinetics of 1,4-Dioxane in Chlorinated Solvent Mixtures. Environ Sci Technol 50:9599-607.

14. Miao Y, Johnson NW, Gedalanga PB, Adamson D, Newell C, Mahendra S. 2019. Response and recovery of microbial communities subjected to oxidative and biological treatments of 1,4-dioxane and co-contaminants. Water Research 149:74-85.

15. Miao Y, Johnson NW, Phan T, Heck K, Gedalanga PB, Zheng X, Adamson D, Newell C, Wong MS, Mahendra S. 2020. Monitoring, assessment, and prediction of microbial shifts in coupled catalysis and biodegradation of 1,4-dioxane and co-contaminants. Water Research 173:115540.

16. Polasko AL, Zulli A, Gedalanga PB, Pornwongthong P, Mahendra S. 2019. A Mixed Microbial Community for the Biodegradation of Chlorinated Ethenes and 1,4-Dioxane. Environmental Science & Technology Letters 6:49-54.

17. Deng D, Li F, Wu C, Li M. 2018. Synchronic Biotransformation of 1,4-Dioxane and 1,1-Dichloroethylene by a Gram-Negative Propanotroph *Azoarcus* sp. DD4. Environmental Science & Technology Letters 5:526-532.

18. Li F, Deng D, Zeng L, Abrams S, Li M. 2021. Sequential anaerobic and aerobic bioaugmentation for commingled groundwater contamination of trichloroethene and 1,4-dioxane. Sci Total Environ 774:145118.

19. Ramalingam V, Cupples AM. 2020. Enrichment of novel Actinomycetales and the detection of monooxygenases during aerobic 1,4-dioxane biodegradation with uncontaminated and contaminated inocula. Appl Microbiol Biotechnol 104:2255-2269.

20. Radajewski S, Ineson P, Parekh NR, Murrell JC. 2000. Stable-isotope probing as a tool in microbial ecology. Nature 403:646-649.

21. Paes F, Liu X, Mattes TE, Cupples AM. 2015. Elucidating carbon uptake from vinyl chloride using stable isotope probing and Illumina sequencing. Applied Microbiology and Biotechnology 99:7735-7743.

22. Aoyagi T, Morishita F, Sugiyama Y, Ichikawa D, Mayumi D, Kikuchi Y, Ogata A, Muraoka K, Habe H, Hori T. 2018. Identification of active and taxonomically diverse 1,4-dioxane degraders in a full-scale activated sludge system by high-sensitivity stable isotope probing. ISME J 12:2376-2388.

23. Cho K-C, Lee DG, Roh H, Fuller ME, Hatzinger PB, Chu K-H. 2013. Application of 13C-stable isotope probing to identify RDX-degrading microorganisms in groundwater. Environmental Pollution 178:350-360.

24. Jayamani I, Cupples AM. 2015. Stable isotope probing reveals the importance of *Comamonas* and *Pseudomonadaceae* in RDX degradation in samples from a Navy detonation site. Environ Sci Pollut Res Int 22:10340-50.

25. Sun W, Sun X, Cupples AM. 2012. Anaerobic methyl tert-butyl ether-degrading microorganisms identified in wastewater treatment plant samples by stable isotope probing. Appl Environ Microbiol 78:2973-80.

26. Fennell DE, Carroll AB, Gossett JM, Zinder SH. 2001. Assessment of Indigenous Reductive Dechlorinating Potential at a TCE-Contaminated Site Using Microcosms, Polymerase Chain Reaction Analysis, and Site Data. Environmental Science & Technology 35:1830-1839.

27. Thelusmond JR, Strathmann TJ, Cupples AM. 2019. Carbamazepine, triclocarban and triclosan biodegradation and the phylotypes and functional genes associated with xenobiotic degradation in four agricultural soils. Sci Total Environ 657:1138-1149.

28. Parales RE, Adamus JE, White N, May HD. 1994. Degradation of 1,4-dioxane by an Actinomycete in pure culture. Appl Environ Microbiol 60:4527-30.

29. Gossett JM. 1987. Measurement of Henry's law constants for C1 and C2 chlorinated hydrocarbons. Environmental Science & Technology 21:202-208.

30. Myers MA, Johnson NW, Marin EZ, Pornwongthong P, Liu Y, Gedalanga PB, Mahendra S. 2018. Abiotic and bioaugmented granular activated carbon for the treatment of 1,4-dioxane-contaminated water. Environ Pollut 240:916-924.

31. Freedman DL, Gossett JM. 1989. Biological reductive dechlorination of tetrachloroethylene and trichloroethylene to ethylene under methanogenic conditions. Appl Environ Microbiol 55:2144-51.

32. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol 79:5112-20.

33. Schloss PD. 2009. A high-throughput DNA sequence aligner for microbial ecology studies. PLoS One 4:e8230.

34. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glockner FO. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188-96.

35. Lahti L, Shetty S. 2017. Tools for microbiome analysis in R, http://microbiome.github.com/microbiome.

36. McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8:e61217.

37. Andersen KS, Kirkegaard RH, Karst SM, Albertsen M. 2018. ampvis2: an R package to analyse and visualise 16S rRNA amplicon data. bioRxiv doi:10.1101/299537:299537.

38. Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

39. Harrell FE. 2020. Hmisc: Harrell Miscellaneous, https://cran.r-project.org/web/packages/Hmisc/index.html.

40. Bates D, Maechler M, Davis TA, Oehlschlägel J, Riedy J. 2019. Matrix: Sparse and Dense Matrix Classes and Methods, https://cran.r-project.org/web/packages/Matrix/index.html.

41. Csardi G, Nepusz T. 2006. The igraph software package for complex network research. InterJournal, Complex Systems 1695.

42. Kassambara A. 2020. ggpubr: 'ggplot2' Based Publication Ready Plots., https://CRAN.R-project.org/package=ggpubr.

43. R Core Team. 2018. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. , https://www.R-project.org/.

44. RStudio Team. 2020. RStudio: Integrated Development for R., RStudio, PBC. Boston, MA, http://www.rstudio.com/.

45. Parks DH, Tyson GW, Hugenholtz P, Beiko RG. 2014. STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics 30:3123-4.

46. Bastian M, Heymann S, Jacomy M. 2006. Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media.

47. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44:D457-62.

48. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI. 2020. PICRUSt2 for prediction of metagenome functions. Nat Biotechnol 38:685-688.

49. Grostern A, Sales CM, Zhuang W-Q, Erbilgin O, Alvarez-Cohen L. 2012. Glyoxylate Metabolism Is a Key Feature of the Metabolic Degradation of 1,4-Dioxane by *Pseudonocardia dioxanivorans* Strain CB1190. Applied and Environmental Microbiology 78:3298.

50. Nishino SF, Shin KA, Gossett JM, Spain JC. 2013. Cytochrome P450 Initiates Degradation of cis-Dichloroethene by *Polaromonas* sp Strain JS666. Applied and Environmental Microbiology 79:2263-2272.

51. Giddings CGS, Jennings LK, Gossett JM. 2010. Microcosm Assessment of a DNA Probe Applied to Aerobic Degradation of cis-1,2-Dichloroethene by *Polaromonas* sp. Strain JS666. Groundwater Monitoring & Remediation 30:97-105.

52. Van Hylckama VJ, De Koning W, Janssen DB. 1997. Effect of Chlorinated Ethene Conversion on Viability and Activity of *Methylosinus trichosporium* OB3b. Appl Environ Microbiol 63:4961-4.

53. Semprini L. 1997. Strategies for the aerobic co-metabolism of chlorinated solvents. Curr Opin Biotechnol 8:296-308.

54. Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32:2847-9.

55. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-20.

56. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW. 2016. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods 102:3-11.

57. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403-10.

58. Huson DH, Beier S, Flade I, Gorska A, El-Hadidi M, Mitra S, Ruscheweyh HJ, Tappu R. 2016. MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. PLoS Comput Biol 12:e1004957.

59. Jennings LK, Chartrand MM, Lacrampe-Couloume G, Lollar BS, Spain JC, Gossett JM. 2009. Proteomic and transcriptomic analyses reveal genes upregulated by cis-dichloroethene in *Polaromonas* sp. strain JS666. Appl Environ Microbiol 75:3733-44.

60. Pugazhendi A. 2015. Biodegradation of 1,4-dioxane by *Rhodanobacter* AYS5 and the role of additional substrates. Annals of Microbiology 65.

61. Sei K, Miyagaki K, Kakinoki T, Fukugasako K, Inoue D, Ike M. 2013. Isolation and characterization of bacterial strains that have high ability to degrade 1,4-dioxane as a sole carbon and energy source. Biodegradation 24:665-674.

62. He Y, Mathieu J, da Silva MLB, Li M, Alvarez PJJ. 2018. 1,4-Dioxane-degrading consortia can be enriched from uncontaminated soils: prevalence of *Mycobacterium* and soluble di-iron monooxygenase genes. Microbial Biotechnology 11:189-198.

63. Nam JH, Ventura JS, Yeom IT, Lee Y, Jahng D. 2016. Structural and Kinetic Characteristics of 1,4-Dioxane-Degrading Bacterial Consortia Containing the Phylum TM7. J Microbiol Biotechnol 26:1951-1964.

64. Ma F, Wang Y, Yang J, Guo H, Su D, Yu L. 2021. Degradation of 1,4-Dioxane by *Xanthobacter* sp. YN2. Curr Microbiol 78:992-1005.

65. Chen D-Z, Jin X-J, Chen J, Ye J-X, Jiang N-X, Chen J-M. 2016. Intermediates and substrate interaction of 1,4-dioxane degradation by the effective metabolizer *Xanthobacter flavus* DT8. International Biodeterioration & Biodegradation 106:133-140.

66. Chen R, Miao Y, Liu Y, Zhang L, Zhong M, Adams JM, Dong Y, Mahendra S. 2021. Identification of novel 1,4-dioxane degraders and related genes from activated sludge by taxonomic and functional gene sequence analysis. J Hazard Mater 412:125157.

67. Xiong Y, Mason OU, Lowe A, Zhou C, Chen G, Tang Y. 2019. Microbial Community Analysis Provides Insights into the Effects of Tetrahydrofuran on 1,4-Dioxane Biodegradation. Applied and Environmental Microbiology 85:e00244-19.

68. Xiong Y, Mason OU, Lowe A, Zhang Z, Zhou C, Chen G, Villalonga MJ, Tang Y. 2020. Investigating promising substrates for promoting 1,4-dioxane biodegradation: effects of ethane and tetrahydrofuran on microbial consortia. Biodegradation 31:171-182.

69. Tusher TR, Shimizu T, Inoue C, Chien MF. 2019. Enrichment and Analysis of Stable 1,4-dioxane-Degrading Microbial Consortia Consisting of Novel Dioxane-Degraders. Microorganisms 8.

70. Hu P, Dubinsky EA, Probst AJ, Wang J, Sieber CMK, Tom LM, Gardinali PR, Banfield JF, Atlas RM, Andersen GL. 2017. Simulation of Deepwater Horizon oil plume reveals substrate specialization within a complex community of hydrocarbon degraders. Proceedings of the National Academy of Sciences 114:7432-7437.

71. Bacosa HP, Erdner DL, Rosenheim BE, Shetty P, Seitz KW, Baker BJ, Liu Z. 2018. Hydrocarbon degradation and response of seafloor sediment bacterial community in the northern Gulf of Mexico to light Louisiana sweet crude oil. ISME J 12:2532-2543.

72. Bock C, Kroppenstedt RM, Diekmann H. 1996. Degradation and bioconversion of aliphatic and aromatic hydrocarbons by *Rhodococcus ruber* 219. Applied Microbiology and Biotechnology 45:408-410.

73. Kampfer P, Kroppenstedt RM. 2004. *Pseudonocardia benzenivorans* sp. nov. Int J Syst Evol Microbiol 54:749-751.

74. Kim YM, Jeon JR, Murugesan K, Kim EJ, Chang YS. 2009. Biodegradation of 1,4-dioxane and transformation of related cyclic compounds by a newly isolated *Mycobacterium* sp. PH-06. Biodegradation 20:511-9.

75. Huang H, Shen D, Li N, Shan D, Shentu J, Zhou Y. 2014. Biodegradation of 1,4-Dioxane by a Novel Strain and Its Biodegradation Pathway. Water, Air, & Soil Pollution 225:2135.

76. Inoue D, Tsunoda T, Sawada K, Yamamoto N, Saito Y, Sei K, Ike M. 2016. 1,4-Dioxane degradation potential of members of the genera *Pseudonocardia* and *Rhodococcus*. Biodegradation 27:277-286.

77. Kohlweyer U, Thiemer B, Schrader T, Andreesen JR. 2000. Tetrahydrofuran degradation by a newly isolated culture of *Pseudonocardia* sp. strain K1. FEMS Microbiol Lett 186:301-6.

78. Vainberg S, McClay K, Masuda H, Root D, Condee C, Zylstra GJ, Steffan RJ. 2006. Biodegradation of ether pollutants by *Pseudonocardia* sp. strain ENV478. Appl Environ Microbiol 72:5218-24.

79. Steffan RJ, McClay K, Vainberg S, Condee CW, Zhang D. 1997. Biodegradation of the gasoline oxygenates methyl tert-butyl ether, ethyl tert-butyl ether, and tert-amyl methyl ether by propane-oxidizing bacteria. Applied and Environmental Microbiology 63:4216.

80. Stringfellow WT, Alvarez-Cohen L. 1999. Evaluating the relationship between the sorption of PAHs to bacterial biomass and biodegradation. Water Research 33:2535-2544.

81. Sun B, Ko K, Ramsay JA. 2011. Biodegradation of 1,4-dioxane by a *Flavobacterium*. Biodegradation 22:651-9.

82. Burback BL, Perry JJ. 1993. Biodegradation and biotransformation of groundwater pollutant mixtures by *Mycobacterium vaccae*. Appl Environ Microbiol 59:1025-9.

83. Masuda H, McClay K, Steffan RJ, Zylstra GJ. 2012. Characterization of three propane-inducible oxygenases in *Mycobacterium* sp. strain ENV421. Lett Appl Microbiol 55:175-81.

84. Whited GM, Gibson DT. 1991. Separation and partial characterization of the enzymes of the toluene-4-monooxygenase catabolic pathway in *Pseudomonas mendocina* KR1. J Bacteriol 173:3017-20.

85. Kukor JJ, Olsen RH. 1990. Molecular cloning, characterization, and regulation of a *Pseudomonas pickettii* PKO1 gene encoding phenol hydroxylase and expression of the gene in *Pseudomonas aeruginosa* PAO1c. J Bacteriol 172:4624-30.

86. Nelson MJ, Montgomery SO, O'Neill E J, Pritchard PH. 1986. Aerobic metabolism of trichloroethylene by a bacterial isolate. Appl Environ Microbiol 52:383-4.

87. Whittenbury R, Phillips KC, Wilkinson JF. 1970. Enrichment, Isolation and Some Properties of Methane-utilizing Bacteria. Microbiology 61:205-218.

88. Yao Y, Lv Z, Min H, Lv Z, Jiao H. 2009. Isolation, identification and characterization of a novel *Rhodococcus* sp. strain in biodegradation of tetrahydrofuran and its medium optimization using sequential statistics-based experimental designs. Bioresource Technology 100:2762-2769.

89. Sharp JO, Sales CM, LeBlanc JC, Liu J, Wood TK, Eltis LD, Mohn WW, Alvarez-Cohen L. 2007. An Inducible Propane Monooxygenase Is Responsible for Nitrosodimethylamine Degradation by *Rhodococcus* sp. Strain RHA1. Applied and Environmental Microbiology 73:6930.

90. Thiemer B, Andreesen JR, Schrader T. 2003. Cloning and characterization of a gene cluster involved in tetrahydrofuran degradation in *Pseudonocardia* sp. strain K1. Arch Microbiol 179:266-77.

91. Smith CA, O'Reilly KT, Hyman MR. 2003. Characterization of the initial reactions during the cometabolic oxidation of methyl tert-butyl ether by propane-grown *Mycobacterium vaccae* JOB5. Appl Environ Microbiol 69:796-804.

92. Rasmussen MT, Saito AM, Hyman MR, Semprini L. 2020. Co-encapsulation of slow release compounds and *Rhodococcus rhodochrous* ATCC 21198 in gellan gum beads to promote the long-term aerobic cometabolic transformation of 1,1,1-trichloroethane, cis-1,2-dichloroethene and 1,4-dioxane. Environ Sci Process Impacts 22:771-791.

93. Lee SW, Keeney DR, Lim DH, Dispirito AA, Semrau JD. 2006. Mixed pollutant degradation by Methylosinus trichosporium OB3b expressing either soluble or particulate methane monooxygenase: can the tortoise beat the hare? Appl Environ Microbiol 72:7503-9.

94. Mattes TE, Alexander AK, Coleman NV. 2010. Aerobic biodegradation of the chloroethenes: pathways, enzymes, ecology, and evolution. FEMS Microbiol Rev 34:445-75.

95. Elango V, Kurtz HD, Freedman DL. 2011. Aerobic cometabolism of trichloroethene and cis-dichloroethene with benzene and chlorinated benzenes as growth substrates. Chemosphere 84:247-253.

96. Zalesak M, Ruzicka J, Vicha R, Dvorackova M. 2017. Cometabolic degradation of dichloroethenes by *Comamonas testosteroni* RF2. Chemosphere 186:919-927.

97. Aulenta F, Verdini R, Zeppilli M, Zanaroli G, Fava F, Rossetti S, Majone M. 2013. Electrochemical stimulation of microbial *cis*-dichloroethene (*cis*-DCE) oxidation by an ethene-assimilating culture. N Biotechnol 30:749-55.

98. Chauhan S, Barbieri P, Wood TK. 1998. Oxidation of trichloroethylene, 1,1-dichloroethylene, and chloroform by toluene/o-xylene monooxygenase from *Pseudomonas stutzeri* OX1. Appl Environ Microbiol 64:3023-4.

99. Elango VK, Liggenstoffer AS, Fathepure BZ. 2006. Biodegradation of vinyl chloride and *cis*-dichloroethene by a *Ralstonia* sp. strain TRW-1. Appl Microbiol Biotechnol 72:1270-5.

100. Yonezuka K, Araki N, Shimodaira J, Ohji S, Hosoyama A, Numata M, Yamazoe A, Kasai D, Masai E, Fujita N, Ezaki T, Fukuda M. 2016. Isolation and characterization of a bacterial strain that degrades *cis*-dichloroethenein the absence of aromatic inducers. The Journal of General and Applied Microbiology 62:118-125.

101. van Hylckama Vlieg JET, Kingma J, van den Wijngaard AJ, Janssen DB. 1998. A Glutathione-Transferase with Activity towards *cis*-1,2-Dichloroepoxyethane Is Involved in Isoprene Utilization by *Rhodococcus* sp. Strain AD45. Applied and Environmental Microbiology 64:2800.

102. Li J, de Toledo RA, Chung J, Shim H. 2014. Removal of mixture of cis-1,2-dichloroethylene/trichloroethylene/benzene, toluene, ethylbenzene, and xylenes from contaminated soil by *Pseudomonas plecoglossicida*. Journal of Chemical Technology & Biotechnology 89:1934-1940.

103. Dolinova I, Strojsova M, Cernik M, Nemecek J, Machackova J, Sevcu A. 2017. Microbial degradation of chloroethenes: a review. Environ Sci Pollut Res Int 24:13262-13283.

CHAPTER 5 Conclusions and Future Research Directions

The key findings from each chapter are summarized below:

Chapter 2. Analysis via shotgun sequencing of groundwater at five SDC-9 bioaugmented chlorinated solvent contaminated sites indicated the presence of DNA from numerous biodegraders, including *Dehalococcoides, Desulfitobacterium* and *Dehalogenimonas*. Further, DNA sequences from both anaerobic (*pceA, tceA, vcrA* and *bvcA*) and aerobic (*etnE, etnC, mmoX* and *pmoA*) functional genes were also detected. Additionally, DNA sequences from hydrogenases and functional genes associated with corrinoid metabolism and 1,4-dioxane degradation were also observed.

Chapter 3. The analysis of groundwater from a biostimulated RDX contaminated site indicated DNA from more than thirty RDX degrading species were present in the pre- and post-biostimulated groundwater samples, with *Variovorax* sp. JS1663, *Pseudomonas fluorescens*, *Pseudomonas putida* and *Stenotrophomonas maltophilia* being the most abundant. From these, nine RDX degrading species significantly ($p<0.05$) increased in abundance following biostimulation. Both shotgun sequencing and qPCR indicated the most abundant RDX degrading functional genes were *xenA* and *xenB*. The relative abundance percentages of three RDX degrading genes (*diaA, nsfI* and *pnrB*) were similar and *xplA* was low or absent in most of the samples.

Chapter 4. This study identified phylotypes associated with 1,4-dioxane and cDCE biodegradation using 16S rRNA gene amplicon sequencing coupled with SIP. In the 1,4-dioxane degrading microcosms two genera (*Rhodopseudomonas* and *Rhodanobacter*) were associated with the majority of ^{13}C assimilation from 1,4-dioxane. In the cDCE degrading microcosms, the

dominant genera for ^{13}C assimilation included *Bacteriovorax, Pseudomonas* and *Sphingomonas.* The predicted functions associated with 1,4-dioxane and cDCE biodegradation were also determined. Overall, the work demonstrated concurrent removal of cDCE and 1,4-dioxane by indigenous soil microbial communities and the enhancement of cDCE removal by lactate. The data generated on the phylotypes responsible for carbon uptake (as determined by SIP) could be incorporated into diagnostic molecular methods for site characterization. The results suggest aerobic concurrent biodegradation of cDCE and 1,4-dioxane should be considered for chlorinated solvent site remediation.

Future research could include manipulating the DNA concentration submitted for sequencing, so that the comparison across samples would be based on the changes of absolute values rather than the relative abundance of a taxon or gene. Further, as additional sequencing data becomes available, data mining activities could improve our understanding of biodegradation potential across sites. Also, future research should include the correlation of geochemical data with molecular data to determine which factors are beneficial or impact the functional genera. In addition, future research could involve RNA-seq (RNA is reverse transcribed to cDNA, and submitted for high throughput sequencing) to reveal the active functions during contaminant biodegradation.