FINAL REPORT

Application of Microarrays and qPCR to Identify Phylogenetic and Functional Biomarkers Diagnostic of Microbial Communities that Biodegrade Chlorinated Solvents to Ethene

SERDP Project ER-1587

January 2012

Lisa Alvarez-Cohen
University of California at Berkeley

SERDP
DOD ▪ EPA ▪ DOE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-2012	Final Report	01-2007-01-2012

4. TITLE AND SUBTITLE

Application of microarrays and qPCR to identify phylogenetic and functional biomarkers diagnostic of microbial communities that biodegrade chlorinated solvents to ethene

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Lisa Alvarez-Cohen

5d. PROJECT NUMBER

ER-1587

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of California at Berkeley
760 Davis Hall, CEE Dept.
Berkeley, CA 94720-1710

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

SERDP/ESTCP
4800 Mark Center Drive, Suite 17D08
Alexandria, VA 22350-3605

10. SPONSOR/MONITOR'S ACRONYM(S)

SERDP/ESTCP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

ER-1587

12. DISTRIBUTION/AVAILABILITY STATEMENT

Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

During this 4-year SERDP project, we've made great progress and success on the application of 16S rRNA- and mRNA-based microarray, along with quantitative PCR tools to monitor Dehalococcoides(Dhc)-contatining microbial communities capable of reductive dechlorination of chlorinated solvents. Dechlorination activities, as well as community structures were characterized in dechlorinating microbial communities, including laboratory-scale enrichment cultures, semi-batch microcosm, continuous-flow chemostat, and environmental samples from contaminated field sites. Firmicutes (Mostly Clostridium spp.), Bacteroidetes (Mostly Bacteroides spp.), as well as Proteobacteria (Mostly sulfate-reducer, i.e. Desulfovibrio spp.) were the most commonly found supportive microorganisms in those characterized communities. Since most of them are capable of fermenting organic compounds such as lactate and whey to acetate and hydrogen, it is indicated that these supportive microorganisms might play a role in providing hydrogen to Dhc.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		74	Aaron J. Slowey
					19b. TELEPHONE NUMBER *(Include area code)* 510-643-8739

Reset

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

TALBE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

Acronym	Full name
Cba	cobamide
Cbi	cobinamide
Cby	cobyric acid
cDCE	cis-dichloroethene
DE195	*Dehalococcoides mccartyi* strain 195
Dhc	*Dehalococcoides*
DMB	5, 6-dimethylbenzimidazole
DNAPL	dense non-aqueous-phase liquid
DVH	*Desulfovibrio vulgaris* Hildenborough
EW	extraction well
FACS	fluorescence-activated cell sorting
FID	flame ionization detector
GC	gas chromatography
HPLC	high performance liquid chromatography
HPR	high plasiticity region
IW	injection well
LC/MS/MS	liquid chromatography tandem mass spectrometry
MC	*Methanobacterium congolense*
MM	mismatch
MW	monitoring well
OTU	operational taxomonic unit
PCA	principal component analysis
PCE	tetrachloroethene
PDS	*Pelosinus, Dendrosporobacter, Sporotalea*
PM	perfect match
qPCR	quantitative polymer chain reaction
RDase	reductive dehalogenase
RGD	reduction gas detector
RT-PCR	real time PCR
SIP	stable isotope probing
SRT	sludge retention time
TCE	trichloroethene
VC	vinyl chloride

ACKNOWLEDGEMENT

This research is supported by Strategic Environmental Research and Development Program (Project number ER-1587).

ABSTRACT

During this 4-year SERDP project, we've made great progress and success on the application of 16S rRNA- and mRNA-based microarray, along with quantitative PCR tools to monitor *Dehalococcoides*(Dhc)-contatining microbial communities capable of reductive dechlorination of chlorinated solvents. Dechlorination activities, as well as community structures were characterized in dechlorinating microbial communities, including laboratory-scale enrichment cultures, semi-batch microcosm, continuous-flow chemostat, and environmental samples from contaminated field sites. Firmicutes (Mostly *Clostridium* spp.), Bacteroidetes (Mostly *Bacteroides* spp.), as well as Proteobacteria (Mostly sulfate-reducer, i.e. *Desulfovibrio* spp.) were the most commonly found supportive microorganisms in those characterized communities. Since most of them are capable of fermenting organic compounds such as lactate and whey to acetate and hydrogen, it is indicated that these supportive microorganisms might play a role in providing hydrogen to Dhc. Moreover, since *Clostridium* spp. and *Desulfovibrio* spp. were reported to have an up-stream corrinoid biosynthesis pathway, these species might also play an important role in providing corrinoids to *Dehalococcoides*. Defined consortia were further constructed by growing pure *Dehalococcoides* species (strain 195) with some of the detected supportive microorganisms. Results indicated that *Desulfovibrio vulgaris* Hildenborough exhibited the most successful syntrophic growth with strain 195. Whole-genome expression microarray results exhibited that in the co-culture of strain 195 and *D. vulgaris*, hydrogenase-encoding genes, cobalamin-associated genes as well as some genes involved in amino-acid biosynthesis were down-regulated comparing to the pure isolate, which indicating a supportive relationship in the aspects of hydrogen transfer, corrinoid and amino-acid availabilities. Furthermore, a genus-wide microarray targeting four known *Dehalococcoides* genomes was designed, validated and applied to pure cultures and communities. The application of mRNA-based whole-genome and genus-wide microarrays combining with qPCR tools indicated that the 16S rRNA gene of *Dehalococcoides* had a good correlation with the sum of well-known functional RDase genes (*tceA*, *bvcA* and *vcrA*). However, it is the expression of functional RDase genes rather than the 16S rRNA gene that determines the dechlorination activity in a community. And the functional RDase genes were not restricted to a specific genome. The method of fluorescence-activated cell sorting (FACS) in combination with whole genome amplification and microarray techniques were successfully established and applied, which would allow us to examine communities with dilute quantities of *Dehalococcoides* without time-consuming and laborious culturing steps. Expression microarray results of strain 195 grown under different corrinoid availabilities indicated corrinoid riboswitches in strain 195 played important roles in corrinoid salvaging mechanisms of strain 195. Furthermore, diverse corrinoid and lower ligand forms were identified and quantified by LC/MS/MS. Results exhibited that cobalamin was the dominant species in ANAS culture, while p-cresol cobamide was the most dominant one produced by supportive microorganisms in groundwater enrichments, and *Dehalococcoides* in these enrichments modified this unfavorable form to the favorable cobalamin in the presence of the associated lower ligand 5, 6-dimenthylbenzimidazole (DMB). The corrinoid species and DMB could also serve as diagnostic "biomarkers" in bioremediation of chlorinated solvents. The findings and the methods established in this project could be applied to prospective studies on continuous-flow reactors, as well as to field site practice in the future. Suggested biomarkers for bioremediation application are summarized in the following table:

Suggested biomarker	Advantage	Limitation	Skills required	Cost
RDase genes	Direct indication of dechlorination capabilities	Not providing other information than dechlorination	Moderate-High	Moderate
Hydrogenase (Hup, Fdh, Hym) genes	Diagnostic of hydrogen utilization	Indirect indicator	Moderate-High	Moderate
Corrinoid biosynthesis genes (*cobT*, *btuF*) and riboswitches	Diagnostic of corrinoid synthesis	Indirect indicator	Moderate-High	Moderate
16S rRNA of Dhc	Indication of dechlorination potential	Not well associated with dechlorination patterns	Moderate	Low
16S rRNA of *Desulfovibrio* spp.	Hydrogen producing microorganisms	Indirect and may be substrate-selected	Moderate	Low
16S rRNA of *Syntrophomonas* spp.	Hydrogen producing microorganisms	Indirect and may be substrate-selected	Moderate	Low
16S rRNA of *Pelosinus* spp.	Possible corrinoid producing bacteria	Indirect and may be substrate-selected	Moderate	Low
16S rRNA of *Clostridium* spp.	Possible hydrogen and corrinoid producing bacteria	Indirect and may be substrate-selected	Moderate	Low
Corrinoids and DMB	Indication of the need of external source of cobalamin	Sensitivity may be low for *in situ* monitoring, concentration step is required	Moderate	Low

PROJECT BACKGROUND

Tetrachloroethene (PCE) and trichloroethene (TCE) have been widely used as industrial solvents and, as a result of poor storage and disposal practices, are now common contaminants of groundwater resources. PCE and TCE can be effectively biodegraded in anaerobic environments by reductive dechlorination processes, and recent progress has been made towards exploiting these processes for bioremediation applications. There remains a need, however, for appropriate and cost-effective biomarkers to assess, monitor, and optimize performance.

Commonly, biomarker development has focused on identifying nucleic acid sequences, peptides, proteins, or lipids of organisms that catalyze biodegradation reactions of interest. Although promising, such approaches are limited in reductive dechlorination processes as they do not address the roles of other organisms required to support and/or enhance the activity of the dechlorinating organisms. Novel biomarkers that quantify the presence, abundance, and activity of supporting organisms are therefore needed to more effectively assess and optimize dechlorination potentials.

The recent advances in molecular techniques, such as quantitative PCR (qPCR) and the application of high throughput microarrays to both RNA and DNA, combined with the sequenced genomes of microorganisms of interest, enable us to investigate genomic and transcriptomic information in a quantitative manner and at a community or systems-levels. Specific primers and probes targeting genes of interest can be designed based on their known DNA sequences. Fluorosent indicators are employed to indicate the presence of a specific gene or expression product (RNA) of interest, indicating the presence and quantity of targeted sequences. Based on this theory, qPCR is designed to target and quantify individual genes of interest and can be expanded with a reverse trascriptase step to quantify the expression of that gene (RNA). Microarrays work on the same principle, but are designed to detect thousands of targeted genes at the same time on a small chip. qQCR is most useful when you know specifically what few genes you are interested in quantifying, whereas microarrays are a high-throughput technique that can be employed when you're not sure about the specific gene, and therefore want to track all possible genes in the genome of a known microbial strain. The application of qPCR and microarrays not only allow us to identify important biomarkers indicative of dechlorination activities from laboratory-scale studies, but also provide tools to apply the identified biomarkers to field site samples for a more effective design and monitoring of bioremediation processes.

OBJECTIVES

It is our goal to identify 16S-rRNA-based phylogenetic and mRNA-based functional biomarkers diagnostic of microbial communities that support the growth and chlorinated ethene-degrading activity of *Dehalococcoides.* To achieve this goal, the following four objectives were pursued.

1. Apply 16S RNA microarrays to identify and quantify organisms that are present and active in a range of different microbial communities capable of the complete conversion of chlorinated

ethenes to ethene. Populations to investigate include *Dehalococcoides*–containing enrichment cultures, environmental microcosms, and groundwater samples.

2. Construct defined consortia that contain *Dehalococcoides* and one or more of the active Bacteria and/or Archaea identified to be key players in objective 1. Develop 16S-rRNA-based phylogenetic biomarkers by identifying specific organisms whose presence enhances the activity of *Dehalococcoides* or imparts a novel metabolic capability.

3. Apply a genomic expression microarray that targets the described *Dehalococcoides* genus to compare transcriptomes when grown within different defined consortia. Identify mRNA-based functional biomarkers by identifying *Dehalococcoides* genes whose expression levels correlate with the presence of specific supporting organisms.

4. Derive quantitative correlations between the rRNA-based phylogenetic biomarkers identified in objective 2, the mRNA-based functional biomarkers identified in objective 3, rates of growth and dechlorination, and the metabolic requirements of the community. Use these correlations to develop tools that can be applied to assess community structure and total biodegradative potentials in unknown microbial populations.

After 4 years of work on this project (including one no-cost extension year), to date, we have successfully completed research on all of the above objectives. Important contributions with promising application prospects have been made: we have, thus far, successfully applied phylogenetic analysis tools, including clone libraries and 16S rRNA-based microarrays to *Dehalococcoides*-containing enrichment cultures, environmental microcosms and groundwater samples; we have constructed stable and robust defined consortia that can syntrophically grow with *Dehalococcoides*; we have also designed, validated and successfully applied the genus-wide microarray targeting four genomes of *Dehalococcoides*, and we have further correlated other dechlorinating bio-markers with dechlorination characteristics. Our accomplishments over the 4 years, including 14 peer-reviewed publications, and a large number of technical presentations, are described in more detail in the section of project accomplishments.

PROJECT ACCOMPLISHMENTS

Objective 1: Identification of supportive microorganisms present and active in *Dehalococcoides*-containing communities.

Recent genomic annotation of four *Dehalococcoides* strains (CBDB1, BAV1, 195, VS) revealed several missing or incomplete metabolic pathways that were assumed to be necessary for their independent survival in the natural environment. Firstly, *Dehalococcoides* spp. are strictly dependent on hydrogen as their sole electron donor for halorespiration processes, but they cannot generate hydrogen via fermentation pathways. Also notable is that none of the strains contains detectable systems for complete *de novo* biosynthesis of cobalamin. Cobalamin is an essential co-factor of reductive dehalogenases (RDases), which are the key enzymes that catalyze the reductive dechlorination reactions, and the growth and activity of *Dehalococcoides* are strictly dependent on their activity. Acquiring a steady supply of the cobalamin from the

environment is stringently regulating the growth and activity of *Dehalococcoides* spp. Finally, all sequenced *Dehalococcoides* strains were annotated to be missing biosynthetic pathways for quinones, some amino acids, and other essential nutrients. These metabolic limitations likely require these *Dehalococcoides* strains to live in close cooperation with organisms that can supply them with these growth-requiring nutrients in the environment. Therefore, identifying the growth-supporting partners of *Dehalococcoides* spp. in dechlorinating microbial communities is of great importance to develop novel biomarkers to effectively assess and optimize dechlorination processes.

In order to identify supportive microorganisms for *Dehalococcoides* strains, we proposed in this research to investigate the microbial ecology of a wide range of *Dehalococcoides*-containing microbial communities grown under different conditions using analytical and molecular tools. Besides previously characterized mixed culture ANAS, which was highly enriched and maintained for over 10 years, we've also constructed similar TCE-dechlorinating enrichments from groundwater samples under different growth conditions, and further constructed and operated a continuous flow chemostat for long-term maintenance of dechlorinating enrichment cultures. 16S rRNA-based phylogenetic analysis tools, such as clone libraries, qPCR, etc. have been applied to these communities, together with physiological measurements. Moreover, we've successfully applied 16S rRNA of *Dehalococcoides*-containing communities from TCE-contaminated field sites to 16S rRNA-based phyloarrays, in order to identify microorganisms that are functioning in TCE-dechlorinating communities. Summaries of publications describing research results derived from this project are provided below:

1.1 Characterization of four TCE-dechlorinating microbial enrichments grown with different cobalamin stress and methanogenic conditions (This study is submitted for publication)

Materials and methods. The groundwater enrichments were originally constructed using groundwater samples from a TCE-contaminated groundwater site in New Jersey. TCE-contaminated groundwater from different monitoring wells were sampled, stored in well-sealed bottles at 4°C, and sent to our lab overnight. Microorganisms in the groundwater were collected by filtering 200 mL groundwater through a 0.22 µm filter. The filter was then put into a 160 mL serum bottle, amended with 50 mL flow-through groundwater, and 50 mL basal medium. All steps were manipulated in an anaerobic chamber. The inoculated bottles were then flushed with N_2/CO_2 headspace and amended with 20 mM lactate as electron donor, 2 µL TCE (~22 µmol) as electron acceptor, and 100 µg/L vitamin B_{12}. The bottle that was found to have the most rapid TCE dechlorination activity was re-amended with more TCE and lactate several times till the total TCE amount reached 20 µL. Then, after all TCE was gone, it was sub-cultured into fresh medium under the same conditions. After two generations, the original culture was split into four subcultures that were further enriched using different conditions: 1) higher initial TCE amount (7 µL, ~77 µmol), in order to inhibit methanogens (methanogens are sensitive to high concentrations of TCE); 2) vitamin B_{12} free medium, in order to enrich microorganisms that can provide cobalamin to *Dehalococcoides*. Thus, 4 different subcultures with similar TCE dechlorinating activities under 4 different enrichment conditions were obtained as follows: 1) with methanogenesis and addition of vitamin B_{12} (denoted "MethB12"), 2) with methanogenesis but without addition of vitamin B_{12} (denoted "Meth"), 3) without methanogenesis but with addition of vitamin B_{12} (denoted "NoMethB12"), as well as 4) without methanogenesis or

addition of vitamin B_{12} (denoted "NoMeth"). Within each feeding cycle, for MethB12 and Meth, the amounts of TCE added from the beginning to the end were 2, 5, 5, 5, 3 µL, for a total of 20 µL (~220 mol) and 20, 2, 2, 2, 1 mM lactate (for a total of 27 mM) correspondingly; while for NoMethB12 and NoMeth, they were maintained by adding 7, 7, 6 µL of TCE and 20, 2, 2.5, 2.5 mM of lactate in time series. For each subculture, 5% of the previous culture was inoculated into fresh medium under the same condition. After 8 subcultures, the TCE dechlorination activity and growth of each enrichment culture became stable. The dechlorination products (i.e. cDCE, VC and ethene), organic acids and hydrogen level at different time points during one feeding cycle were measured using GC-FID, HPLC and GC-RGD, respectively. 16S rRNA gene of *Dehalococcoides* and 3 functionally defined RDase genes (i.e. *tceA*, *vcrA*, *bvcA*) were used as biomarkers of microorganisms responsible for TCE dechlorination to VC and ethene. The changes of gene copy numbers of each enrichment culture were determined using quantitative PCR (qPCR). In order to investigate the phylogenetic structure of each enrichment culture, 16S rRNA gene clone libraries were constructed for all four cultures using a universal bacterial 16S rRNA gene primer set (8F and 1492R) and TOPO TA cloning kit (with the pCR2.1-TOPO vector) (Invitrogen, Carlsbad, CA). The identified OTUs were also quantified by qPCR using primer sets designed according to 16S rRNA gene sequences with Genbank accession numbers of JQ004083-JQ004090. In addition to primers, a probe was designed for the OTU closest to *Clostridium* to obtain a specific target. PCR product of plasmid inserts corresponding to each 16S rRNA gene sequence was used as quantitative standards. The copy numbers were normalized to the Dhc_GW copy number in each enrichment.

Chloroethenes and ethene were determined with a gas chromatograph equipped with a flame ionization detector (Agilent Technologies, Santa Clara, CA). 100 µL of headspace was loaded onto a 30-m J&W capillary column with a 0.32-mm inside diameter. Detector and injector temperatures for the gas chromatography were held at 250 and 220 °C, respectively. A gradient temperature program ramped the column temperature from 45 to 200 °C at the maximum rate, and held it at 200 °C for 2 min. Standard curves and Henry's constants were used to calculate total micromoles (µmol) of each compound in culture bottles. Standard curves were generated from individual standards for TCE, cDCE, VC, and ethene that were created by adding known amounts of each compound to 160 ml serum bottles containing 100 ml of sterile medium.

Concentrations of organic acids, such as acetate, lactate, propionate and butyrate were analyzed with a high-performance liquid chromatograph equipped with a UVD 170S UV detector (set to 210 nm). The eluent was 5 mM aqueous H_2SO_4, which was pumped at a flow rate of 0.5 ml/min through an Aminex HPX-87H ion exclusion organic acid analysis column (300 by 7.8 mm; Bio-Rad, Hercules, CA). Aqueous samples (1 mL) withdrawn from the cultures were acidified with 50 μL of 1 M H_2SO_4 and frozen immediately at -20°C. Before analysis, solids were removed from the samples by centrifugation (14,000 x *g*, 10 min) and the supernatants were transferred to autosampler vials for analysis. Five-point calibration curves were established for each analysis.

Results and discussion. All four enrichments were able to dechlorinate TCE to VC and ethene with VC. MethB12 and Meth both completely dechlorinated the first dose of 2 µL (c.a. 22 µmol) TCE in three days (Figure 1 A & B) with a dechlorination rate of ~73 µmol L^{-1} d^{-1}, while NoMethB12 and NoMeth both completely dechlorinated the first dose of 7 µL (c.a. 77 µmol)

TCE in 6 days (Figure 1 C & D) with a dechlorination rate of ~130 µmol L^{-1} d^{-1}. After the first dose of TCE was dechlorinated, subsequent doses were dechlorinated more rapidly, within 1-2 days by all four cultures. For the methanogenic cultures MethB12 and Meth, the total 20 µL (c.a. 220 µmol) TCE fed prior to sub-culturing, was completely dechlorinated to VC and ethene within 14 days, while for the non-methanogenic NoMethB12 and NoMeth, complete dechlorination of 20 µL TCE took 11 days (Figure 1). VC was not completely dechlorinated by these four cultures and the percentages of VC dechlorinated to ethene at the end of each feeding cycle were 5.5 ± 1.4%, 3.0 ± 0.5%, 56 ± 0.5% and 34 ± 1% for MethB12, Meth, NoMethB12 and NoMeth, respectively. In all four enrichments, lactate was fermented to acetate, propionate, butyrate and hydrogen within two days of addition. Acetate and propionate were the dominant fermentation products, and they kept increasing and were built up at the end to 26.9 ± 1.1, 28.5 ± 0.4, 18.1 ± 0.9, 18.0 ± 1.1 mM acetate, 26.2 ± 1.2, 23.7 ± 0.1, 26.7 ± 0.5, 27.0 ± 1.3 mM propionate in MethB12, Meth, NoMethB12 and NoMeth, respectively, while only a small portion of butyrate was detected in non-methanogenic enrichments (0.5-0.6 mM). The fermentation of lactate in methanogenic enrichments (MethB12 and Meth) resulted in an acetate to propionate ratio of 1:1, while 1:1.5 in the non-methanogenic ones (NoMethB12 and NoMeth). Concentrations of aqueous hydrogen reached a peak within 1-2 days in both methanogenic and non-methanogenic cultures (Figure 1 E & F), however, the peak height differed between the Meth and NoMeth conditions (1.2 µM versus 2.2 µM) and hydrogen consumption was more rapid in the Meth cultures. The aqueous hydrogen remained below 0.15 µM after 3 days in the Meth cultures (Figure 1E), while it remained above 0.5 µM throughout the entire feeding cycle for the NoMeth cultures (Figure 1F).

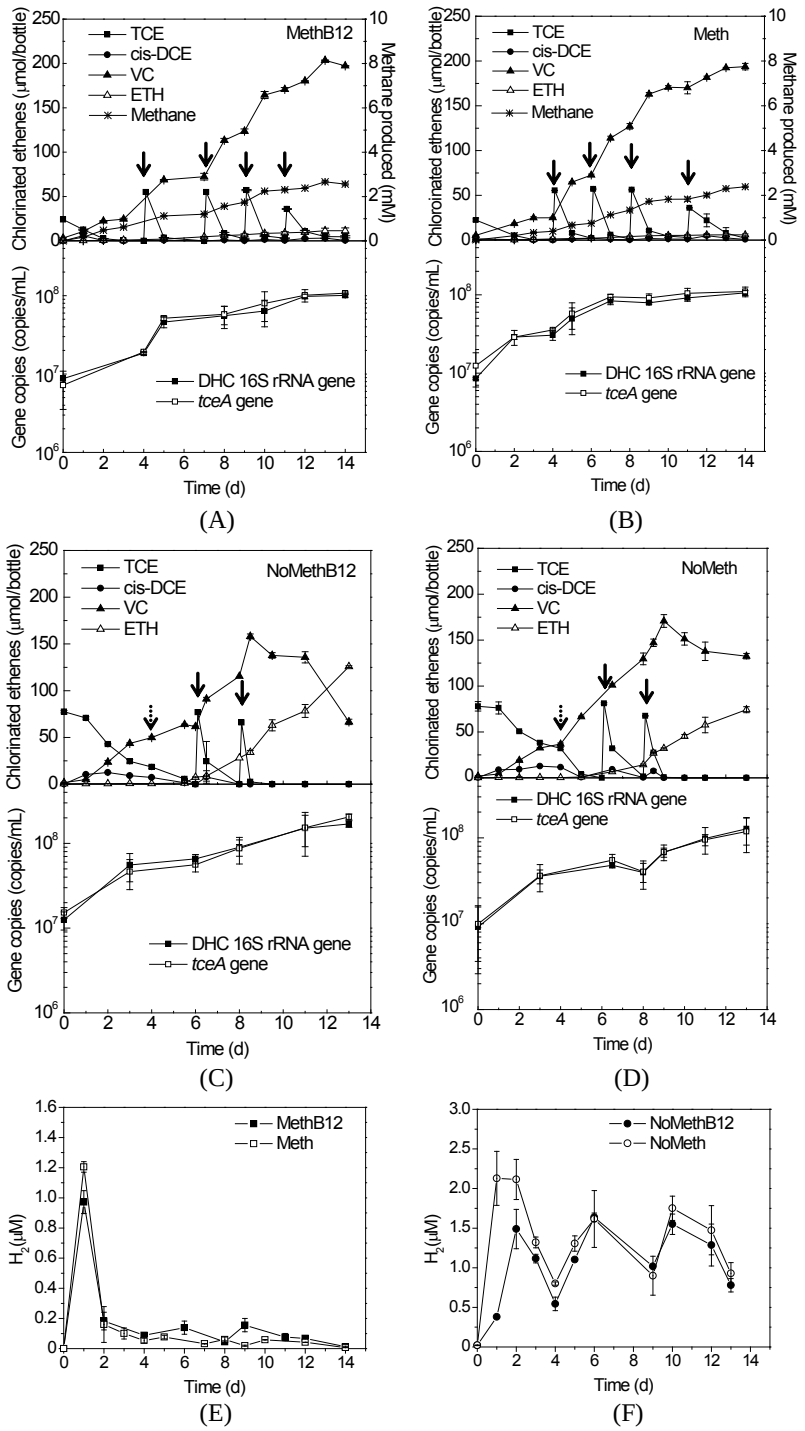


Figure 1. Dechlorination and growth curves of Dhc (A: MethB12; B: Meth; C: NoMethB12; D: NoMeth, ↓ indicates the re-amendment of lactate and TCE, ⇓ indicates the re-amendment of lactate according to the feeding regimes) and H_2 production in the enrichment cultures (E: H_2 in MethB12 and Meth; F: H_2 in NoMethB12 and NoMeth; Note: the H_2 concentration scales are different)

Results of quantitative PCR (qPCR) targeting the 16S rRNA gene of *Dehalococcoides* spp. and 3 functionally-described RDase genes (i.e. *tceA*, *vcrA*, *bvcA*) showed that the

Dehalococcoides 16S rRNA gene and the functional *tceA* gene were present in the 4 enrichments at approximately a 1:1 ratio, while *bvcA* and *vcrA* were not detected. The dechlorination and growth activities of each enrichment culture are shown in Figure 1. The yields of *Dehalococcoides* 16S rRNA gene copies in MethB12, NoMethB12, Meth, and NoMeth were $(2.3 \pm 0.1) \times 10^7$, $(2.7 \pm 0.2) \times 10^7$, $(3.3 \pm 0.3) \times 10^7$, $(2.3 \pm 0.7) \times 10^7$ copies per µmol Cl^- released, respectively.

Table 1. Summary of clone library results of 4 enrichment cultures.

OTUs (NCBI accession number), closest neighbor	Maximum identity to the closest neighbor (query coverage)(%)	% of clones in each library			
		MethB12	Meth	NoMethB12	NoMeth
*Dehalococcoides*_GW (JQ004083), *Dehalococcoides mccartyi* strain 195	99 (99)	8.5	4.2	11.4	16.5
*Pelosinus*_GW (JQ004084), *Pelosinus* sp. UFO1	93 (99)	27.7	59.7	40.4	24.3
*Dendrosporobacter*_GW(JQ004085), *Dendrosporobacter quercicolus* strain DSM 1736(T)	100 (89)	10.6	4.2	11.4	14.8
*Sporotalea*_GW (JQ004090), *Sporotalea propionica* strain TM1	91 (99)	7.4	2.8	3.5	0.88
*Clostridium*_GW (JQ004086), *Clostridium propionicum*	98 (99)	2.1	6.9	8.8	14.8
*Desulfovibrio*_GW (JQ004087), *Desulfovibrio oryzae*	99 (99)	19.1	2.8	N.D[1]	4.3
*Spirochaetes*_GW (JQ004088), *Spirochaetes bacterium* SA-8	99 (97)	3.2	N.D	1.8	4.3
*Bacteroides*_GW (JQ004089), *Bacteroides* sp. strain Z4	99 (99)	N.D	N.D	0.88	4.3
Total clones		94	72	114	115

[1]N.D: not detected

In order to identify supportive microorganisms in these enrichment cultures grown under different conditions, bacterial 16S rRNA gene clone libraries of all these four cultures were constructed. Table 1 gives a summary of bacterial OTUs identified in the four enrichments and their closest phylogenetic affiliates. Unsurprisingly, all four libraries contained OTU with sequence similarity (99%) to Dhc. Besides Dhc, the next four commonly detected OTUs had the highest sequence similarity to *Pelosinus* sp., *Dendrosporobacter quercicolus*, *Sporotalea propionica* and *Clostridium propionicum,* which all belong to the *Firmicutes*. Other than that, OTUs with similarity to the phyla of δ-proteobacteria, *Spirochaetes*, *Bacteroidetes* were also detected, with the closest cultured species of *Desulfovibrio oryzae*, *Spirochaetes bacterium* and *Bacteroides* sp.

qPCR results exhibited that the greatest change among the four enrichments occurred in *Clostridium*_GW whose normalized copy number was 5 times higher in Meth (~ 1×10^7 copies per 10^8 copies of Dhc_GW) compared to MethB12 (~ 2×10^6 copies per 10^8 copies of Dhc_GW) (Figure 2). In the non-methanogenic enrichments, normalized copy numbers of *Clostridium*_GW (0.8-1.0 × 10^8 copies per 10^8 copies of Dhc_GW) were one order of magnitude higher than in both the methanogenic cultures (Figure 2). Normalized copy numbers of *Desulfovibrio*_GW did not exhibit significant difference between the B_{12}-amended enrichments and the unamended, however *Desulfovibrio*_GW was 5-6 times lower in non-methanogenic enrichments than in methanogenic ones. P.D.S were the most dominant members in all enrichments, but their overall numbers were not significantly different between conditions (Figure 2).

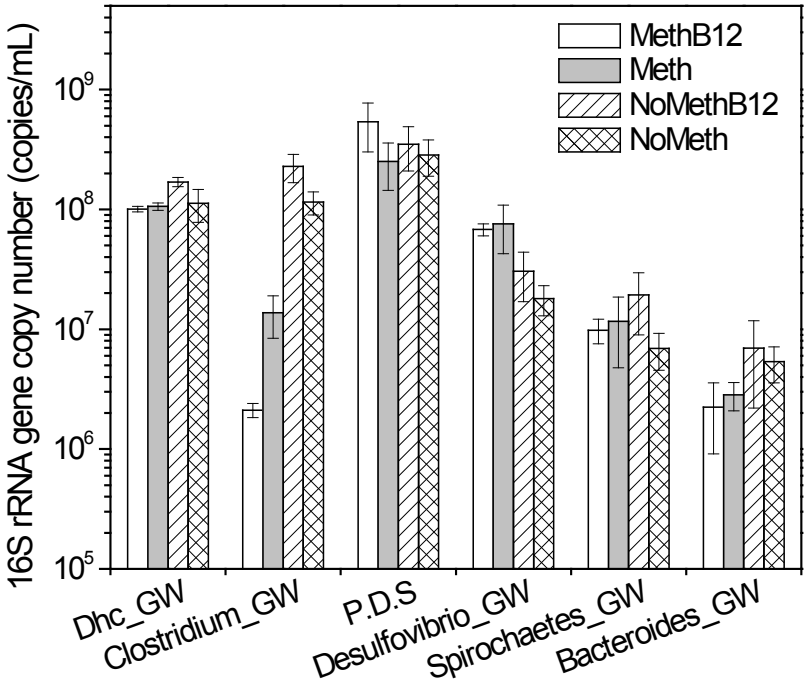


Figure 2 16S rRNA gene copy numbers of dominant OTUs in the four enrichments at the end of a feeding cycle. (Note: P.D.S represents a group of *Pelosinus*_GW, *Dendrosporobacter*_GW and *Sporotalea*_GW)

Principal component analysis (PCA) was applied to 16S rRNA gene copy numbers of different OTUs in the four enrichments (Figure 3) collected over a time course of approximately 15 days. Different feeding-cycle time points of the four enrichments clustered into two groups correlated with the methanogenic condition (MethB12 and Meth vs. NoMethB12 and NoMeth). The PCA results suggest that the methanogenic condition exerted more influence on community structure than B_{12} amendment.

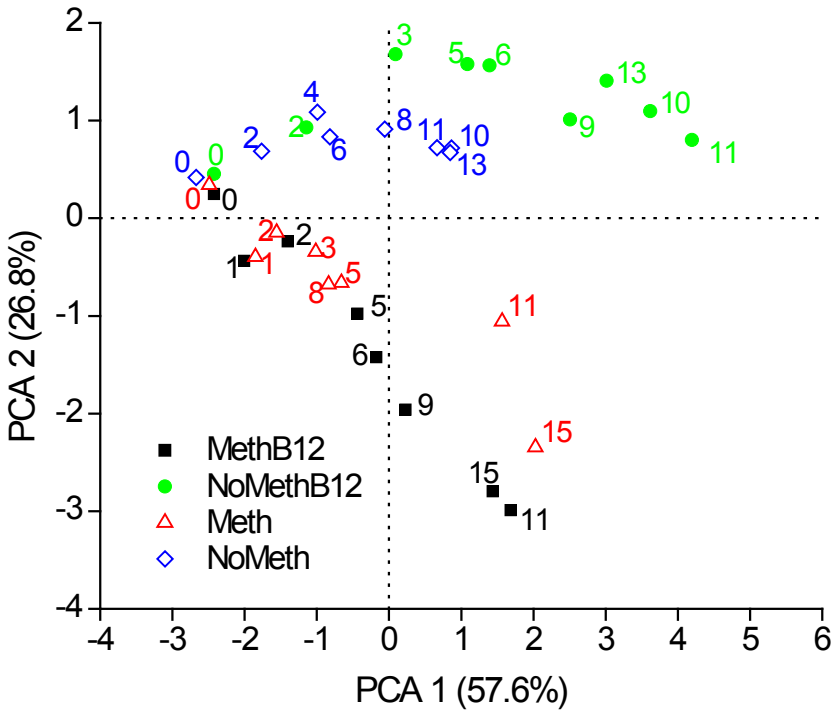


Figure 3. Principal component analysis (PCA) plot based on a time-course 16S rRNA gene copy numbers of 6 OTUs in the 4 enrichments (numbers next to each dot represents incubation days)

Firmicutes (mostly *Clostridia*), δ-proteobacteria (mostly *Desulfovibrio*), *Spirochaetes*, and *Bacteriodetes* that were commonly found in most dechlorinating enrichment cultures were also detected in the four groundwater enrichments (Table 4), while the other three *Firmicutes* OTUs (i.e. *Pelosinus*_GW, *Dendrosporobacter*_GW and *Sporotalea*_GW) detected in this study were rarely reported in dechlorinating microbial communities. The first *Pelosinus* strain was isolated from Fe (III)-reducing enrichments, capable of Fe (III)-reduction only in the presence of a fermentable substrate (40). Its phylogenetic position is within the *Sporomusa–Pectinatus–Selenomonas* phyletic group. *Pelosinus* sp. strain UFO1 was isolated from sediments, characterized as a fermentative bacterium capable of metal transformation (33). Interestingly, the most recent isolated strain *Pelosinus defluvii* sp. nov. was from chlorinated solvent contaminated groundwater by Moe *et al*. (31). In the same study, the authors also proposed to transfer *Sporotalea propionica*, which was first isolated from termite gut (7), to the genus *Pelosinus* as *Pelosinus propionicus* comb. nov. based on phylogenetic, chemotaxonomic, and phenotypic properties. Similar to *Pelosinus* spp., *Dendrosporobacter quercicolus* was reported as a fermentative bacterium belonging to *Sporomusa–Pectinatus–Selenomonas* phyletic group. Therefore, P.D.S reported in this study is representative of two closely related members, *Pelosinus* spp and *Dendrosporobacter* spp. *Clostridium* sp. are known to not only have isomerases such as glutamate mutase, methylmalonyl-CoA mutase and methyltransferase, which require a cobalamin co-factor (5), but they also possess complete up-stream corrinoid biosynthesis pathways (http://www.genome.jp/kegg/pathway.html). Moreover, *Clostridium thermoaceticum* is reported to produce an alternate corrinoid form, 5-methoxybenzimidazolylcobamide as well as cyanocobalamin, and *Clostridium formicoaceticum* can generate corrinoid 5-methoxy-6-methylbenzimidazolylcobamide (34). In addition, sequenced *Desulfovibrio* sp. and *Bacteroides* sp. genomes also have partial up-stream corrinoid

biosynthesis pathways (http://www.genome.jp/kegg/pathway.html). Two other corrinoid forms, guanylcobamide and hypoxanthylcobamide are formed by *Desulfovibrio vulgaris* (19). *Bacteroides* sp. are reported to possess vitamin B_{12}-dependent fermentation pathways for glucose (10). Little is known about whether the other detected species (i.e. *Pelosinus* sp., *Dendrosporobacter* sp. and *Spirochaetes* sp.) are capable of *de novo* corrinoid biosynthesis.

One syntrophic B_{12} relationship that has been previously studied is between algae and bacteria (12). However, few studies have evaluated corrinoid syntrophy between bacterial strains. The survey of 326 algal species by Croft *et al*. (12) revealed that half are B_{12}-dependent and require an exogenous cobalamin source. They provided convincing evidence that cobalamin required by the microalgae were provided by the bacteria. The B_{12}-dependent algae *Amphidinium operculatum* and *Porphyridium purpureum* were both able to grow with *Halomonas* sp. capable of *de novo* cobalamin synthesis. Moreover, their study revealed enhanced growth and up-regulated B_{12} biosynthesis in *Halomonas* sp. in the presence of algal extracts. In this study, enhanced growth was observed for *Clostridium*_GW in the methanogenic enrichment without B_{12} amendment (Meth) compared to in the methanogenic, B_{12}-amended enrichment (MethB12). This suggests that *Clostridium* sp. might contribute to providing corrinoids to the methanogenic enrichment.

These enrichment cultures serve as representative study materials, which were applied to different molecular and analytical tools in some of the following objectives, in order to identify biomarkers related to corrinoid salvaging by *Dehalococcoides* and corrinoid production by other supportive microorganisms.

1.2 Phylogenetic microarray analysis of a microbial community performing reductive dechlorination at a TCE-contaminated site (This study is published in Lee et al., Appl. Environ. Microbiol., 2008 & Lee et al., Eviron. Sci. & Technol., 2011)

Materials and methods. We performed phylogenetic microarray analysis to monitor dynamic changes of microbial community in the groundwater of the TCE-dense non-aqueous-phase liquid (DNAPL) site (Ft. Lewis, WA) that was sequentially subjected to biostimulation and bioaugmentation treatments. The site was located within the area designated NAPL Area3 (Figure 4). Each approximately 40-foot by 20-foot treatment plot was constructed with an injection, extraction, and four monitoring wells (Figure 4). Periodic whey powder injections were formed as part of the treatment strategy within the two treatment plots (Table 2).

The 16S rRNA microarray that was being applied in this research was developed using the Affymetrix GeneChip photolithography platform and targeted both the Bacterial and Archaeal domains. The array consists of 9,000 sets of 25-mer oligonucleotide probes; each set is specific for one 16S rRNA gene of a particular species or group of related species, referred to as an Operational Taxonomic Unit (OTU).

Two treatment plots (approximately 40-foot by 20-foot each) were constructed within the site to evaluate enhanced bioremediation using different treatment strategies. Groundwater samples were collected from monitoring wells on each treatment plot for chemical and molecular analysis prior to biostimulation with whey amendment as a baseline (April 2005), one month

following whey amendment (July 2005), one month following bioaugmentation (August 2005), four months following bioaugmentation and modification of whey amendment (November 2005), seven months following bioaugmentation (February 2006), and finally nine months following bioaugmentation and 2 months after cessation of whey amendments (April 2006).

Both DNA and RNA of the groundwater microorganisms were collected for further analysis. The collection strategies are described in details here. Groundwater samples intended for genomic DNA (gDNA) isolation were collected from monitoring wells (MW) into autoclaved 1-liter bottles and were shipped overnight on ice to the laboratory at the University of California at Berkeley. Samples were stored at 4°C upon arrival and processed within 48 h. Between 100 ml and 1,000 ml of groundwater was filtered through a 0.2-µm, surfactant-free cellulose acetate filter (Fisher, Houston, TX) that was subsequently vortexed vigorously for 5 min in 2 ml of the sample groundwater in order to dislodge filtered particles. The resultant liquid was centrifuged at 21,000 x g at 4°C for 10 min, the supernatant was discarded, and the cell pellet was stored at –80°C until extraction. The gDNA was isolated from frozen cell pellets using the UltraClean microbial DNA isolation kit (Mo Bio Laboratories, Carlsbad, CA) according to the manufacturer's instructions. A negative control of 0.2-µm-filtered and autoclaved Milli-Q water was extracted in parallel with each set of samples. The gDNA concentration was quantified using the PicoGreen double-stranded-DNA quantitation kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions.

Groundwater samples intended for RNA isolation were filtered on site at the time of sampling. Between 275 ml and 8 liters of groundwater was pressured through a sterile 142-mm-diameter, 0.22-µm hydrophilic polyvinylidene fluoride filter (Millipore, Billerica, MA) via a tripod collection system. The filtration process typically took less than 5 min to prevent RNA degradation. The filters were then placed into sterile 50-ml conical tubes and immediately frozen on dry ice. Samples were shipped overnight on dry ice to the University of California at Berkeley and upon arrival stored at –80°C until extraction. Studies have shown that overnight shipping on dry ice is effective in preserving RNA and that this method did result in significant RNA degradation when samples shipped overnight on dry ice were compared to samples treated on site (1, 21). Total RNA was isolated from the frozen filters using the RNA PowerSoil total RNA isolation kit (Mo Bio Laboratories, Carlsbad, CA) according to the manufacturer's instructions. Prior to the first step of the extraction procedure, the filters in the 50-ml conical tubes were cut into smaller pieces with scissors while the tubes were submerged in liquid nitrogen. After RNA isolation, contaminating DNA was removed by applying two successive DNase treatments via the DNA-free kit (Ambion, Austin, TX) according to the manufacturer's recommendation. The purified total RNA concentration was quantified with the RiboGreen RNA quantification kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. In order to deal with the limitations inherent in imperfect DNA and RNA extraction efficiencies in this study, a single method was applied throughout the study (by the same operator for consistency) and results were compared over time for relative interpretation rather than absolute quantification.

Extracted DNA and RNA were fragmented into 50 – 200 base pair fragments, labeled with biotin, and hybridized to custom-made Affymetrix 16S-rRNA microarrays according to the manufacturer’s recommended protocols. After the overnight incubation at 60 rpm at 48°C and

50°C for DNA and RNA, respectively, chips were washed and stained according to standard Affymetrix protocols. Each chip was scanned and recorded as a pixel image, and initial data acquisition and intensity determination were performed using GeneChip microarray analysis suite software (version 5.1, Affymetrix). Each OTU is represented by a set of 11 or more target probe pairs. One probe in the pair is referred to as the perfect match (PM) probe because it has a sequence of a representative member of the target OTU and the sequence is dissimilar from members outside of the OTU. The second probe in the probe pair is a control probe, referred to as the mismatch probe, identical in sequence to the perfect match probe in all but one nucleotide of the 25-mer sequence. A probe pair is scored as positive if the intensity of the perfect match probe is significantly higher than that of the mismatch probe (PM/MM>1.3 and PM-MM>130N). An OTU is considered present when at least 90% of its probe pairs were positive. The OTUs from the DNA hybridized arrays were used to identify the microbial composition in groundwater and the OTUs from the RNA hybridized arrays were compared among different treatment phases in order to identify the microbial population that enriched from enhanced bioremediation.

Chlorinated ethene concentrations were measured in groundwater samples collected from eight sampling locations throughout each treatment plot during the enhanced bioremediation demonstration conducted at the site. Molar concentrations of TCE, cDCE, VC, and ethene measured in groundwater samples are presented in Figure 4. Three sampling events were conducted prior to whey amendment, two or three events were conducted in subsequent phases of whey amendment, and two sampling events were conducted 1 and 2 months following cessation of whey injections. Because of poor mass balances in the open system, the temporal changes in chlorinated ethenes were represented as a percentage of the total ethenes.

Results and discussion. The predominant chlorinated ethene observed during the three baseline sampling events (March to April 2005) conducted prior to whey injection was TCE, representing an average range of 78 to 82% of the molar concentration in plot 1 (Figure 5) and 68 to 69% of the molar concentration in plot 2 (Figure 5). No VC or ethene was detected in either treatment plot. In addition, the average chlorinated ethene concentration was a factor of 2.9 higher in plot 2 (202 ± 26 µM) than in plot 1 (69 ± 23 µM). In order to stimulate more reducing conditions that would favor reductive dechlorination, whey was injected into the subsurface as an exogenous electron donor for biostimulation beginning in June 2005. One month following 30 g/liter whey injection (July 2005), the molar percentage of chlorinated ethenes in the form of cDCE increased significantly in plot 1 (88%) and plot 2 (96%). Immediately following the July 2005 sampling, each treatment plot was bioaugmented with 18 liters of a *Dehalococcoides*-containing enrichment culture along with injection of 100 g/liter whey into plot 1 and 10 g/liter whey into plot 2. One month following these injections (August 2005), the molar percentage of chlorinated ethenes in the form of cDCE again increased, to an average of 94% in plot 1 and 98% in plot 2, a trend that continued for 4 months after bioaugmentation (November 2005), with cDCE reaching 99% of the molar percentage in both plots 1 and 2. In plot 1, no significant VC or ethene was observed during this sampling, but in plot 2, quantifiable VC was observed, representing 0.3% of the molar concentration.

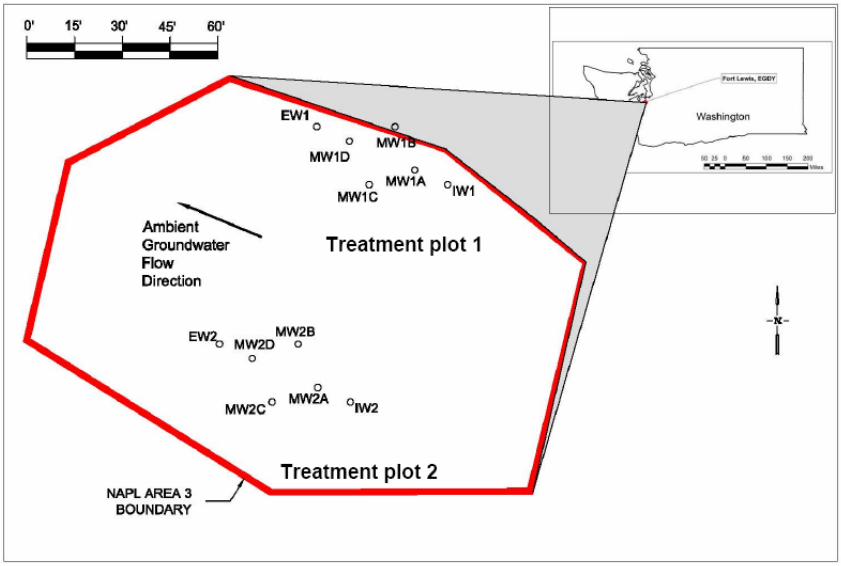


Figure 4. A schematic of the treatment plots at Ft. Lewis EGDY (Tacoma, WA). MW = monitoring well; IW = injection well; EW = extraction well.

Table 2. Dates and concentration of whey injected at treatment plots 1 and 2.

Date	Treatment plot 1		Treatment plot 2	
	Vol of water (gallons)	Concn of whey[b] (g/liter)	Vol of water (gallons)	Concn of whey (g/liter)
19 June 2005	3,200	30	3,900	30
26 June 2005	3,200	30	3,200	30
July 2005	1,700	100	4,000	10
Aug. 2005	0[a]	0[a]	1,800	10
Sept. 2005	1,700	100	4,000	10
Oct. 2005	1,900	100	1,800	10
Nov. 2005	1,800	10	1,800	100
Dec. 2005	1,800	10	1,800	100
Jan. 2006	1,800	10	1,800	100
Feb. 2006	1,300	10	1,800	100

[a] No recirculation or injection of whey due to equipment difficulties.
[b] Whey was a complex substrate with ~10 to 13% protein and ~70 to 75% lactose

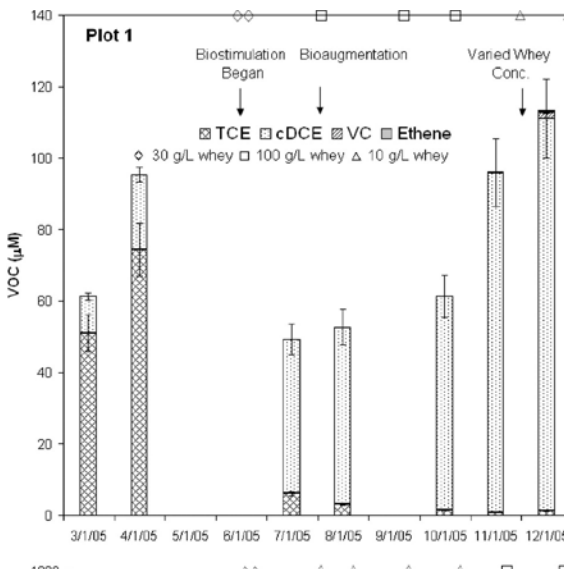


Figure 5. Molar concentrations of chlorinated ethenes and ethene in groundwater samples over the 1-year monitoring period for treatment plots 1 and 2. Concentrations are reported as averages for eight sampling locations within each treatment plot, and error bars represent analytical error.

The date and concentration of whey injected into the respective treatment plots are indicated on the top axis, and manipulations implemented at the site are indicated by arrows on the graphs. For clarity, only one of the two sampling events performed within 2 weeks of each other during March 2005 is plotted on the graph.

Approximately 5 months after the initial whey injection (November 2005), the whey injection strategy was altered so that plot 1 received 10 g/liter whey injections while plot 2 received 100-g/liter whey injections. Three months after this operational change and 7 months after bioaugmentation (February 2006), significant increases in VC and ethene were observed, with VC and ethene in both treatment plots representing an average of 24% and 4%, respectively, of the molar concentration in plot 1 and 12% and 1%, respectively, of the molar concentration in plot 2. The increase in VC and ethene fractions continued even after cessation of whey

injections, with VC and ethene representing an average of 34% and 10%, respectively, of the molar concentration in plot 1 and 36% and 4%, respectively, of the molar concentration in plot 2 during the last sampling event (April 2006). Up to the April 2006 sampling event, residual carbon persisted within the treatment area (as evidenced by chemical oxygen demand concentrations) were measuring between 67 to 538 mg/liter in plot 1 and 183 to 821 mg/liter in plot 2, indicating that the electron donor was available for reductive dechlorination.

In order to identify active bacterial members in the microbial community, a clone library was constructed based on the expressed 16S rRNA genes. Total RNA collected from selected monitoring wells during the February 2006 sampling was pooled and two 0.24 μg-samples were reverse-transcribed in parallel using the SuperScript III First-Strand synthesis system (Invitrogen, Carlsbad, CA) as previously described. Following reverse-transcription (RT), the two cDNA samples were combined and 2 μL of cDNA products were amplified in eight separate reactions with annealing temperatures between 48 to 58 °C using the universal bacterial primers. The amplified products were combined and the 16S rRNA band was quantified. A parallel no reverse-transcriptase sample was prepared and no band was visible on the gel electrophoresis after RT-PCR, indicating no DNA contamination. The clone library of the 16S rRNA gene PCR products was constructed with the TOPO TA cloning kit (with the pCR2.1- TOPO vector) (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. A total of 384 transformants were randomly picked and inserts were sequenced bidirectionally using M13-vector specific primers at the University of California, Berkeley sequencing center. Paired-end sequencing reads were vector trimmed, assembled, and quality-checked via the Joint Genome Institute's GeneLib software package (Kirton, E. unpublished data).

Besides clone library construction, phylogenetic microarray was also applied to analyze the microbial community in the groundwater samples collected from both treatment plots. A total of ten PhyloChips were analyzed in this study, eight for DNA and two for RNA. Technical replication on the PhyloChip has an average coefficient of variation of 10% for an OTU across arrays, indicating low variation and high reproducibility, and technical replicates were therefore not analyzed here. DNA-PhyloChip analyses were performed on the five samples collected within the treatment plot (July 05, August 05, November 05, February 06, and April 06), the upstream (background) and downstream samples, and the *Dehalococcoides*-containing enrichment culture used for bioaugmentation. An insufficient mass of PCR products was generated from the baseline sample collected within the treatment plot prior to biostimulation (April 05) to analyze. RNA-PhyloChip analyses were performed on the samples collected during the February 2006 and April 2006 sampling periods. Throughout this study, the samples were referred to by their date or location with "April" referring to the 2006 sample and "FTB" referring to the bioaugmentation culture.

The 237 nonchimeric sequences were grouped into 102 OTUs with 97% sequence homology across 14 phyla (Figure 1), indicating a broad diversity and supporting the richness detected via PhyloChip (Table 1). In the clone library, a majority of the phylotypes (60%) belonged to the four classes of Proteobacteria with β-Proteobacteria being the disproportionately dominant group (Figure 6). A relatively large number of phylotypes were also identified from the phyla *Firmicutes* and *Bacteroidetes*. Two phylotypes from the Phylum *Chloroflexi* were

identified, but neither was identified as *Dehalococcoides* sequences (closest database matches were to uncultured environmental sequences).

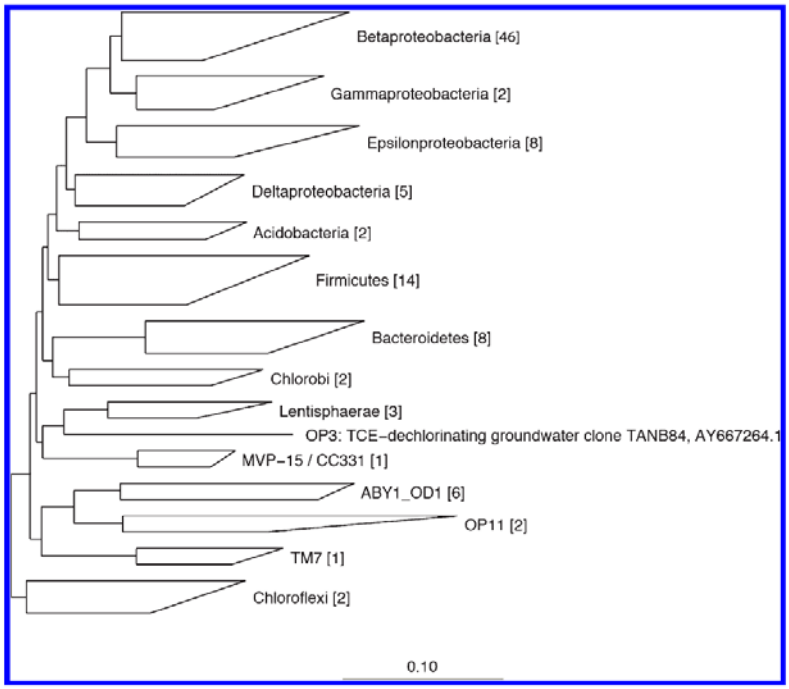


Figure 6. Phylogenetic tree of the expressed 16S rRNA sequences. Neighbor-joining algorithm with the filter lanemaskPH was used for phylogenetic analysis in order to exclude highly variable positions of the alignment. A total of 102 phylotypes were found (number in brackets) based on 97% sequence homology. Scale bar indicates 10% estimated sequence divergence.

PhyloChip analyses of the DNA in groundwater samples from the Ft. Lewis site were used to comprehensively evaluate the phylotypes represented there, including those present at low abundances. With 8741 bacterial and archaeal taxa on the G2 PhyloChip, over 1300 OTUs were positively detected in DNA from the groundwater, with a similar richness detected in the laboratory-grown bioaugmentation culture containing *Dehalococcoides*. While analyzing a microbial community via DNA reveals the overall genetic diversity and relative abundance of the constituent OTUs, the metabolic activity of the community members are expected to differ. The activity of different populations can be inferred from the expressed 16S rRNA. When RNA from groundwater samples were hybridized onto the PhyloChip to analyze for the expressed 16S rRNA genes, ~600 OTUs or ~180 subfamilies were detected in the February or April samplings. There were sequences from five bacterial phyla that were detected in the RNA fraction but not in the DNA fraction, but only one subfamily was detected for each of these phyla.

Principle component analysis (PCA) of the DNA hybridization intensity (HybScore) that included any subfamilies that were positively detected in at least one of the samples showed that the eight analyzed samples were unique and that 66% of the data set variance could be explained by two axes (Figure 7). The two communities that were closely associated were the background sample and the sample collected downstream of the treatment plot, outside the zone of direct whey influence, highlighting the significant influence of exogenous carbon amendment on the microbial community. The time series samples that were collected within the treatment plot

showed that over the course of treatment, the community structure was changing continuously, with separation along both axes on the PCA plot relative to the July sample (Figure 7). Injection of the laboratory-grown bioaugmentation culture did not result in significant immediate changes as the July and August samples were relatively similar to each other according to the PCA plot. Further, the PCA plot shows that the bioaugmentation culture was different from any of the samples collected at Ft. Lewis throughout the study.

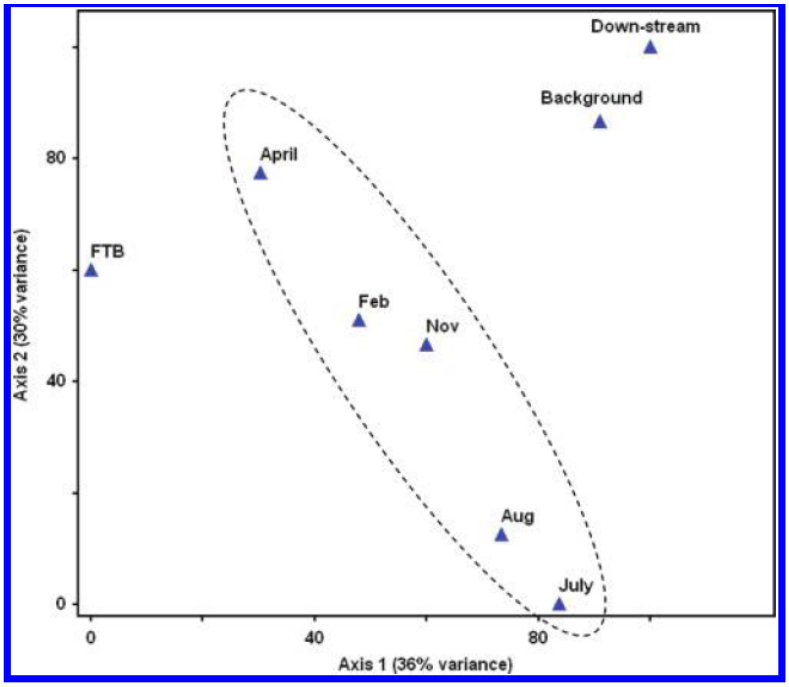


Figure 7. Principle component analysis (PCA) of the PhyloChipanalyzed samples. The dotted circle highlights the samples that were collected within the treatment plot at different time points.

The immediate responses of the microbial community to biostimulation can be seen in changes in HybScores between the background sample and the July treatment plot sample, the first sampling period after whey injection (Figure 8). Changes in HybScores are positively correlated to relative abundance, with estimates that a 1000-unit change in HybScore is roughly proportional to a 10-fold change in abundance (8). Of the 393 subfamilies that responded to whey, 193 showed changes in HybScore of 500 units or more with 121 and 72 subfamilies showing increases and decreases, respectively. Subfamilies from *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Cyanobacteria*, *Firmicutes*, and all classes of Proteobacteria accounted for 76% of this dynamic subset. Specifically, bacteria in the *Bacteroidetes* and *Firmicutes* and those with diverse metabolic capabilities to utilize electron acceptors in the δ-Proteobacteria and ε-Proteobacteria were enriched.

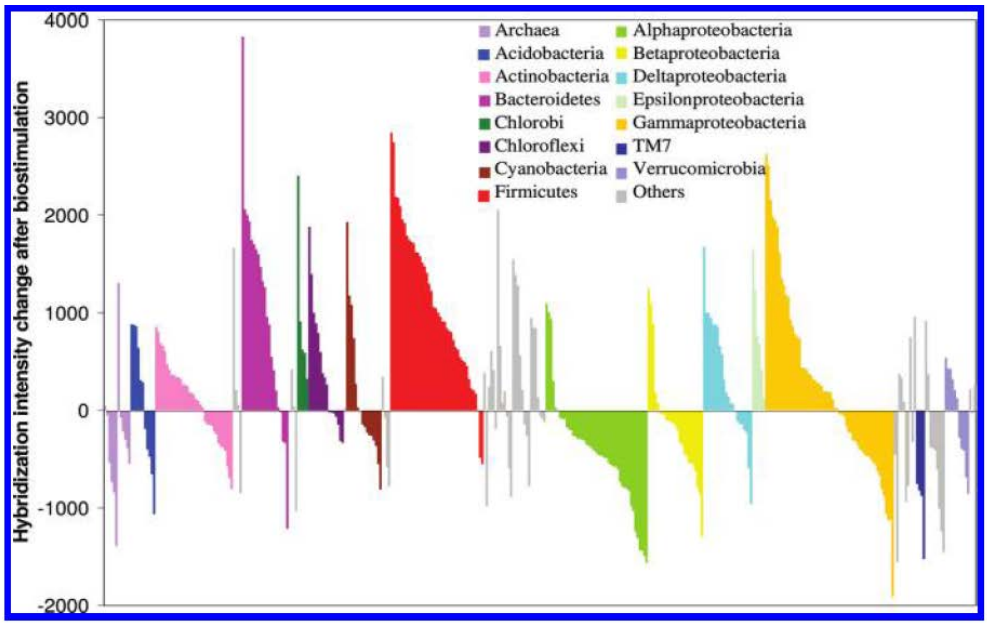


Figure 8. Changes in hybridization intensity relative to the background sample of the 393 subfamilies immediately after biostimulation with whey (July 2005) in the treatment plot. Phyla are color-coded and ordered alphabetically from left to right starting with the archaeal domain followed by the bacterial domain. Each bar represents a subfamily with positive bars indicating subfamilies that increased in abundance relative to the background after receiving whey and negative bars represent subfamilies that decreased in abundance. A 1000-unit change in hybridization intensity is equivalent to a 10-fold change in relative abundance.

Clustering analysis of the HybScores of each of the samples indicates that the bacterial populations within the treatment plot after biostimulation with whey were significantly different from the background sample, with the latter forming a separate cluster (Figure 9, top-axis). As is consistent with the PCA analysis, the July and August samples form a cluster that is separate from the three later samples. Overall, clustering analysis of all the 478 bacterial subfamilies found during the course of treatment identified three distinct groups (Figure 9, y-axis).

Cluster group 1 consisted of sequences from 157 subfamilies from 26 phyla whose amplicons exhibited a high relative abundance in the background sample followed by a sharp decline upon whey injection that continued over the course of treatment (Figure 9). The dominant members in cluster group 1 were from different classes of Proteobacteria (88 subfamilies), especially the α-Proteobacteria (36 subfamilies). Members from the orders of Bradyrhizobiales, Consistiales and Rhizobiales within the α-Proteobacteria, Burkholderiales within the β-Proteobacteria, and Legionellales and Thiotrichales within the γ-Proteobacteria were within this cluster.

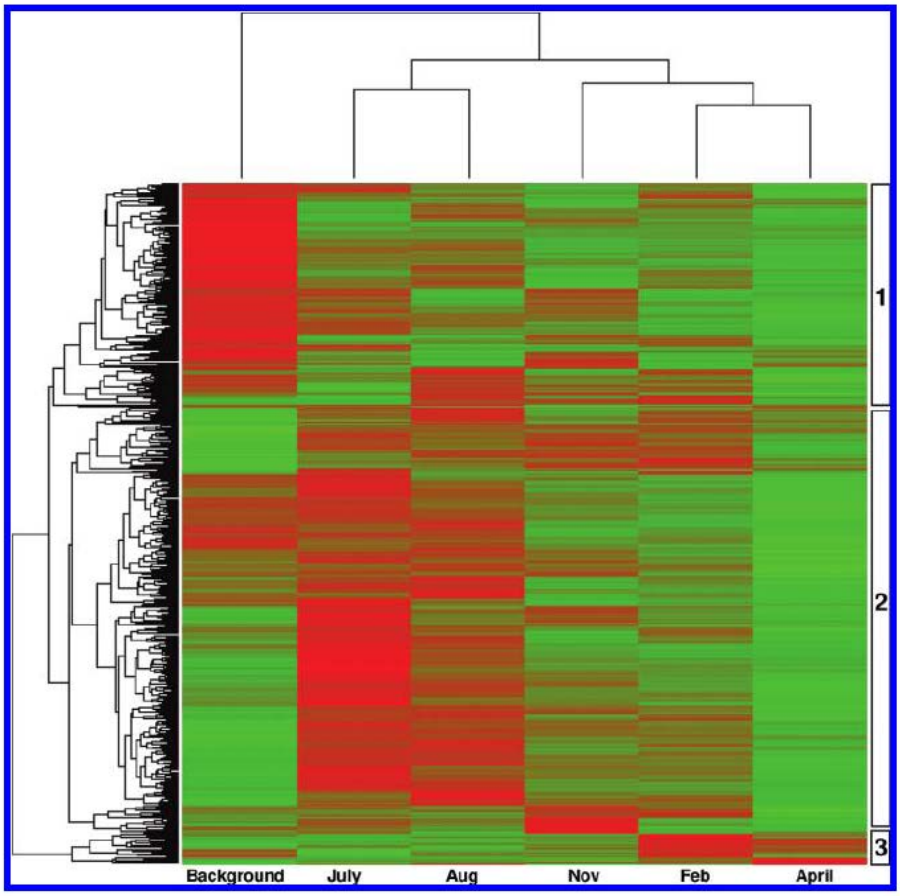


Figure 9. Hierarchical clustering analysis of samples and subfamilies in the bacterial domain over the course of treatment. The color gradient from green to red of the heatmap represents increasing array hybridization intensity. Each row represents a subfamily and each column represents a sample with labeling at the bottom. Three main dynamic groups were identified and labeled on the right.

Cluster group 2 is the largest group consisting of 300 subfamilies distributed over 35 phyla. This cluster had a low relative abundance initially, followed in the first month by an increase then faded toward the end, indicating that these bacteria were initially stimulated by whey but were not sustained at higher relative abundances over the course of treatment (Figure 4). The dominant members in cluster group 2 were from Actinobacteria (24 subfamilies), Bacteroidetes (25 subfamilies), Firmicutes (66 subfamilies), γ-Proteobacteria (40 subfamilies), and δ-Proteobacteria (25 subfamilies). Fermenters within the orders of Bacteroidales, Flavobacteriales, and Sphingobacteriales in the Bacteroidetes along with Bacillales, Clostridiales, and Lactobacillales in the Firmicutes were within this cluster. Bacteria in the families of Enterobacteriaceae and Pseudomonadaceae within γ-Proteobacteria and Desulfoarculaceae, Geobacteraceae, and Nitrospinaceae in the δ-Proteobacteria were also in cluster group 2.

Cluster group 3 is relatively small with only 21 subfamilies and they remained at a low relative abundance until the February sampling when significant increases were observed (Figure 4). Members in the Campylobacteraceae family within the ε-Proteobacteria and the Desulfobulbaceae family within the δ-Proteobacteria were representatives of this group. Interestingly, three subfamilies within the candidate phyla OP11 were also in cluster group 3.

Cluster group 3 was the only set of bacterial subfamilies that were present at a high density relative to the background toward the end of treatment.

As observed, members within the five classes of Proteobacteria (α, β, δ, ε, γ) were a dynamic component of the microbial community. Members of Proteobacteria tended to show mixed responses after the injection of whey (Figure 3) in contrast with the Bacteroidetes and Firmicutes, which tended to show significant increases in all subfamilies. In general, a great variety of physiology and metabolism are found within the Proteobacteria phylum, and many are typical heterotrophic soil microbes that can respire different terminal electron acceptors for growth (oxygen, nitrate, sulfate, iron) . Significant enrichment in some of the Proteobacteria subfamilies was concomitant with the decreases in dissolved oxygen, nitrate, and sulfate concentrations as well as the generation of ferrous iron after whey injection, suggesting that the diverse microbes were respiring the available terminal electron acceptors. Perhaps because the injected whey contained a background level of sulfate, sulfate-reducers from families such as Desulfobacteraceae, Desulfohalobiaceae, Desulfomicrobiaceae, and Desulfuromonaceae in the δ-Proteobacteria became more active toward the end of treatment.

After the initial injection of whey, only a small fraction of Archaea showed responses (Figure 10a). A hierarchical clustering analysis of all the 27 archaeal subfamilies detected in the DNA fraction showed that most of the Archaea remained at a relatively low abundance during the first four months of operation (up to August), but increased significantly from November to April (Figure 10a). Cluster group 1 contained many methanogens from the Euryarchaeota phylum within the Methanosaetaceae, Methanomicrobiaceae, Methanosarcinaceae, Methanocorpusculaceae, and Methanobacteriaceae families. This cluster exhibited significant increases toward the later part of treatment (Figure 10a), corresponding with over 2 orders of magnitude increases in methane concentrations detected in the later samples (Figure 10b) from less than 10 μg/L prior to November to over 7500 μg/L by the end of treatment, and demonstrating the strong correlation between PhyloChip results and corresponding community metabolism. Other members in cluster group 1 included representatives in the families of Thermococcaceae and Halobacteriaceae. Cluster groups 2 and 3 represent a relatively small fraction of the archaeal populations from the C1 and Thermoprotei classes of Crenarchaeota that decreased in relative abundance over time (Figure 10a). Methanogens have commonly been observed in dechlorinating microbial communities (28, 36) and can be physically collocated with *Dehalococcoides* cells as biofloc (38). Although previous research has shown that hydrogen-consuming methanogens are potential competitors of *Dehalococcoides* spp. for hydrogen (47), some methanogens (e.g., *Methanosarcina* spp.) are known to synthesize corrinoids (30) that might benefit *Dehalococcoides* spp. as important cofactors for RDases. Interestingly, at this contaminated site, the apparent delay in methanogenesis might have been caused by the lower pH (5.2) during the first few months of operation.

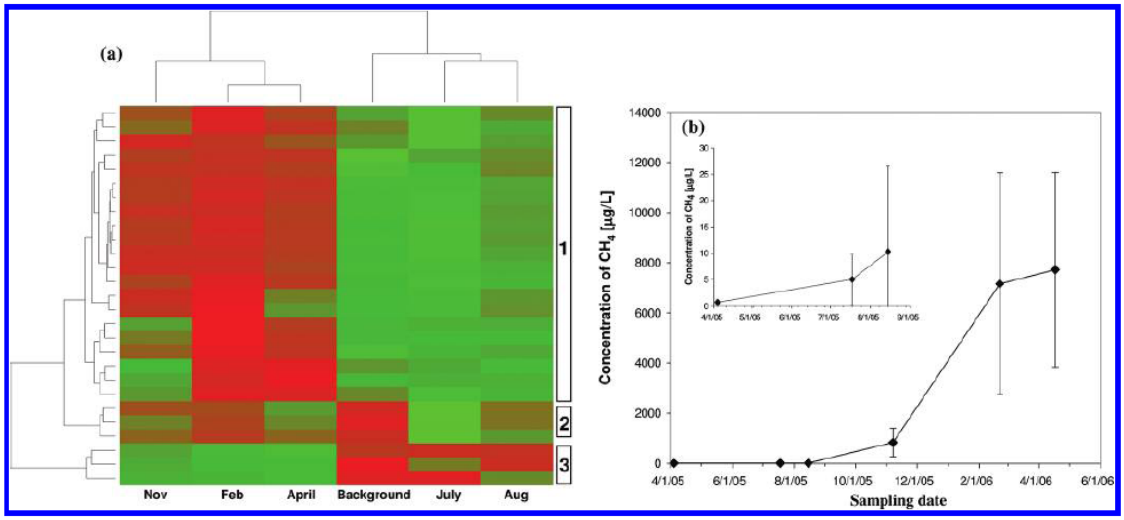


Figure 10. (a) Hierarchical clustering analysis of samples and subfamilies in the archaeal domain over the course of treatment. The color gradient from green to red of the heatmap represents increasing array hybridization intensity. Each row represents a subfamily and each column represents a sample with labeling at the bottom. Three main dynamic groups were identified and labeled on the right. (b) Methane concentration over the course of treatment at Ft. Lewis. The inserted graph highlights the differences in scale during the early part of treatment. The concentration at each time point is the average measurement of the 16 samples taken from monitoring wells that were spatially separated and screened to different depths on both treatment plots. The error bars represent standard deviation of the 16 concentration measurements.

Overall, the combination of PhyloChip analyses of DNA and RNA, together with clone library construction performed in this study have provided insights into the in situ microbial ecology and population dynamics at the TCE-contaminated field site undergoing biostimulation and bioaugmentation. Whey was injected into the treatment plots in this study to stimulate the activity of the subsurface microbial community. Of concern initially was the fact that the site was originally aerobic, but based on the PhyloChip results, diverse groups of microorganisms responded and their metabolic activities resulted in a favorable reducing condition for reductive dechlorination. Furthermore, the generated fermentation products supported growth of *Dehalococcoides*. These results are consistent with previously reported data which showed that *Dehalococcoides* concentrations increased following biostimulation while TCE was converted to mainly cDCE. Subsequently, during the later part of the treatment, VC and ethene were formed, and further increases in *Dehalococcoides* were observed. By using data from advanced molecular tools such as PhyloChips, practitioners will be able to obtain a comprehensive time-series view of the subsurface microbial community. Such data would complement qPCR results targeting specific key functional dechlorinators (such as *Dehalococcoides*). Ultimately, by tracking the overall microbial community together with key functional players, informed decisions can then be made regarding how to best manipulate the field conditions to achieve effective bioremediation of chlorinated ethenes.

1.3 Construction and operation of the continuous-flow chemostat

Materials and methods. A TCE-dechlorinating chemostat with 3 L operating volume and without headspace was constructed and operated under continuous flow conditions, being fed by 9.74 mM TCE at a mean cell residence time of 45 days (Figure 11). The chemostat is fed with excess lactate (60 mM) to achieve complete dechlorination of TCE to ethene under steady-state conditions. At the very beginning, 5% of the ANAS enrichment culture, which was continuously maintained in a sequence batch reactor, was inoculated into the chemostat. During the first stage (0-55d) after inoculation, methane production was observed after 18 days, and ethene, the benign dechlorination product of TCE, was detected after 25 days. The reactor was then being operated under batch reactor condition for one SRT and it was able to successfully dechlorinate TCE all the way to ethene. Then a series of step-wise transitions was performed on the chemostat, during which the influent and effluent flow rates were gradually increased from 0 mL/d to 50 mL/d. After 1.5 SRT of intermittent flow, the reactor transitioned to fully continuous flow, under which the chemostat was being operated for 24 hours with a flow rate of 50 mL/d (Figure 12).

Results and discussion. After the initial start-up process, we investigated the reductive dechlorination characteristics of the continuous-flow chemostat under steady-state operation. The continuously added TCE was completely dechlorinated to ethene at a dechlorination rate of 100 µmol TCE/ (10^9cells·d). Approximately 40 mM propionate was produced as a fermentation product of lactate added in excess. *Dehalococcoides* species were present as the main dechlorinators in the chemostat, and two well-characterized reductive dehalogenase (RDase) genes (*tceA* and *vcrA*) were found as the main functional genes. By applying quantitative PCR, we observed that under steady-state condition, the copies of *Dehalococcoides* 16S rRNA gene, *tceA* and *vcrA* in the chemostat were 1.3×10^9 copies/ml, 8×10^8 copies/ml, and 5.7×10^7 copies/ml, respectively.

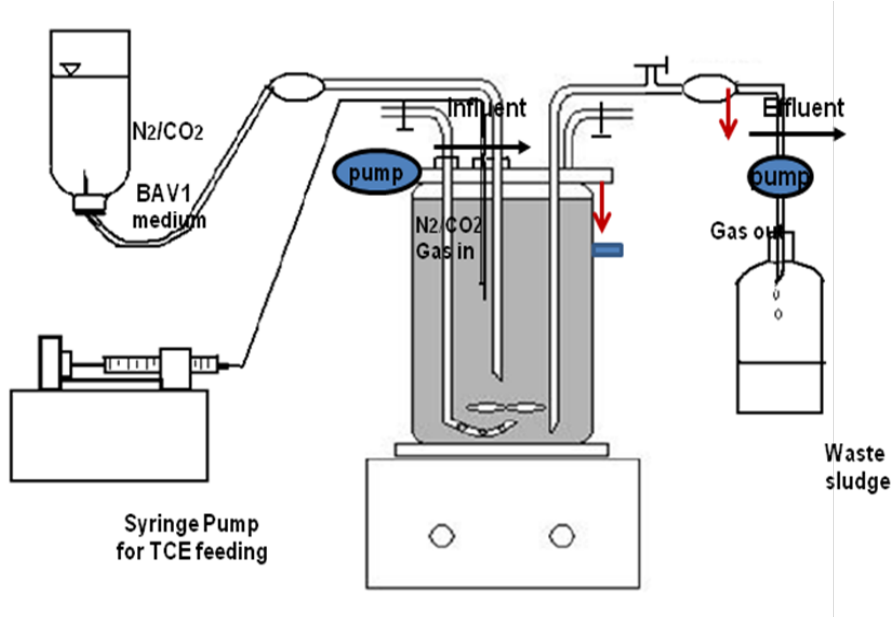


Figure 11. The scheme of continuous-flow chemostat.

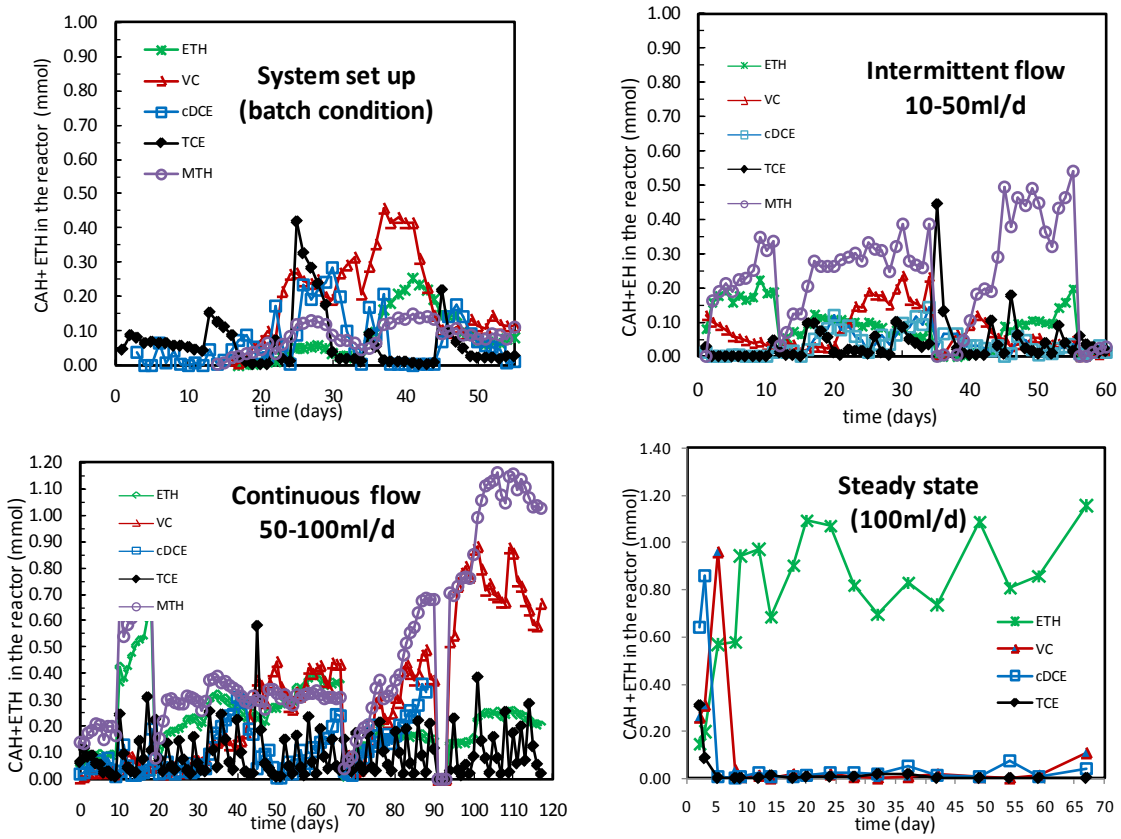


Figure 12. Dechlorination performance of different operating stages during the start-up of the chemostat.

The operation of this continuous-flow chemostat will allow us to conduct investigations in a flow-through system more like the contaminated field site, thus to extend our understanding of reductive dechlorination to a more practical prospect in the future.

Objective 2: Construction of defined consortia containing *Dehalococcoides* and identification of rRNA-based phylogenetic biomarkers. (Part of this study was published in Men *et al.*, *ISME J*, 2012)

We have constructed defined microbial consortia containing *D. mccartyi* strain 195 as a key member and one or two potentially function-supporting microorganisms. The candidate partners for strain 195 were selected primarily based on two criteria: 1) partners who can supply strain 195 with its sole electron donor hydrogen, and 2) partners who can produce essential growth factors like cobalamin (Table 3).

Materials and method. We conducted studies with constructed co-cultures that contain *D. mccartyi* 195 and various organisms, such as a species capable of generating hydrogen via fermentation of organic substrates, and a species able to produce vitamins and cofactors. The targeted fermenting species are those that have been found in high frequency in dechlorinating enrichment and environmental samples, e.g., *Desulfovibrio*, *Clostridium*, *Eubacterium*, *Bacteroides*, and *Syntrophomonas* species. The targeted vitamin/cofactor-producing species are

those have also been frequently detected in dechlorinating microbial communities, e.g. *Acetobacterium, Methanobacterium, Methanococcus, Methanosarcina and Methanosaeta*. Active inocula (1% to 5%, vol/vol) of these cultures were transferred along with *D. mccartyi* 195 culture (5% to 10% vol/vol) into the defined mineral salt medium with 5 mM organic acids (i.e. lactate or acetate), ~40 µmol of TCE, and a N_2-CO_2 or H_2-CO_2 headspace regarding to different co-growing strains. The resulting cultures were incubated at 34°C and periodically sampled to monitor the chloroethenes and ethene, as well as the concentrations of lactate and acetate.

Chlorinated ethenes and ethene, organic acids as well as cell growth were determined by GC, HPLC and qPCR, respectively, as described in "Objective 1".

Microarrays targeting the genome of D. *mccartyi* strain195 (DE195) were designed and applied as previously described by West *et al*. (45) and Johnson *et al*. (23). The microarray platform (GPL6336) was deposited to the NCBI Gene Expression Omnibus database (23). cDNA was synthesized from 10 µg RNA, then fragmented, labeled, and hybridized to each array. The hybridized arrays were stained, washed and then scanned with an Affymetrix Scan 3000 scanner. All procedures were performed according to the protocol outlined in Section 3 of the Affymetrix GeneChip Expression Analysis Technical Manual (www.affymetrix.com). Three replicate arrays were analyzed for each condition. The microarray analyses were performed using the R statistical program (www.r-project.org) (R 2005) with packages available from Bioconductor version 1.9 (www.bioconductor.org) (18) as described by Johnson *et al*. (23, 24). The Benjamini and Hochberg procedure (6) was applied to control the false discovery rate (FDR) below 0.05. In addition, only genes with absolute hybridization signal intensities greater than 200 for at least one condition and with more than two-fold changes between two conditions could be considered to be significantly regulated and used for further analyses.

DE195 and DE195/DVH (co-culture containing DE195 and *D. vulgaris* Hildenborough) cell samples were analyzed via 2-dimensional nanoES 2D LC-MS/MS as described by (42). Briefly proteomes were digested in peptides with sequencing grade trypsin (Promega, Madison WI) then separated via 2-dimensional HPLC (SCX/RP) into a nanoelectrospray source connected to a Linear ion Trap Orbitrap with full scans at 30K in Orbitrap and five subsequent data dependent MS/MS scans in the linear ion trap. MS/MS spectra were searched with SEQUEST (16) against a database created with both DE195 and DVH predicted protein sequences, common contaminants as well as common lab protein standards. Peptide identifications were filtered and sorted into proteins with DTASelect (Tabb *et al*., 2002) as previously described (42). Contrast (41) was used to display all proteins across runs, differentially expressed proteins were identified based on the following criteria (11): at least under one condition, > 40% sequence coverage, >5 unique peptides, and ≥ 2-fold difference in spectral counts identified between DE195/DVH and DE195 isolate. Data identification as well as the actual MS/MS spectra from every peptide and accessory scores are available at http://compbio.ornl.gov/Dehalo195_CoCulture/.

Results and discussion. Among the defined consortia we constructed and evaluated so far, the co-culture of *D. mccartyi* strain 195 and lactate-fermenting hydrogen-generating bacteria *Desulfovibrio vulgaris* strain Hildenborough or *Desulfovibrio desulfuricans* showed the most robust synergetic activities, and both triggered an increased TCE dechlorination rate and growth rate of strain 195 (Figure 13 A&B). *D. mccartyi* strain 195 was found able to reduce 80 µmole

(per bottle) TCE to ethene within a week in the co-culture with either *D. vulgaris* or *D. desulfuricans* while the reduction of the same amount of TCE in the pure *D. mccartyi* 195 culture took about three weeks. In addition, the maximum cell density of the *D. mccartyi* 195 in both co-cultures increased 1.5-fold while the time strain 195 needed to reach the density shortened 10-day compared with it grew in pure culture (Figure 15).

In addition, we have expanded our construction and investigation of defined consortia to include other hydrogen-generating fermentation bacteria, such as crotonate/butyrate fermenting *Syntrophomonas wolfei*, propionate fermenting *Syntrophobacter fumaroxidans*, and lactate fermenting *Bacteroides* spp. Except for *Syntrophomonas*, other genus of *Clostridium*, *Eubacterium*, *Syntrophobacter* and *Bacteroides* had no significant improvement in the dechlorination and growth rates of strain 195 (Table 3).

The microbes that can supply essential factors for strain 195 have also been screened, especially for the species that potentially generate corrinoids (the precursors of cobalamin) in their metabolic processes. We found that the present of homoacetogenic *Acetobacterium woodii* could enhance both growth and dechlorination of strain 195, while methanogenic *Methanosarcina* spp. *and Methanosaeta harundinacea* showed no effects under methanogenesis using either acetate or methanol as substrates (Table 3). It has been revealed that *Methanosarcina* and *Methanosaeta* strains used in this study grew insufficiently under acetoclastic methanogenesis condition. Thus, we hypothesized that the insufficient growth caused these strains to produce limited corrinoids, so that strain 195 could not benefit by growing with these strains.

In addition to acetoclastic methanogens, we also screened several hydrogenotrophic methanogens because they were frequently detected to be active community members coexisting with *Dehalococcoides* spp. in different dechlorinating environments. We therefore added various hydrogenotrophic methanogens into the co-cultures of strain 195 and *D. vulgaris* Hildenborough to investigate the role of these methanogens in synergetic and symbiotic activities with strain 195. Hydrogenotrophic methanogens selected in this study included *Methanobacterium congolense, Methanobacterium formicicum and Methanococcus maripaludis.* TCE degradation profiles showed that a triculutre of *D. mccartyi* strain 195, *D. vulgaris* strain Hildenborough and *M. congolense* (DE195/DVH/MC) faster dechlorination than the co-culture of *D. mccartyi* strain 195 and *D. vulgaris* strain Hildenborough. However, addition of *M. formicicum* and *M. maripaludis* to the co-culture of *D. mccartyi* strain 195 and *D. vulgaris* strain Hildenborough resulted in no improvement in the dechlorination rates and sometime undermined the beneficial effects of *D. vulgaris* strain Hildenborough to *D. mccartyi* strain 195, perhaps by competing with strain 195 for available hydrogen.

Table 3. Hypothetically function-supporting microorganisms that were selected in this study to construct defined consortia with *D. mccartyi* strain 195.

Hypothetically function-supporting microorganisms		Cultivation conditions				Sustain. growth & dechlo.[c]
Major Function	Strain(s)	Headspce	ED	EA	Carbon Source	
Lactate fermentation to produce H_2 and acetate	*D. vulgaris* Hildenborough	N_2/CO_2	Lactate	TCE	Acetate[a]	+
	D. desulfuricans Essex6					+
	D. desulfuricans G20					+
	E. limosum/					–
	C. propionicum					–
	C. freundii					–
	C. cellulovorans					–
	B. galacturonicus					–
Corrinoids generating	*A. woodii*	H_2/CO_2	H_2	TCE	Acetate[b]	+
	M. barkeri/mazei/ M. harundinacea (Acetoclastic methanogenesis)					–
	M barkeri/ mazei (Methylo-trophic methanogenesis)					–
1. Lactate fermentation to produce H_2 and acetate 2. Corrinoids generating	*D. desulfuricans* Essex6 + *A. woodii*	N_2/CO_2	Lactate	TCE	Acetate[a]	+
	D. vulgaris Hildenborough + *A. woodii*					+
	D. vulgaris Hildenborough + *M. barkeri*					–
	D. vulgaris Hildenborough + *M. Mazei*					–
	D. vulgaris Hildenborough + *M. harundinacea*					–
1. Lactate fermentation to produce H_2 and acetate 2.Hydrogenotrophic methanogenesis	*D. vulgaris* Hildenborough +*Methanobacterium congolense*	N_2/CO_2	Lactate	TCE	Acetate[a]	+
	D. vulgaris Hildenborough + *M. formicicum*					–
	D. vulgaris Hildenborough + *Methanococcus maripaludis* S2					–
Other organic acid fermentation	*S. wolfei*	N_2/CO_2	crotonate/ butyrate	TCE	Acetate[a]	+

S. halorespirans	N_2/CO_2	propionate	TCE	Acetate[a]	–	

ED: electron donor; EA: electron acceptor.
[a] Acetate was produced by supporting bacteria. [b]Acetate was added exogenously.
[c] Sustained growth and dechlorination.

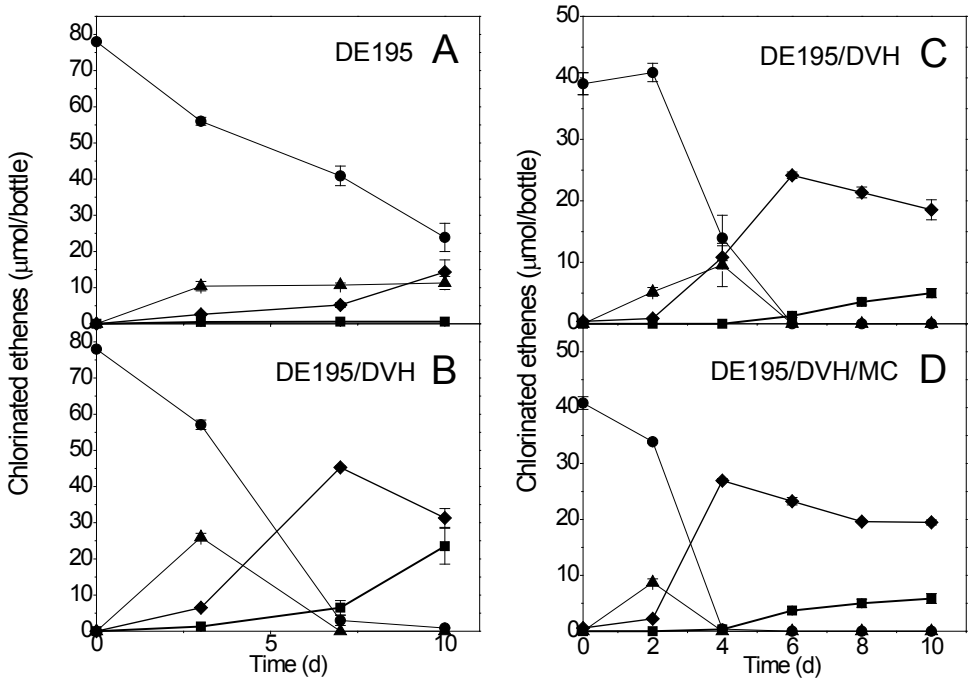


Figure 13. Temporal changes in the quantities of solvents for (A) DE195 fed ~78 µmol TCE, (B) DE195/DVH fed ~78 µmol TCE, (C) DE195/DVH fed ~40 µmol TCE, and (D) DE195/DVH/MC fed ~40 µmol TCE. All measurements are averages from three biological replicates and error bars are the standard deviation; (●) TCE, (▲) c-DCE, (♦) VC, (■) ethene.

A comparison between strain 195 isolate, DE195/DVH co-culture and DE195/DVH/MC tri-culture was made on the dechlorination rate and cell growth. Results indicated that TCE degradation was substantially faster in the sytrophic cultures versus the isolate. That is, while it took 20 days for DE195 to dechlorinate 78 µmol TCE to VC and ethene when grown alone (ca. 3.8 ± 0.1 µmol per day), it took only 7 days (ca. 11.0 ± 0.01 µmol per day) in the co-culture, the DE195/DVH/MC tri-culture took 4 days to dechlorinate 40 µmol TCE (ca. 10.1 ± 0.3 µmol per day), compared to 6 days (ca. 7.9 ± 0.5 µmol per day) in the co-culture control (Figure 13). The amount of ethene produced from VC in the co-culture was higher compared with the isolate, and similar compared with the tri-culture (Figure 13). By the 20th subculture of the co-culture, the syntrophic growth of DE195 was substantially more consistent and robust than growth in isolation. For example, while subculturing 10 vials of DE195 isolate resulted in only 6 successful cultures, all of the 10 vials inoculated with DE195/DVH subcultures grew successfully. Similar results were observed over multiple subsequent subcultures. The density of DE195 grown in the 20th subculture of the co-culture was approximately 2 times higher than when DE195 was grown alone (Figure 14A) (two-tailed student's t-test $p=0.002$). The density of DE195 grown in the 16th subculture of the tri-culture was lower than the co-culture control

(Figure 14A) (two-tailed student's t-test p=0.04). The growth of DVH in the tri-culture was 5 fold greater than that in the co-culture control. The ratio of DE195 to DVH cells in co-culture remained approximately 5:1 (Figure 14B) over multiple subsequent subcultures, while the ratio of DE195 to DVH cells in the tri-culture remained around 1:1.5 (Figure 14B). Further, the ratio of DE195 to MC and DVH to MC remained at approximately 4:1 and 6:1, respectively (Figure 14B). The cell yield of DE195 in the co-culture was higher ($9.0 \pm 0.5 \times 10^7$ cells per μmol Cl^- released) than that of DE195 in pure culture ($6.8 \pm 0.9 \times 10^7$ cells per μmol Cl^- released) (two tailed student's t-test p=0.02), while the cell yield of DE195 in the tri-culture ($7.3 \pm 1.8 \times 10^7$ cells per μmol Cl^- released) had no significant difference from that of the co-culture control ($10 \pm 1.6 \times 10^7$ cells per μmol Cl^- released) (two tailed student's t-test p=0.12).

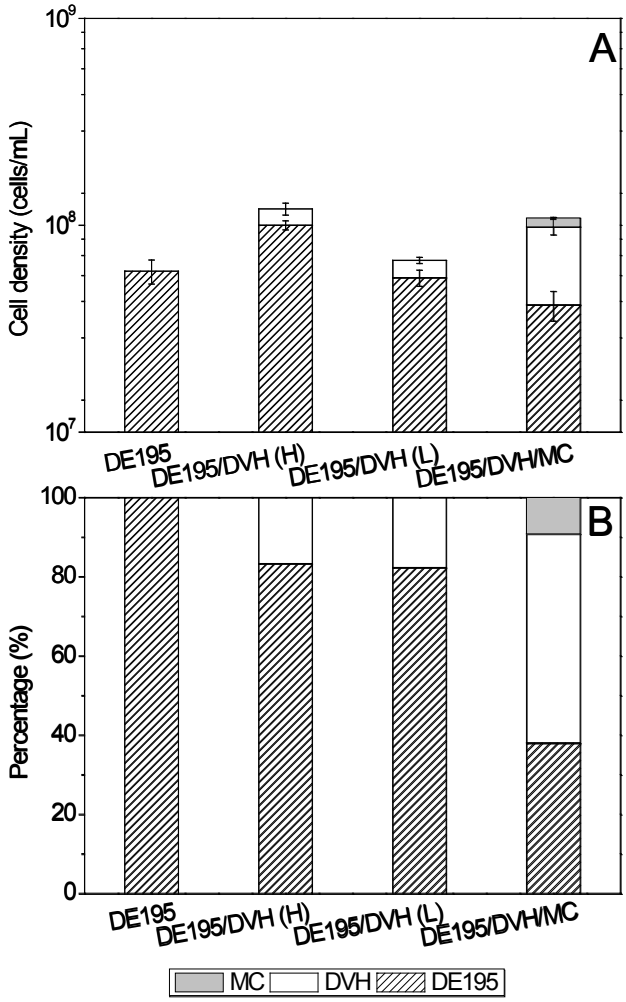


Figure 14. Cell density (A) and percent of each species in syntrophic cultures (B) for DE195 fed ~78 μmol TCE; DE195/DVH(H) fed ~78 μmol TCE; DE195/DVH(L) fed ~40 μmol TCE).

When DE195 was grown in isolation, 500 μmoles acetate was added to the basal medium, far exceeding the stoichiometric requirement of 0.03 μmoles for growing DE195 to 10^8 cells mL^{-1} and assuming a cell composition of $C_5H_7O_2N$. In the co- and tri-cultures, lactate was provided as the electron donor and carbon source, with the expectation that DVH would ferment it into acetate and hydrogen with the stoichiometry given in, thus supporting the hydrogen and acetate

requirements of DE195 and MC. However, the growth of DVH on lactate is only thermodynamically favorable when sufficient hydrogen is consumed by another strain (DE195 for dechlorination and/or MC for hydrogenotrophic methanogenesis), such that sustained survival of the constructed consortia requires syntrophic association among the species. The 500 µmoles lactate provided per bottle was partially consumed in the co-culture but was completely consumed in the tri-culture (Figure 15), with near stoichiometric acetate production in both. The co-culture also generated 140 µmoles (Figure 15A) hydrogen indicating that approximately 40% of the lactate consumed provided electrons for dechlorination. In the tri-culture, hydrogen was completely consumed by day 6 and 250 µmoles methane were generated (Figure 15B), indicating that approximately 90% of the lactate electrons were consumed by methanogenesis while only 10% were used for dechlorination. Although MC consumed most of the generated hydrogen in the tri-culture, the aqueous hydrogen concentration never dropped below 5 nM, remaining above the hydrogen threshold (2 nM) for dechlorination by *Dehalococcoides*. Consequently, competition for hydrogen between MC and DE195 was not observed to affect dechlorination in this study.

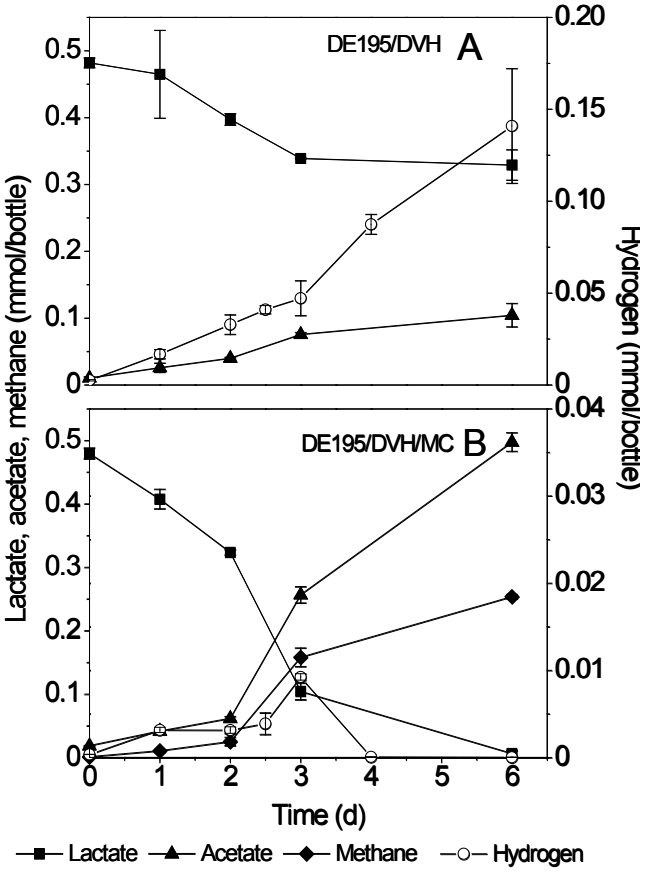


Figure 15. Consumption of lactate and production of acetate, hydrogen, and methane in (A) DE195/DVH and (B) DE195/DVH/MC (both fed by ~40 µmol TCE). Note: different H_2 scales in (A) and (B).

Scanning electron micrographs in Figure 16 exhibit the closely grown cells of *D. mccartyi* strain 195, *D. vulgaris* Hildenborough, and *M. congolense* in the co-culture and tri-culture.

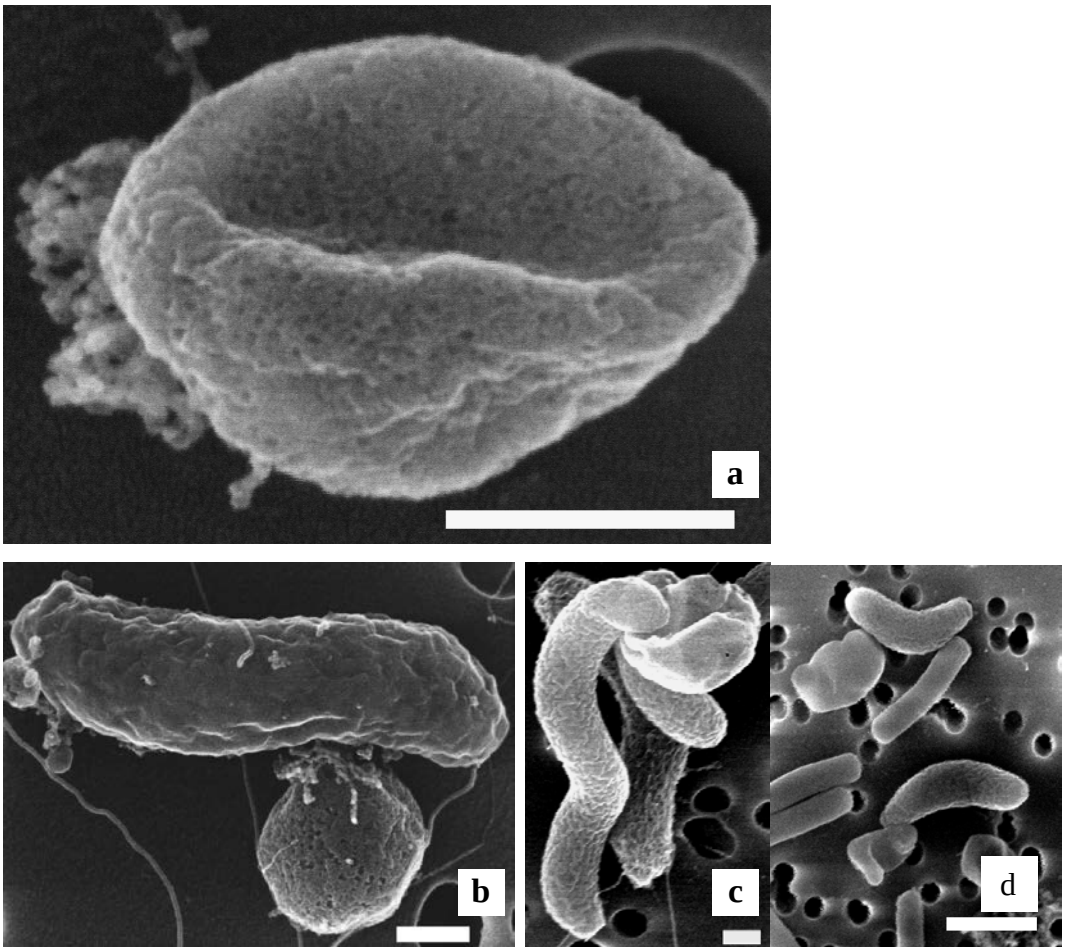


Figure 16. Scanning electron micrograph of *D. mccartyi* strain 195 in pure culture (a), DE195/DVH (b) and (c), DE195/DVH/MC (d). Bar, 0.2 μm.

Differential gene expression in the co-culture DE195/DVH, as well as in tri-culture DE195/DVH/MC, comparing to the pure culture DE195 was investigated using a microarray target at the strain 195 genome. Results exhibited 102 significantly up- or down-regulated genes in the co-culture, compared to DE195 grown in isolation (Figure 17), while no significant transcriptomic difference was observed between co- and tri-cultures. In addition, proteomic analysis showed that 120 proteins were differentially expressed in the co-culture compared to DE195 grown in isolation. Physiological, transcriptomic and proteomic results indicate that the robust growth of DE195 in co- and tri-cultures is due to the advantages associated with the capabilities of DVH to ferment lactate to provide H_2 and acetate for growth, along with potential benefits from proton translocation, cobalamin-salvaging, and amino acid biosynthesis, while MC in the tri-culture provided no significant additional benefits beyond those of DVH.

The cobalamin co-factor salvage and transport genes (DET0650-0652/DET0684-0686), as well as the genes annotated to construct and attach the lower ligand base to cobyric acid (DET0657-0660/DET0691-0694) and its associated riboswitch were significantly down-regulated in the co-culture compared with in isolation (Figure 18C), suggesting that cobalamin might be produced and released in sufficient quantities by DVH to down-regulate the cobalamin-

salvaging genes of DE195. Interestingly, DET0125 and DET0126, two putative cobalamin riboswitches were also down-regulated in the co-culture.

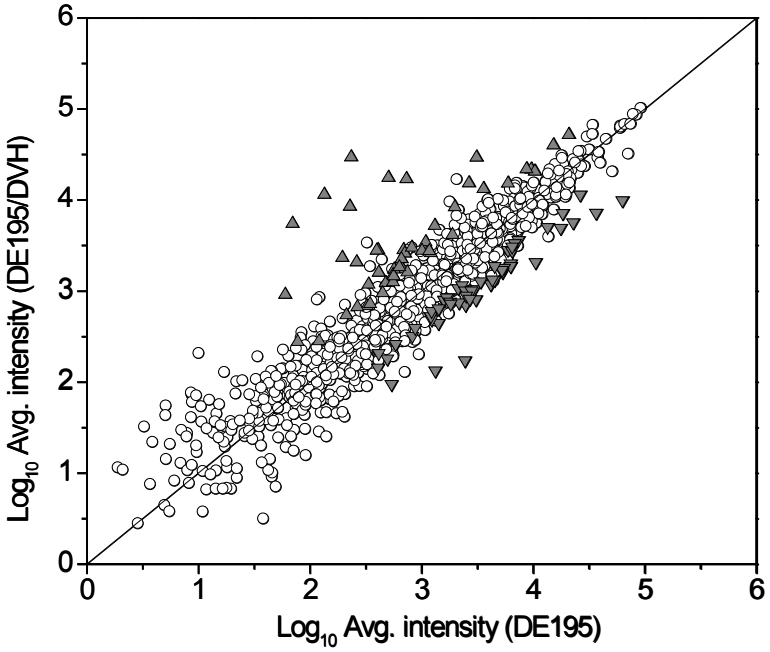


Figure 17. Plot of signal intensities of transcripts from DE195 grown alone vs. signal intensities of transcripts from DE195/DVH (colored data points represent statistically significant differential transcription, avg. intensity > 200, $p < 0.05$, > 2-fold difference; genes significantly up-regulated (▲) or down-regulated (▼) in DE195/DVH compared to DE195. All measurements are averages from three biological replicates.

Among the 15 genes that are predicted to encode SAM proteins within the genome of DE195, 8 genes were actively transcribed and 4 genes (DET0622, DET1280, DET1368, and DET1629) were down-regulated in the co-culture (Figure 18C) while one gene (DET1314) was significantly up-regulated in the co-culture as compared to the DE195 isolate. Predicted oxidoreductase complexes identified in the genome of DE195 include: molybdopterin oxidoreductase (Mod), putative formate dehydrogenase (Fdh), NADH-ubiquinone oxidoreductase (complex I, Nuo), as well as six hydrogenase complexes (i.e. Hup, Hym, Hyc, Vhu, Ech and Hyp). In this study, the entire operon for the Nuo (DET0923-933) was significantly down-regulated in the co-culture compared with the isolate (Figure 18A). Within the six hydrogenase complexes, the Hym operon (DET0145-148) was observed to be significantly down-regulated, whereas a gene predicted to encode a HymA subunit but not associated with other hydrogenase genes (DET0446) was significantly up-regulated in the co-culture. Other genes within the hydrogenase complexes did not show significant regulation (Figure 18B), or were not actively transcribed in either culture condition (i.e. Hyc operon).

The genome sequence of DE195 revealed 19 potential RDase genes, 5 of which (DET0079, DET0162, DET0180, DET0318, and DET1559), were actively transcribed in the isolate with an additional 4 up-regulated in the co-culture (DET0311, DET1171, DET1522, and DET1545). Although the transcript levels of *tceA* (DET0079) were at the same level between the isolate and co-culture, all the other actively transcribed RDase genes were up-regulated in the co-culture (Figure 18A).

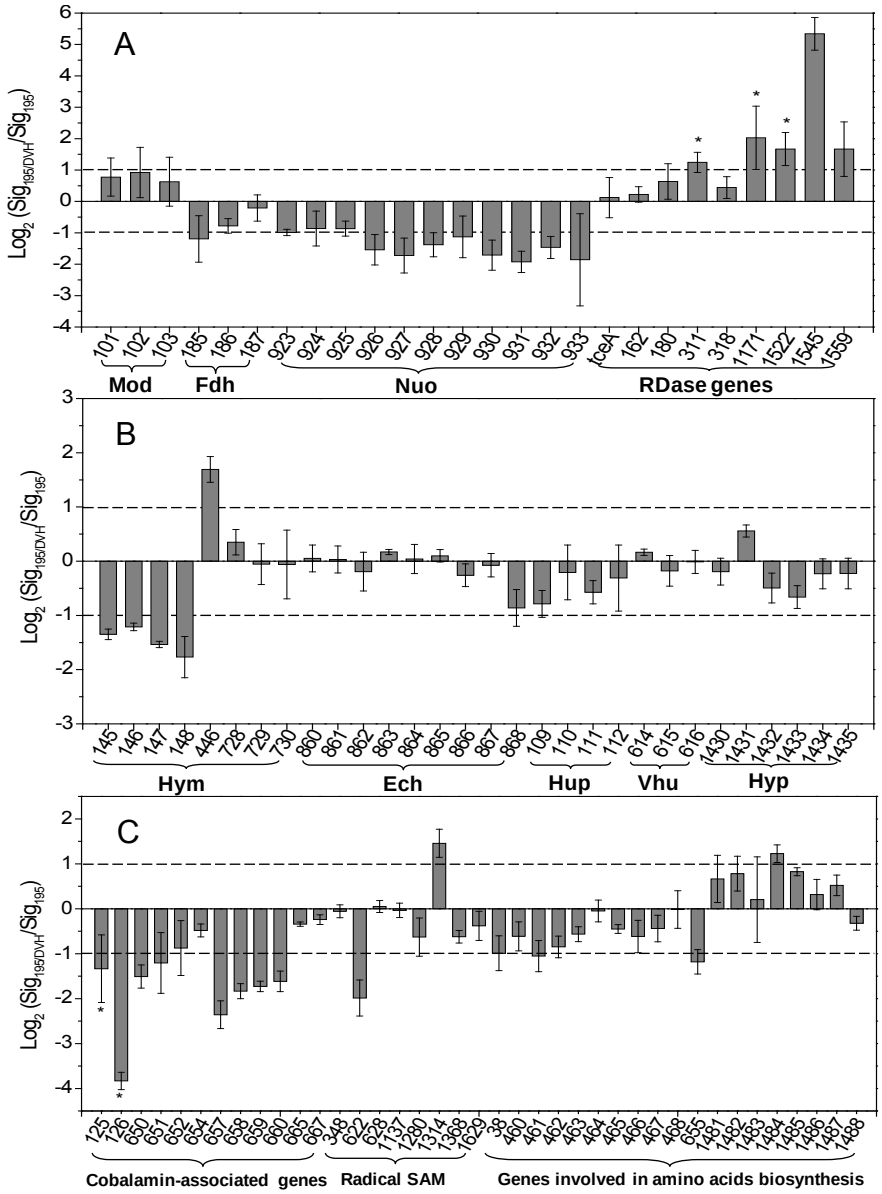


Figure 18. Log_2 signal intensity ratio between genes of strain 195 in the co-culture and strain 195 isolate. All measurements are averages from three biological replicates and error bars are the standard deviation. All X axis labels are designated for strain 195 gene loci (e.g. DET0101). Dashed lines indicate the 2-fold difference in signal intensities. * indicates genes not actively transcribed (signal intensity < 200) in one of the two cultures.

Several genes involved in amino acids biosynthesis were significantly differentially transcribed. Genes predicted to encode glutamate synthase and chorismate mutase/prephenate dehydratase (DET0038 and DET0461) were significantly down-regulated in the co-culture (Figure 18C). DET0461 is a gene involved in the biosynthesis of chorismate, a precursor of all aromatic amino acids (i.e. tryptophan, tyrosine and phenylalanine). The other genes within the chorismate operon (DET0462-DET0468), except DET0464 and DET0468 were also down-regulated, but less than two-fold difference (Figure 18C). The only amino-acid synthesis gene

significantly up-regulated in the co-culture was DET1484, predicted to encode indole-3-glycerol phosphate synthase which is associated with tryptophan synthesis.

One interesting observation that may be related to the increase in dechlorination activity and robustness of DE195 grown in the co- and tri-culture is the increased concentration of the corrinoid co-factor in the cultures containing DVH. Although the vitamin B_{12} concentration detected in the co-culture was only 10 μg L^{-1} higher than that of the isolate (60 μg L^{-1} for the co-culture versus 50 μg L^{-1} for the isolate), the amount of vitamin B_{12} required by *Dehalococcoides* might be up-taken and utilized immediately, such that what was detected in the supernatant is the B_{12} concentration at steady-state. It is not specifically known what form of corrinoid DE195 prefers. DVH possesses the full set of genes required for the biosynthesis of adenosylcobalamin, a derivative of vitamin B_{12} (37), however mechanism for the transport of adenosylcobalamin has not been identified. Because of the important role of cobalamin as a co-factor for the RDases, the uptake and potential transformation of this corrinoid needs to be elucidated. Both transcriptomic and proteomic analyses showed that cobalamin-associated genes, including genes predicted to encode a corrinoid ABC-type transport system (DET0650-0655/DET0684-0689) and a corrinoid salvage system (DET0657-0660/DET0691-0694), were significantly down-regulated in the co-culture, which suggests that DE195 may exert less energy salvaging corrinoids when DVH is present. DET0650-0654 were also found to be significantly down-regulated when DE195 is grown with excess vitamin B_{12}, as well as with spent medium of a *Dehalococcoides*-containing enrichment culture (ANAS) compared with DE195 grown with limited vitamin B_{12} (24). Interestingly, another significantly down-regulated gene in both this study and the study by Johnson *et al*. (2009) was DET0126, which is annotated as an anthranilate phosphoribosyltransferase (TrpD) and reported to have an upstream putative cobalamin-binding riboswitch (24). Although the function of DET0126 remains unclear, its differential regulation within different growth conditions suggests that it is of biological importance to DE195.

Since the radical SAM proteins have functions in the biosynthesis of vitamins and co-factors, the down-regulation of several SAM protein-encoding genes could also indicate an increase in the availability of the cobalamin co-factor for DE195 grown with DVH.
Another possible reason for more robust growth and faster dechlorination of DE195 grown in the co- and tri-cultures is facilitated hydrogen transfer between DVH and DE195 due to physical proximity under the unshaken conditions. Membrane-bound oxidoreductase operons, such as Fdh, Nuo, as well as Hym and Hup operon were down-regulated in DE195 grown in the co-culture compared with the DE195 isolate. These operons are predicted to be components of the electron transport chain, and are important in hydrogen transfer for DE195 (39). Moreover, the down-regulation of TatC gene, whose product functions in the secretion of folded proteins such as the putative hydrogenase complexes and Fdh, could be a corresponding effect of a down-regulation of those proteins. The down-regulation of these genes suggests that the syntrophic growth of DE195 and DVH might have a different hydrogen transfer system than the DE195 isolate. Transcriptomic analysis of DVH syntrophically grown with a hydrogenotrophic methanogen compared with DVH grown in sulfate-limited monocultures showed that genes encoding hydrogenases Coo, Hyd and Hyn were among the most highly expressed and up-regulated genes (44). Therefore, differential regulation of genes encoding hydrogenases of DVH grown in DE195/DVH compared with the DVH isolate should be further investigated, in order to elucidate the interspecies hydrogen transfer during syntrophic growth.

Finally, aromatic amino acids (i.e. tryptophan, tyrosine, and phenylalanine) might play important roles in the robust syntrophic growth of DE195 with DVH. In the co-culture, the down-regulation of genes involved in the biosynthesis of chorismate, a precursor for aromatic amino acid biosynthesis, suggests that DVH might support the growth of DE195 by decreasing the need for *de novo* aromatic amino acid biosynthesis, consistent with our recent detailed amino acid study.

In summary, *Dehalococcoides mccartyi* strain 195 exhibits faster dechlorination and more robust growth when growing syntrophically with DVH using lactate as carbon and energy source than when grown in isolation on the gaseous substrate hydrogen. The difference in gene transcription and protein expression levels between DE195 grown in isolation and with DVH suggests that effective transfer of hydrogen, cobalamin and some amino acids may contribute to the enhanced dechlorination and robust growth of DE195 in the co-culture. These results provide an improved method for rapid and robust growth of *Dehalococcoides* strains using lactate as an inexpensive, widely available substrate.

Objective 3: Identification of mRNA-based functional biomarkers by application of genus-wide microarray on different TCE-dechlorinating microbial communities.

Although the 16S-rRNA-based phylogenetic biomarkers are powerful tools for quantifying the presence and activity of supporting organisms, they do not definitively reveal the functional role(s) of the supporting organisms. This limitation may lead to several types of erroneous predictions. One type of erroneous prediction may result if functional redundancies are not recognized among the identified supporting organisms. For example, two phylogenetically distinct organisms may supply *Dehalococcoides* with the same essential vitamin cobalamin. If only a 16S-rRNA biomarker approach is taken, we may conclude that both cobalamin-producing organisms are needed for optimal growth and activity when only one of the two organisms is actually needed. Another type of erroneous prediction may result if unidentified organisms can effectively fill the functional role of identified organisms. For example, we may identify a group of organisms that support *Dehalococcoides* by fermenting organic acids to hydrogen. When probing an uncharacterized community for these stimulatory organisms, we may falsely assume that at least one of our identified organisms must be present for optimal activity when in fact other unknown organisms may already be filling the role. To overcome both of these limitations, we complement our set of 16S-rRNA-based phylogenetic biomarkers with mRNA-based functional biomarkers that are identified using a genus-wide microarray.

3.1 Design, validation and application of a genus-wide microarray (This study is published in Lee et al., ISME J, 2011)

We have effectively identified biomarkers indicative of cobalamin (vitamin B_{12}) and nitrogen stress from both pure culture and defined *Dehalococcoides*-containing consortia by designing and applying a whole-genome expression array for *D. mccartyi* strain 195. In addition, we've further designed and validated a new genus-wide microarray containing 4 genomes of *Dehalococcoides* species (strain CBDB1, BAV1, 195 and VS), and applied this genus-wide

microarray to un-sequenced *Dehalococcoides* strains, as well as to different *Dehalococcoides*-containing microbial communities, in order to investigate genome presence and expression patterns.

Materials and methods. The microarrays targeting all genes from the four sequenced *Dehalococcoides* genomes (195, CBDB1, BAV1 and VS) were produced by Affymetrix (Santa Clara, CA, USA) as prokaryotic midi format (format 100) photolithographic microarray chips. The design strategy was to target each gene within the genomes with a unique probe set that would not cross hybridize with closely related genes. If a unique probe set could not be found for a gene, multiple genes were targeted by the same probe set. Probe sets consisted of 11 probe pairs (22 total probes) of 25-mer oligonucleotide probes that are distributed along the length of the respective gene. Each probe pair is made up of a perfect-matched probe and a corresponding single mismatch probe in which the thirteenth nucleotide is a mismatch with the target to control for nonspecific hybridization. The microarrays also contained 24 positive control and 21 negative control probe sets as described previously (45) to facilitate calibration and to resolve background signals. The input sequences for the microarrays were downloaded from the National Center for Biotechnology Information and included *Dehalococcoides* sequences as well as all the GenBank sequences targeting putative RDases, known hydrogen-producing nitrogenases and hydrogenases, and a variety of enzymes involved in vitamin B12, biotin and methionine biosynthesis pathways.

gDNA of two isolated strains from ANAS (ANAS1 and ANAS2), the ANAS enrichment culture, as well as those four groundwater enrichments (see "Objective 1") was applied on the genus-wide microarray according to Affymetrix's protocol. 1 µg of gDNA was prepared for each microarray and triplicate analyses were performed for each culture. The gDNA along with a positive control spike-mix was fragmented, biotin end-labeled and hybridized to the microarray. Following hybridization, the microarray was processed according to the instructions given in section 3 of the Affymetrix GeneChip Expression Analysis technical manual. Data analysis was performed using Affymetrix GeneChip software and the MAS5 algorithm. Each microarray was normalized by scaling the signal intensities of the positive control spike-mix to a target signal intensity of 2500 to allow comparison between microarrays. A gene was considered "present" in a culture if the probe set across all three replicate samples had signal intensities greater than 140 and a P value less than 0.05.

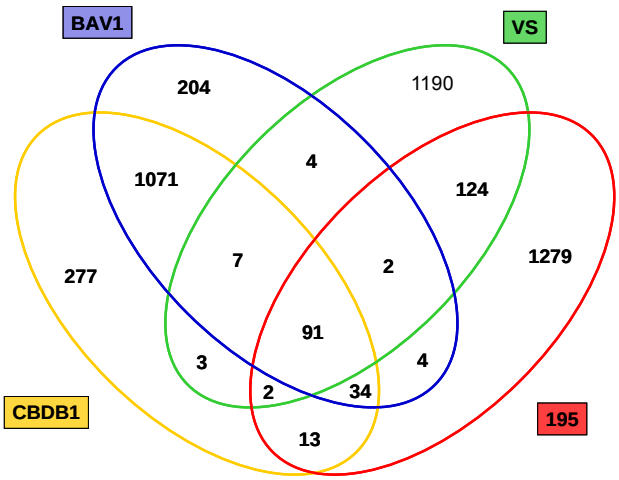


Figure 19. A Venn diagram showing the genome target distribution for the 4305 probe sets on the microarray.

Results and discussion. In the design of the microarrays, unique probe sets targeting genes from only a single genome were sought to allow strain differentiation. In total, 4305 probe sets were designed to represent 6010 Dehalococcoides genes, including 68.5% unique probes targeting genes from a single genome (Figure 1). Genes in strains BAV1 and CBDB1 are targeted by a large number of common probe sets (1203), whereas strains 195 and VS are represented by a relatively large number of unique probe sets (1279 and 1190, respectively) (Figure 19). The limited number of unique probe sets for strains BAV1 and CBDB1 mostly target genes within the HPRs or integrated elements (IEs).There are also 348 probe sets representing genes outside of the four *Dehalococcoides* genomes with functions of cobalamin biosynthesis, biotin synthesis, methionine synthesis (cobalamin dependent), hydrogen production, as well as nitrogen fixation.

We've validated this genus-wide microarray with gDNA of strain 195 and BAV1 as positive controls, and gDNA of non-*Dehalococcoides* species, *Anaeromyxobacter dehalogenans* strain 2CP-C, as negative control. Results are shown in Figure 20, which indicates the qualitative response of the microarray. As expected, probe sets responded correspondingly when positive and negative controls were appropriately hybridized. Titrated gDNA of strain 195 and BAV1 was applied on the array. $R^2 > 0.96$ was observed, which indicated a good quantitative representation of gene presence and expression on this microarray.

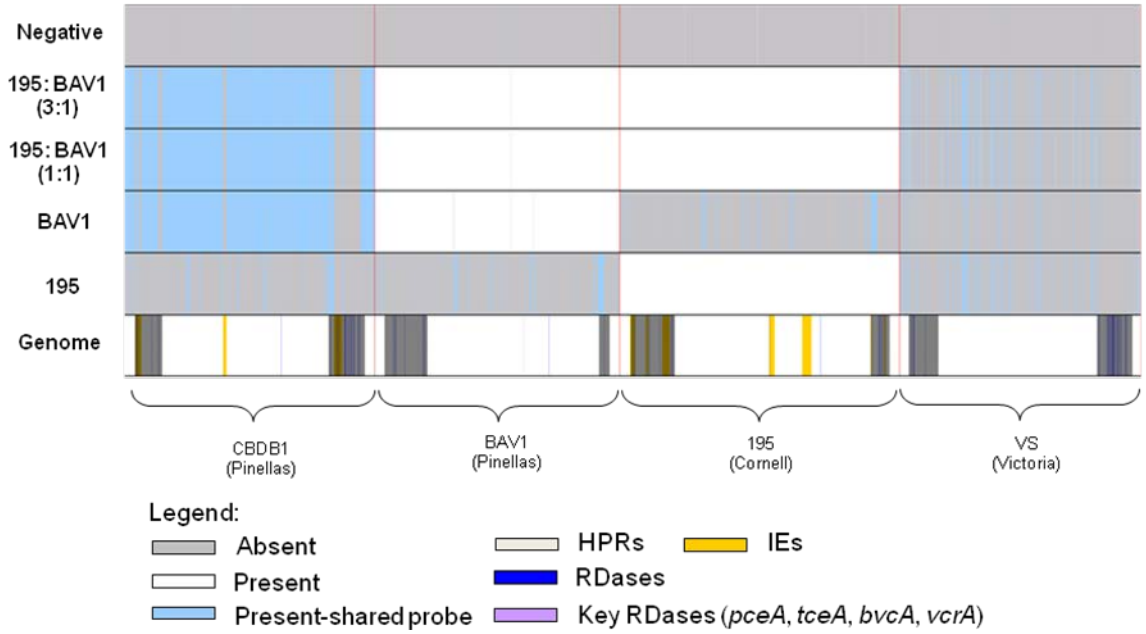


Figure 20. Linear representation of the four *Dehalococcoides* genomes that are tiled on the array with gDNA of both negative control and 3 positive controls.

Figure 21 shows the genes of the four *Dehalococcoides* genomes tiled on the array, which are present or absent in ANAS culture, ANAS1 and ANAS 2 isolates, with strain 195 as a positive control. ANAS1 and ANAS2 are the two *Dehalococcoides* strains isolated from ANAS culture. ANAS1 contains *tceA* , and is capable of dechlorinating TCE to VC; while ANAS2 contains *vcrA*, and is capable of dechlorinating VC to ethene. Thus, together with these two strains, ANAS culture is able to degrade TCE beyond VC to ethene. Although ANAS1 and ANAS2 have different dechlorination characteristics, both of them resembled a strain 195-like genome, with the absence of genes in high plasticity regions (HPR) and integrated element (IE) regions. ANAS1 only has *tceA*, while ANAS2 has *vcrA*, which is in the genome of strain VS. Interestingly, the tryptophan operon of strain 195 was missing in both ANAS1 and ANAS2, while that of strain VS was present, instead.

gDNA of the four groundwater enrichment cultures were also hybridized onto the genus-wide microarray to compare their genomic content against the sequences of the four sequenced strains tiled on the array. Results are shown in Figure 21. Similarly, all four cultures resembled a strain 195-like genome with only *tceA* found to be present, and most genes in HPR and IE regions were missing. Moreover, the tryptophan operon of strain 195 was also missing in all four cultures, while that of strain VS was present. The percentage of present and absent genes from strain 195 in different enrichment cultures and isolates are shown in Figure 22. Different percentages of genes present or absent in ANAS1 and ANAS2 were observed. More genes were absent within/outside HPR/IE regions in ANAS2 than ANAS1. For the four groundwater enrichment cultures, the presence/absence of genes within strain 195 genome was similar. Interestingly, genes that encode tRNA for Ala and Val were missing in the non-methanogenic cultures (NoMethB12 and NoMeth) compared with the methanogenic ones (MethB12 and Meth).

This suggests that these two amino acids (Alanine and Valine) might play an important role in methanogenic environments.

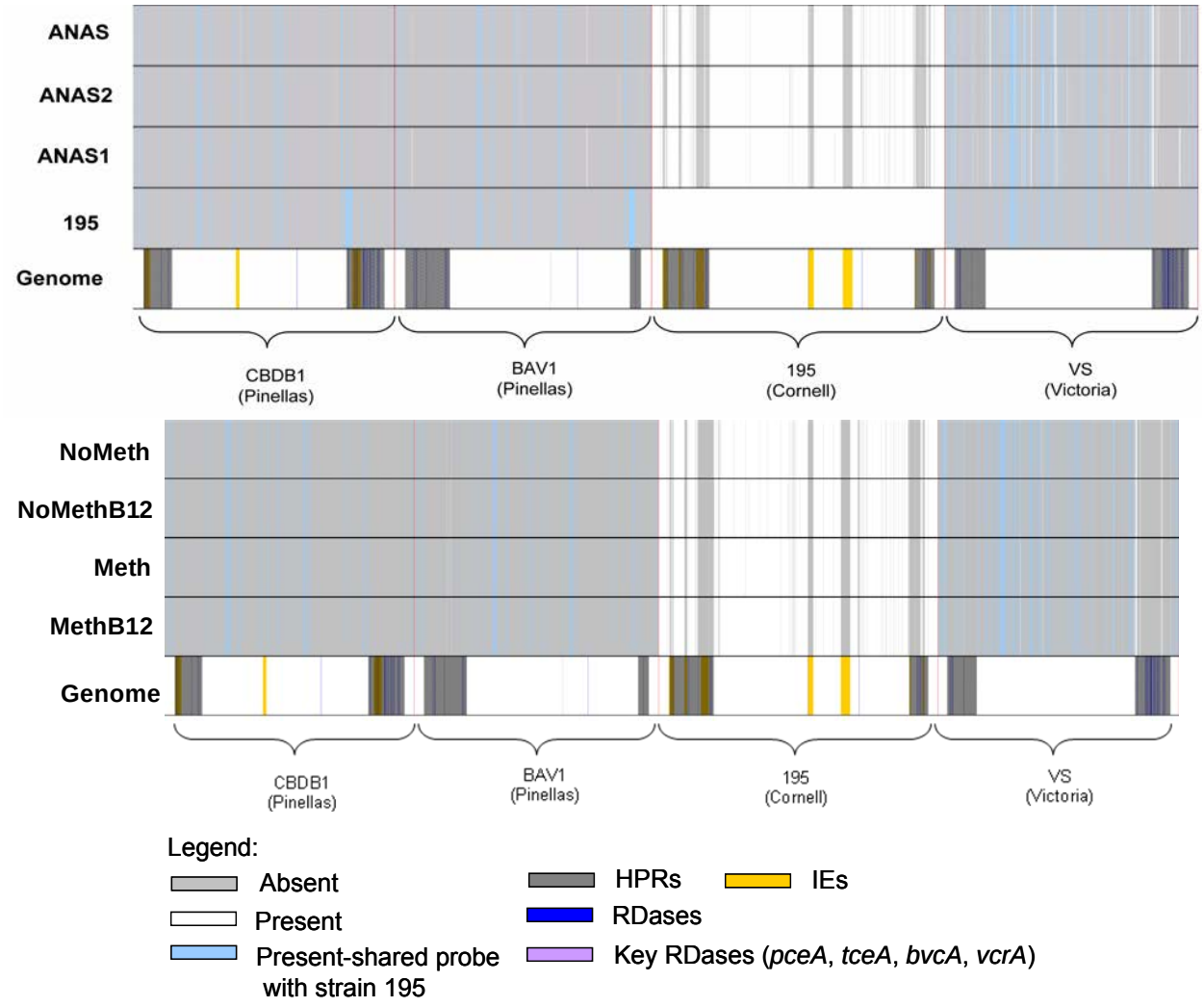


Figure 21. Linear representation of the four *Dehalococcoides* genomes that are tiled on the array with gDNA of different isolated *Dehalococcoides* strains and enrichment cultures. Each row represents a sample as indicated on the y-axis. The last row is the genome annotation with the orange color indicating the strain 195's integrated elements, dark gray indicating the high plasticity region (HPR) and dark blue indicating RDases. The gray and white colors in the corresponding samples represent genes that are absent and present, respectively, light blue indicates genes outside of the strain 195 genome that are present due to sharing a probe set with genes in strain 195. The purple color indicates key RDase genes.

There are three important findings from this study that highlight the biology of *Dehalococcoides*. First, even though a novel isolate can have identical dechlorination capabilities as a sequenced strain, their genomes might not necessary be the same and instead, the genome of the new isolate can resemble that of a different sequenced strain. We believe the reason for this observation has to do with our second key finding. We observed that the key functional reductive-dehalogenase (RDase) genes with assigned dechlorination function (i.e., *pceA*, *tceA*, *bvcA*, *vcrA*) may not necessary be restricted to the genome where the RDases were

first identified. For example, the *tceA* gene was first identified in strain 195 but we found that this gene can in fact be present in an un-sequenced strain that has a genome resembling that of the sequenced strain BAV1. Furthermore, if a *Dehalococcoides* strain lacks one of the key RDases, the typical dechlorination functions can be lost. We believe horizontal gene transfer is a major driver for the observed results and suggest that the overall genomic content of a *Dehalococcoides* strain may play a lesser role in dictating dechlorinating functions in comparison to the key functional RDases. The third important result of this study is the fact that many genes within the integrated regions in the sequenced strains are missing in the novel strains, suggesting that different populations of *Dehalococcoides* likely contain unique integrated elements.

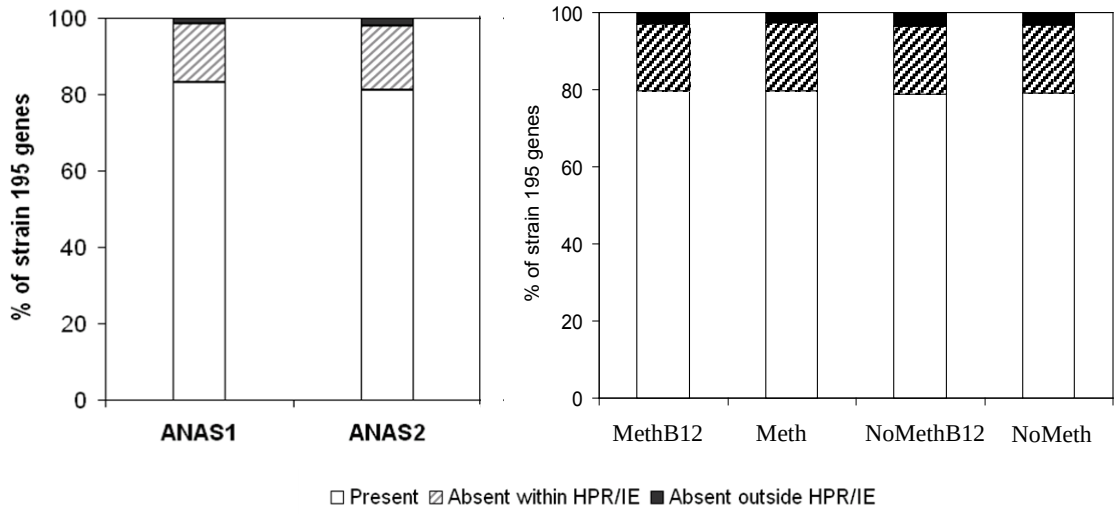


Figure 22. Percentage of strain 195 genes detected in different enrichment cultures.

3.2 The development of fluorescence-activated cell sorting (FACS), whole genome amplification (WGA) and microarray method (This study is in preparation for submission)

The application of microarray approaches to target a specific subpopulation within a microbial community is inherently challenging as the target genomes can be masked by other sequences, especially if the relative concentration of the target organisms is low. Therefore, one of our extended objectives to Objective 3 is to enhance the resolution of genomic analysis, fluorescence-activated cell sorting (FACS) techniques followed by whole genome amplification (WGA) have been established to separate and recover *Dehalococcoides* populations from a microbial community. The approach of coupling cell sorting with WGA and microarray was first tested with a pure culture and later two enrichment cultures containing ~10% and ~1% *Dehalococcoides* as proof of concept.

Materials and methods. *D. mccartyi* strain 195, enrichment cultures MethB12 and AD14 containing ~10% and ~1% *Dehalococcoides* cells, respectively, were all grown in defined medium as described previously. Enrichment AD14 was developed using sludge from an anaerobic digester of a pesticide factory in Gehua (Hubei Province, China). The culture is maintained with lactate and dechlorinates TCE to ethene.

Actively dechlorinating cells from pure culture (~10^9) or enrichment cultures (~10^{10}) were harvested by centrifugation (21,000× g, 3 min, 4°C), the supernatant removed, and re-suspended in 200 µL of 1× PBS. Cells were fixed with 600 µL of 4% paraformaldehyde (in 1× PBS) at 4°C for 13 hours. The fixed cells were washed three times with 800 µL of 1×PBS and re-suspended in 180 µL of hybridization buffer prepared according to a described protocol with 30% formamide (46). Duplicate cultures were set up for each experiment and three samples were prepared from each culture for the hybridization of different probes. Two were for negative controls where no probe or the non-specific probe Non338 (3) was added and the third was incubated with a published *Dehalococcoides*-specific probe (5'-AGCTCCAGTTCRCACTGTTG- 3') targeting the 16S rRNA (46). Probes labeled with Alexa Fluor 488 were synthesized by Invitrogen (Carlsbad, CA) and added to a final concentration of 5 ng/µL in the hybridization reaction. Samples were incubated at 37°C in the dark for 2 h and then stained with 20 ng/µL (46) of the DNA-intercalating dye 4′,6-diamidino-2-phenylindole (DAPI; Invitrogen, Carlsbad, CA) for 15 min. After hybridization, cells were pelleted (21,000× g, 10 min, 25°C) and incubated with 500 uL of pre-warmed hybridization buffer at 37°C for 20 min. Subsequently, cells were pelleted and washed with 500 uL of pre-warmed wash buffer (46) at 48°C for 20 min. Finally, cells were re-suspended and homogenized in 50 uL of 1× PBS for microscopic observations and cell sorting. Cells were observed using a Zeiss AxioImager M1 epifluorescence microscope (Thornwood, NY) and a 100× oil immersion objective with the appropriate filter set at the UC Berkeley Biological Imaging Facility. Images were captured separately from the same examining field with an Orca-03 CCD digital camera (Hamamatsu, Japan) and an exposure time of 0.25 s for both DAPI and Alexa Fluor 488. Images were viewed using the iVision software (BioVision Technologies, PA, USA).

Cell sorting was performed on a MoFlo high speed flow cytometer (Dako-Cytomation, Carpinteria, CA) at the UC Berkeley Flow Cytometry Facility. The instrument was stringently sterilized prior to each run. Forward and side light scattering as well as the Alexa Fluor 488 fluorescence of microbial cells were measured. The sorting gate was determined using the two negative controls with no probe and the non-specific probe in conjunction with the sample labeled with the *Dehalococcoides*-specific probe. Cells displaying a fluorescence signal above the cutoff were gated and collected into 1.5 mL centrifuge tubes.

Triplicate samples of one µg of amplified DNA or gDNA from cultures were prepared for microarray hybridization according to previously described protocols. The resulting data were normalized and analyzed according to previously described procedures. Briefly, for a gene to be considered "present", the corresponding probe set across all three replicate samples had to have signal intensities greater than 140 and a *P* value less than 0.05. The minimum signal requirement was relaxed for the amplified DNA from enrichment MethB12.

Results and discussion. The applicability of FACS was first tested with the isolate *D. mccartyi* 195 as positive control. Applying a *Dehalococcoides*-specific probe targeting the 16S rRNA (46), we verified that fluorescence signals from *Dehalococcoides* cells were detectable with microscopic observations and flow cytometry (Figure 23).

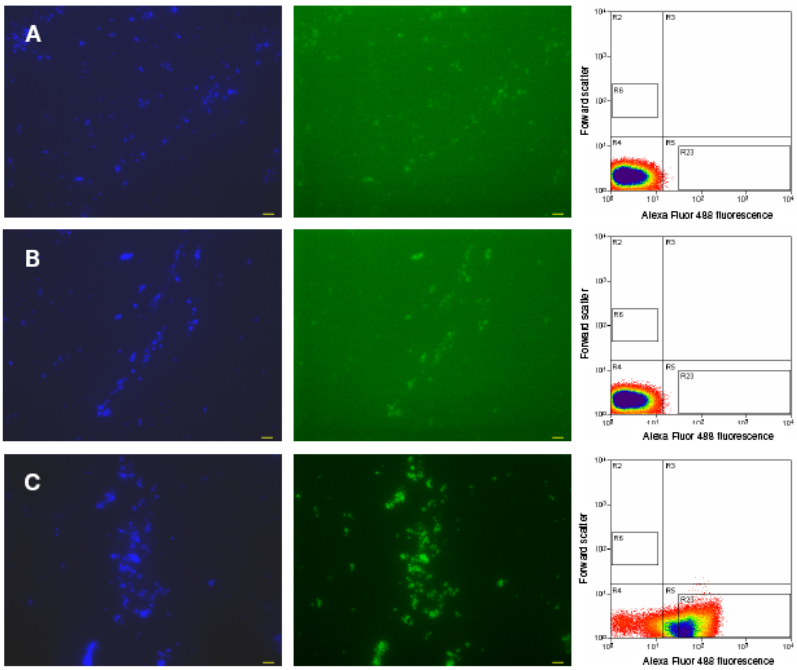


Figure 23. DAPI (left column) and Alexa Fluor 488 (middle column) fluorescence microscopy of the same examining field, and flow cytometry analysis (right column) of strain 195. The scale bar represents 5 μm. Each row indicates a sample: (A) no probe, (B) Non338 and (C) *Dehalococcoides*-specific probe. Each flow cytometry plot contains ~120,000 events.

Using a commercially available WGA kit, two identically sorted samples from the same bottle of strain 195 were amplified and hybridized onto microarrays (Figure 24A). Based on the hybridization results, 1 ng of starting DNA yielded high quality amplification. In fact, with the exception of only four genes (DET1089, DET1091, DET1092, DET1094) in one of the two samples, all genes across the genome were deemed present but genes near genome positions 1100 and 1400 was noticeably under-represented when compared against gDNA (Figures 24A and 24B). Although nearly all genes across the genome were amplified, intra-sample variability in amplification was observed as the coefficient of variation (CV) across all probe sets for gDNA of strain 195 (Figure 24B) was only 15.9%, but it was 35.7% and 33.7% for the two amplified samples. Besides intra-sample variability, differences in signal intensity for respective genes between the two identically prepared samples also illustrated inter-sample variability, mostly on a region-to-region basis across the genome (Figure 24C). In order to minimize amplification biases, two samples were sorted from each of the duplicate cultures and equal masses of the amplified DNA were pooled prior to microarray analysis (Figure 24D). With this procedure, the CV of probe sets improved slightly to 30.9%, all genes were positively detected, and this averaging process improved the signal intensity near genome position 1100 and elsewhere when compared to results from a single amplification (Figures 24A and 24D).

After verifying the FACS-WGA method with a pure culture, the protocol was applied to two enrichment cultures containing different concentrations of *Dehalococcoides*: 1) previous described groundwater enrichment MethB12 (containing ~ 10% *Dehalococcoides* whose genome is similar to strain 195); 2) enrichment AD14 (containing ~ 1% *Dehalococcoides* whose genome is similar to strain VS). With a starting of 10^5 cells of MethB12, FACS-WGA sample gave

21.6% false negative results on a quad-genome microarray comparing with results of directly applying gDNA on the microarray, while a starting of 10^6 cells gave 4% false negative results. Therefore, a quantity of 10^6 cells suggests a good start point using FACS-WGA method (Figure 25A). With a starting of 10^6 cells of AD14, the sample going through FACS-WGA process has an essentially similar genomic distribution comparing to the gDNA sample, with only 1.4% false-negative results detected (Figure 25B).

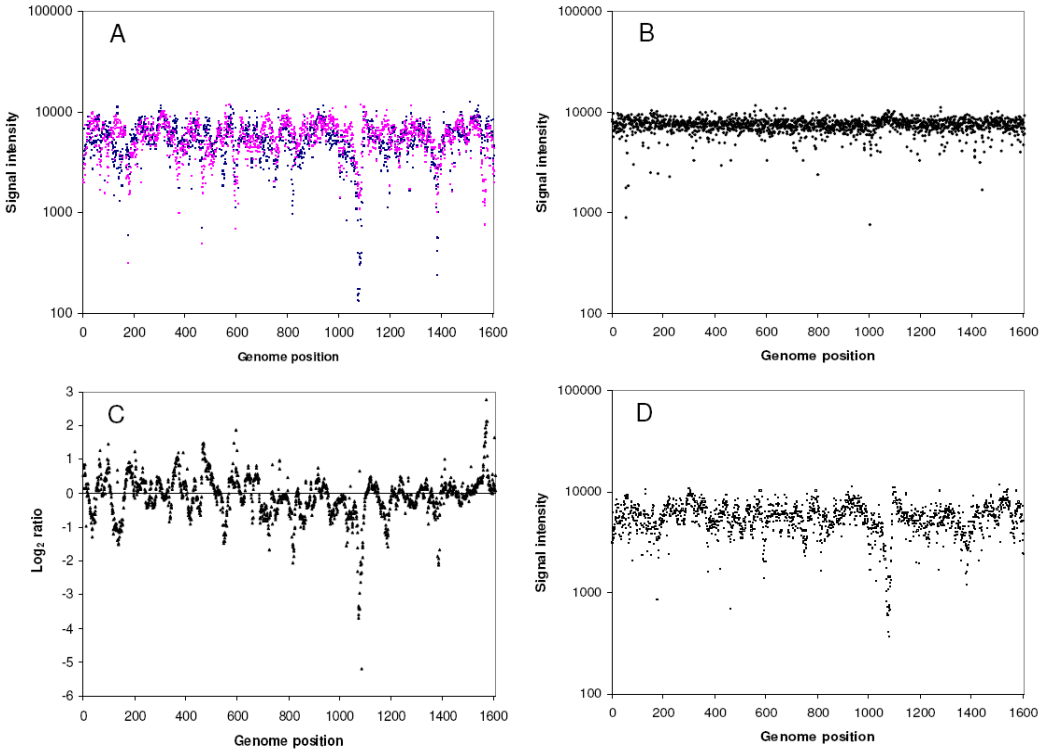
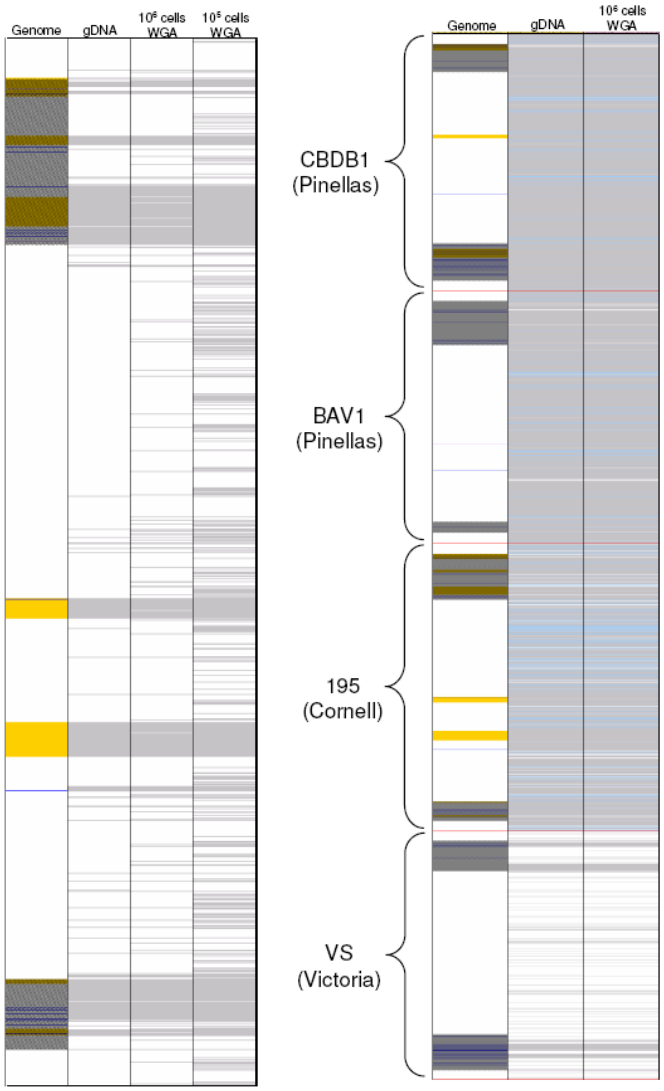


Figure 24. (A) Microarray signal intensity of two (pink and dark blue) independently sorted and amplified samples (10^6 cells) from the same bottle of strain 195. For each panel, a data point represents a gene of strain 195 arranged according to its location in the genome from DET0001 going left to right. (B) Signal intensity from hybridization of gDNA of strain 195. (C) Ratio of signal intensity between the two amplified samples shown in (A). (D) Signal intensity from pooling equal proportion of four independently sorted and amplified samples from duplicate cultures.



(A) (B)

Figure 25. (A) Microarray results from hybridization of gDNA and amplified DNA from 10^6 and 10^5 cells of MethB12. Each column represents a sample as indicated on the top and each row represents a gene of strain 195 where all the genes that are targeted by the microarrays (except a small number that tends to cross hybridize to nonspecific DNA) are depicted and arranged according to their location in the genome from top to bottom starting from DET0001. (B) Microarray results from hybridization of gDNA and amplified DNA from 10^6 cells of enrichment AD14. Genes in the four sequenced genomes (marked on the left along with the subgroup classification and separated by a red horizontal line) are shown and the color light blue is used to indicate genes that are targeted by the same probe sets as genes in strain VS. Genes that are considered present are colored white and grey indicates those that are absent. The 'genome' column on the left depicts annotations that are of interest, including the two HPRs (dark shade), IEs I to IX (orange), putative RDases (dark blue), and the *pceA* and *tceA* genes (purple).

We will extend the method to examine enrichment cultures that only contain dilute quantities of *Dehalococcoides* and samples collected directly from the environment. The developed method in combination with the genus-wide microarrays will enable the identification

of *Dehalococcoides* genes that are present in environmental samples in a high-throughput manner as no time-consuming and laborious culturing steps are required.

3.3 Metagenomic analysis of a TCE degrading microbial community (ANAS) (This study is published in Brisson et al., ISME J, 2012)

As one of our extended objectives, we applied metagenomic analysis to investigate functional genes present in dechlorinating microbial communities. In addition to 16S rRNA-based and mRNA-based microarray approaches, which are targeting at phylogenetic characteristic and functions within *Dehalococcoides* species, respectively, metagenomic data provide a broad view of the genetic composition of a community, which includes information about both the identity and potential metabolic capabilities of all the members.

Materials and methods. Dhc contigs were identified in a two stage sequence similarity (SS) process. In the first stage, contigs were identified by comparison to previously sequenced Dhc genomes. This analysis was performed using the genomes of Dhc strains 195, BAV1, CBDB1, VS and GT to identify as many Dhc contigs as possible in the metagenome. The genome sequences of reference Dhc strains were retrieved from the National Center for Biotechnology Information (NCBI) genome database [ftp://ftp.ncbi.nlm.nih.gov/genomes] and blastn (48) was used to compare the full reference genome sequence against a database of all metagenome contig sequences. The top 250 BLAST hit contigs for each reference genome were selected for the second stage of comparison, where their identities were checked by comparison to the NCBI chromosome database using megablast (48). All contigs whose top BLAST hit (based on expect value) in the chromosome database was to Dhc were identified as Dhc contigs.

ANAS contigs were grouped by tetranucleotide frequencies (TF) using a procedure based on one described by Dick *et al*. (15) with some modifications described here. Briefly, clustered regularly interspaced short palindromic repeat (CRISPR) and rRNA gene sequences were removed from contig sequences prior to classification because these sequence regions are known to have atypical nucleotide compositions compared to their genomes (15, 35). Next, all contig sequences larger than 2500 bp were selected for classification, with contigs > 7500 bp divided into 5000 bp fragments. Sequence fragments were classified based on TF using the Databionics ESOM Tools program. Dick et al.'s method (15) was used for clustering with the following modifications. Online training was used instead of the k-batch algorithm because online training provides more accurate, albeit slower performance. A map size of 120x196 and an initial radius of 60 were selected based on the size of the dataset.

Classes were given tentative taxonomic identifications based on 16S and 23S rRNA gene sequences associated with contigs in each class based on BLAST searches using the NCBI website [http://blast.ncbi.nlm.nih.gov/Blast.cgi]. Searches were performed in August 2010 using the megablast program and the non-redundant nucleotide collection database (48).

Results and discussion. Classes were identified for *Clostridium, Dehalococcoides, Desulfovibrio, Methanobacterium, Methanospirillum,* a Spirochaete, and a Synergistete. 250 contig sequences were identified as *Dehalococcoides*. *Dehalococcoides* sequence found in the metagenome was compared to the results of analysis of ANAS DNA on a *Dehalococcoides*

genus-wide microarray. Based on this comparison, the metagenome was estimated to have greater than 95% coverage of the *Dehalococcoides* genomes present in ANAS. The gene annotations of all *Dehalococcoides* contigs were searched for genes not present in known *Dehalococcoides* genomes. Results revealed 403 genes not found in other known and fully sequenced Dhc genomes. The most notable finding of this search was the presence of several genes that appear to constitute a complete or near complete cobalamin synthesis pathway, since no known Dhc strains are capable of synthesizing this important cofactor. PCR and sequencing were used to confirm the presence of these genes in a *Dehalococcoides* strain previously isolated from ANAS. The possible cobalamin synthesis capabilities of this *Dehalococcoides* strain are being further studied. In addition to Dhc, several other members of this community appear to have complete or near-complete cobalamin biosynthesis pathways. Moreover, eighteen putative reductive dehalogenases were identified, including 10 novel ones, all associated with Dhc. Genes for hydrogenases (271 in total) were widespread across all classes in this community, which is indicative of the important role that hydrogen metabolism plays in this community. DNA sequences from other major microorganisms in ANAS community were identified within the metagenome. This includes sequences from *Clostridium* (at least 2 strains), *Desulfovibrio*, *Methanobacter*, and *Methanospirillum*.

The description of the community composition provided by analysis of the metagenomic sequence is consistent with those of previous 16S rRNA gene clone library studies (27, 36) and the Phylochip study presented here. Overall, the clone libraries and metagenome sequencing agreed on the most abundant bacterial taxa, which were also detected by the Phylochip. Because it is much more sensitive to low abundance organisms, the Phylochip also detected many other taxa (8, 14). The archaeal taxa found in the metagenome were also detected by the Phylochip, along with several other archaea. No Archaeal clone libraries have yet been prepared for this community.

One notable discrepancy between the bacterial clone libraries and the metagenome was in the relative abundance of taxa detected by the two methods. Specifically, the Spirochaete exhibited only low abundance (1-2% of clones) in both of the clone library experiments (27, 36). Based on the median contig length and average read depth in the metagenome data, however, the Spirochaete appears to be one of the more abundant organisms in this community. The use of clone libraries to estimate the relative abundance of different taxa is known to be susceptible to PCR and cloning biases (43), and some studies have found Spirochaetes in particular to be underrepresented in some clone libraries (9, 22). However, recent studies also suggest that estimates of relative abundance based on metagenomic sequencing read depth may also be biased (4, 32).

Each of the three approaches to understanding community composition discussed here has inherent advantages and disadvantages. Clone libraries provide a survey of known and unknown organisms, emphasizing the quantitatively dominant community members. This approach has a bias toward organisms with high 16S rRNA gene copy numbers, biases associated with differential sequence primer affinity and PCR conditions, as well as cloning biases (43). It also is likely to miss sequences from very low abundance organisms. While the Phylochip is much more sensitive to low abundance community members, it only allows identification of taxa included on the chip, so any novel organisms would be missed (8,

14). Further, due to variation in probe hybridization across taxa, the Phylochip does not provide information on the relative abundance of different organisms. Metagenomic sequencing does not involve either a PCR step or *a priori* knowledge of bacterial sequences, so it avoids biases associated with primer or probe binding and novel organisms or genes can be identified. However, metagenomic sequencing only provides good coverage of the more highly abundant organisms, so identifying sequences for low abundance organisms may be missed by this approach (26). Metagenome sequences also provide a small amount of information on relative abundance from the read depths of different sequences (26), although these estimates can also be biased (4). One recent study examined reads from both Sanger sequencing and pyrosequencing of DNA from a constructed mix of microorganisms (32). This study found that the two methods produced read compositions that differed from each other and from the predictions based on the known composition of the sample. Finally, clone libraries, the Phylochip, and metagenomic sequencing are all subject to biases induced by DNA extraction and preparation methods (32, 43).

One major advantage of metagenomic sequencing is that it provides information on novel genes. This study identified what appears to be a complete or near-complete cobalamin synthesis pathway in ANAS Dhc. Both the SS and TF methods identified the contig containing genes for this pathway as belonging to Dhc. This finding was unexpected because known Dhc strains are not capable of cobalamin synthesis (20, 25, 39). However, the possibility of miss-assembly of metagenome contigs means that cobalamin synthesis genes from some other organism could have been co-assembled with Dhc sequence. This possibility was investigated by comparing the TF for the sequence regions containing these genes to the original TF classification of contigs. The regions containing the cobalamin synthesis genes grouped with the Dhc sequences and not with any of the other classes of sequences. This indicates that these genes were not mis-assembled or recently horizontally transferred to Dhc, but have been maintained in the ANAS Dhc for some time.

This analysis of metagenomic sequence data has advanced our understanding of this dechlorinating microbial community. The phylogenetic composition of ANAS described by the metagenomic sequencing generally confirms the composition described by Phylochip and previous 16S clone library studies. However, the metagenomic sequencing and the 16S clone libraries disagree on the relative abundances of some taxa. In particular the metagenomic data indicates a much higher abundance of Spirochaetes than the clone libraries. Our analysis of functional genes relevant to dechlorination provides insight into the capabilities of the microbial community members. This analysis indicated that Dhc appear to be the only members of ANAS capable of reductive dechlorination. Genes related to the synthesis of cobalamin, an important cofactor for reductive dechlorination, are widespread, highlighting the importance of this cofactor in the function of the ANAS microbial community.

Objective 4: Evaluating correlations between quantitative biomarker detection and solvent degrading activity.

This objective is to coalesce the quantitative measurements of the identified rRNA-based phylogenetic biomarkers, the mRNA-based functional biomarkers, as well as physiological characteristics into predictive models that can be applied to uncharacterized microbial communities. This includes the correlation between *Dehalococcoides*-related biomarkers and TCE dechlorination activity, and the correlation among supportive microorganism-related biomarkers, the growth and dechlorination activity of *Dehalococcoides* species.

4.1 Quantification and correlation of RDase gene copies with 16S rRNA gene of Dehalococcoides at a TCE-contaminated site at Ft. Lewis East Gate Disposal Yard (EGDY) (Tacoma, Washington) (This study is published in Lee et al., Appl. Environ. Microbiol., 2008)

Materials and methods. Materials and methods used in this study were described in the relevant section in "Objective 1".

Results and discussion. qPCR was used to monitor the concentrations of 16S rRNA, *tceA*, *bvcA*, and *vcrA* genes of *Dehalococcoides* spp. in treatment plots over a 1-year monitoring period spanning biostimulation and bioaugmentation activities (Figure 26). At the baseline sampling event (April 2005), the *Dehalococcoides* 16S rRNA gene was detected at $1.2 \pm 0.20 \times 10^4$ copies/liter and $1.1 \pm 0.13 \times 10^5$ copies/liter in plots 1 and 2, respectively (Fig. 13). Redox conditions at the site during the baseline sampling event indicated predominantly aerobic conditions, with plot 2 exhibiting slightly more reducing conditions than plot 1. Genes associated with VC degradation, *bvcA* and *vcrA*, were present in both treatment plots, while the *tceA* gene was detectable only in plot 2 initially. One month after the commencement of whey injection (July 2005), significant increases in gene copies for *Dehalococcoides* spp. were detected. Notably, the 16S rRNA genes increased 32-fold in plot 1 and only 2.5-fold in plot 2. Furthermore, the *tceA* gene became detectable in plot 1, and the *bvcA* gene in both plots increased by over an order of magnitude. The dominance of *bvcA* among the three monitored RDase genes was amplified following biostimulation, accounting for the majority of the total RDase genes measured.

Following the July 2005 sampling, bioaugmentation was implemented at the site with an enrichment culture that had a *Dehalococcoides* 16S rRNA gene concentration of $4.3 \pm 0.72 \times 10^7$ copies/ml, and it contained the *tceA* ($5.1 \pm 0.28 \times 10^7$ copies/ml) and *vcrA* ($3.6 \pm 0.12 \times 10^7$ copies/ml) genes but no *bvcA* gene. For 3 months after bioaugmentation (August to November 2005), while all monitored genes increased to some extent, the most significant change was the order-of-magnitude increase in the *vcrA* gene in both plots (Figure 26). By February 2006, 3 months after the whey injection strategy was altered, the 16S rRNA gene concentrations in plot 1 increased by 2 orders of magnitude while those in plot 2 increased by 19-fold (Figure 26).

In parallel with the significant increase in the 16S rRNA gene was the increase in the three RDase genes, with the *bvcA* gene concentration continuing to dominate in numbers. During the 1-year biostimulation-and-bioaugmentation period, the 16S rRNA gene concentrations increased by more than 3 orders of magnitude in both plots.

The 16S rRNA gene sequences of distinct *Dehalococcoides* strains are highly conserved, making it impossible to correlate different metabolic functions with *Dehalococcoides* strains using this gene. Primers and probes that target the 16S rRNA gene in this study were therefore designed as universal primers to target 16S rRNA sequences of all known *Dehalococcoides* strains.

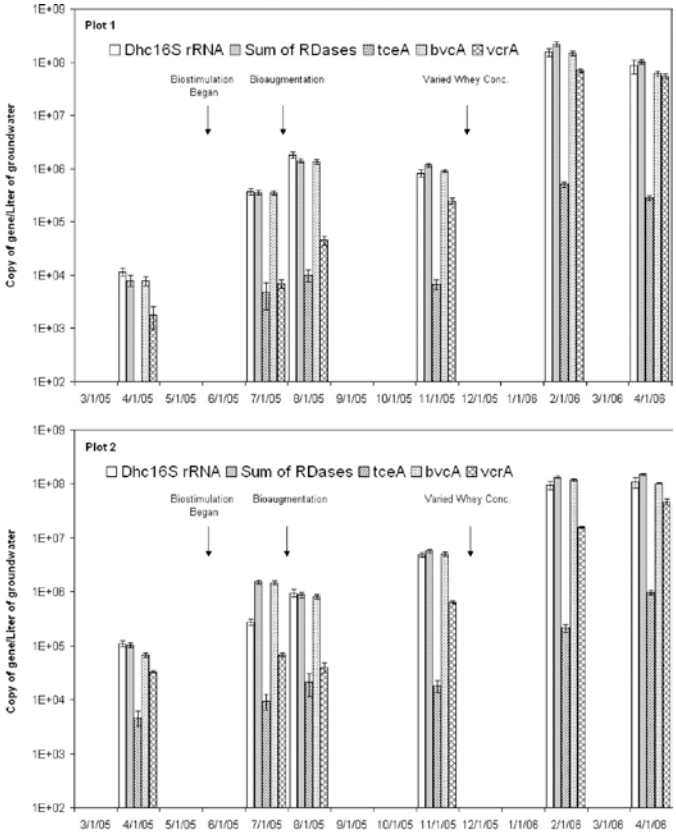


Figure 26. Dynamics of the *Dehalococcoides* (Dhc) 16S rRNA gene and RDase genes concentrations over the 1-year monitoring period for treatment plots 1 and 2. Manipulations implemented at the site are indicated by arrows on the graphs. Data at each time point are averages for samples from the eight monitoring wells in the respective plot, and each error bar represents one standard deviation for the qPCR method.

It was one of our goals in this study to evaluate whether *tceA*, *vcrA*, and *bvcA* genes can be used to correlate different functional *Dehalococcoides* groups and if the sum total of the three RDase genes can adequately predict the total *Dehalococcoides* populations from environmental samples. A good linear correlation ($r^2 = 0.99$) between the 16S rRNA gene and the sum total of the three RDase gene concentrations suggested that their quantification was highly representative of the RDase genes for the major *Dehalococcoides* metabolic functions at this site (Figure 27). The close tracking of copy numbers between RDase and 16S rRNA genes was observed both during the early phases of treatment, when *bvcA* was the only dominant RDase gene, and during the later phases of treatment, when both the *vcrA* and *bvcA* genes were dominant. The close agreement observed between the 16S rRNA gene concentration and the RDase gene sum across several orders of magnitude of detection suggests that the majority of *Dehalococcoides* cells at

this field site carry a *tceA*, *vcrA*, or *bvcA* gene and that the number of *Dehalococcoides* cells with none of these RDase genes is relatively insignificant. Furthermore, the three RDase genes were determined to be present in separate *Dehalococcoides* cells, since independent variation in the ratio of gene concentration over time indicates that co-location of any of the genes is unlikely.

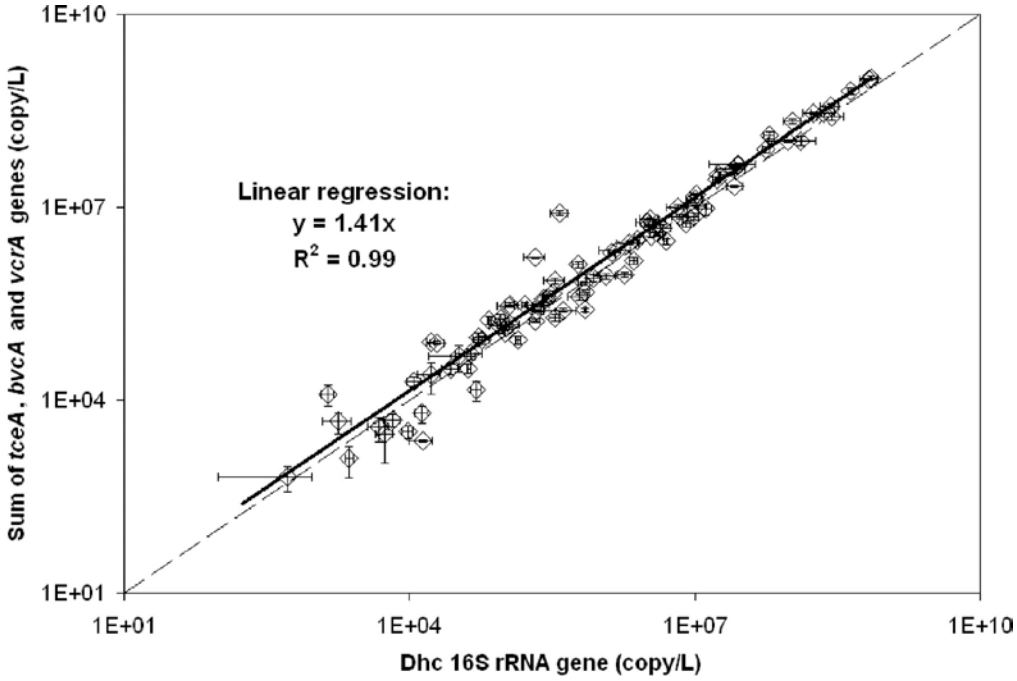


Figure 27. The sums of RDase genes concentrations are compared against the *Dehalococcoides* (Dhc) 16S rRNA gene concentrations on a log-log scale. Each data point represents a sample collected at a monitoring well from treatment plots 1 and 2 during each sampling event. Each error bar represents one standard deviation for the qPCR method. The dashed line represents the hypothetical 1:1 correlation between the two variables.

The results summarized above were based on DNA tracking methods which can effectively determine the presence or absence of particular bacteria and their associated functions, but do not necessarily indicate the functions that are actually active. In order to verify that the detected genes (16S rRNA gene and RDase genes) were being expressed, RT-qPCR was applied to the more labile RNA molecules for the final two sampling events (February 2006 and April 2006). The 16S rRNA genes of *Dehalococcoides* were highly expressed in the selected monitoring wells for both sampling events, with concentrations between $9.7 \pm 2.1 \times 10^1$ and $8.7 \pm 1.4 \times 10^2$ transcripts/gene (Figure 28A). The *tceA*, *vcrA*, and *bvcA* transcripts were also detected, but the copies per gene were about 1 to 3 orders of magnitude lower than those for the 16S rRNA transcripts. Samples without detectable *tceA* gene copies (e.g., MW 1A4, February 2006) (Figure 28B) unsurprisingly had no detectable transcripts, but there were two samples (MW 2D4, February 2006, and MW 1A4, April 2006) with no detectable transcripts despite the presence of a detectable gene. As negative controls, samples collected from two up-gradient aerobic monitoring wells that were not impacted by whey injection had no detectable *Dehalococcoides* 16S rRNA or RDase transcripts.

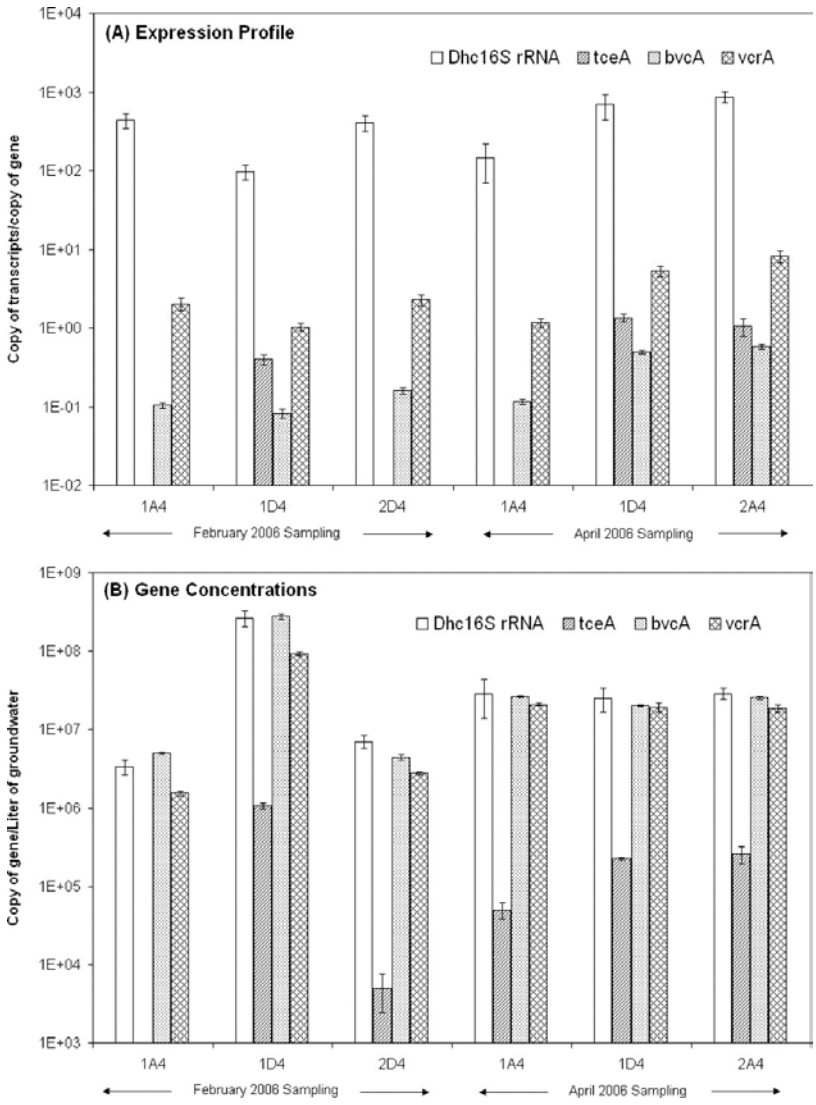


Figure 28. Expression profile of the *Dehalococcoides* (Dhc) 16S rRNA gene and the RDase genes during the February 2006 and April 2006 sampling events (A). Data were calculated from triplicate RT-qPCRs, and each error bar represents one standard deviation. Labels on the *x* axis represent monitoring well designations. Gene concentrations at the corresponding monitoring wells. Data were calculated from triplicate qPCRs, and each error bar represents one standard deviation (B).

This study demonstrates that quantifying genes and transcripts of the 16S rRNA gene and three characterized RDase genes of *Dehalococcoides* during different stages of bioremediation can provide important and complementary information about the dynamics and activity of these organisms.

4.2 Investigation of the correlation between growth phases and identified biomarkers (this manuscript is currently under preparation for submission)

Materials and methods. In order to correlate growth phase to certain biomarkers, changes of gene expression of the ANAS culture at three phases throughout one feeding cycle, from nutrient-excess to nutrient-limited, were investigated by applying RNA extracted from cells grown after 27 hours, 57 hours and 13 days on the genus-wide microarray (Figure 29). cDNA was first synthesized from RNA samples using superscript II reverse transcriptase (Invitrogen, Carlsbad, CA) according to instructions given in section 3 of the Affymetrix GeneChip Expression Analysis technical manual. The following procedures are the same as described previously.

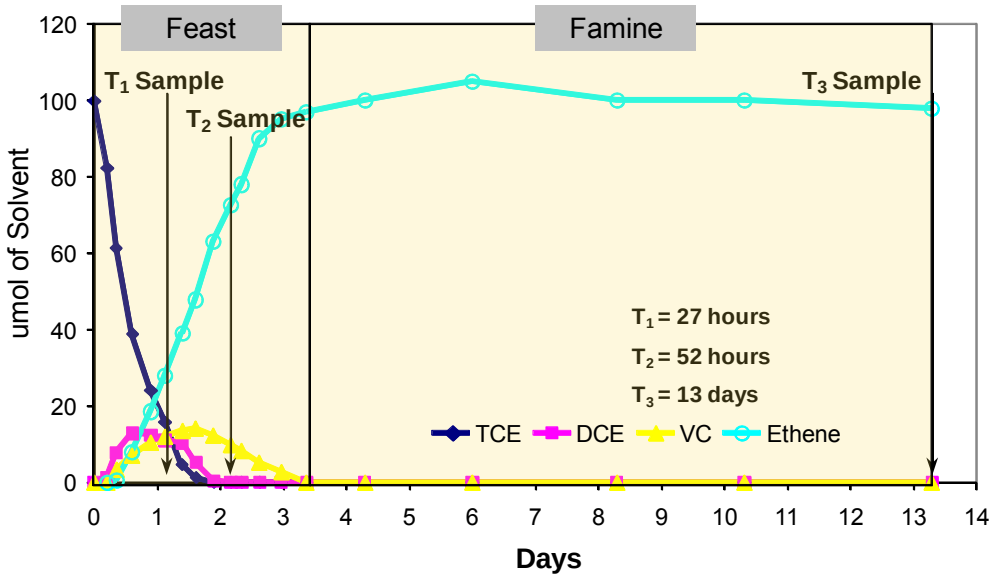


Figure 29. The feeding cycle of ANAS and sampling points.

Results and discussion. First, gDNA was applied on the genus-wide microarray querying what genes were present in ANAS culture. The three most abundant groups of functional genes present in ANAS were 1) translation, ribosomal structure and biogenesis; 2) energy production and conversion; 3) general function prediction only. Secondly, RNA samples taken from different phases were applied on the genus-wide microarray, and gene expression was quantified. 35% of the genes detected in ANAS were differentially expressed throughout a feeding cycle, among which the most changed ones were those with the functions of translation, ribosomal structure and biogenesis; energy production and conversion; amino acid transport and metabolism (Figure 30). This suggested that these functional genes could serve as potential biomarkers indicative of different dechlorination phases and the nutrient stress.

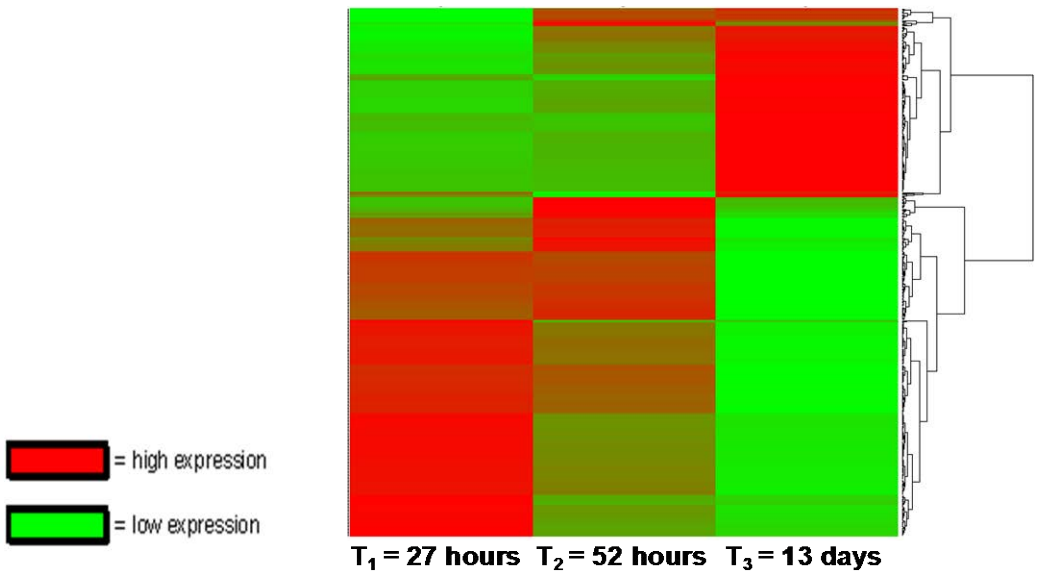


Figure 30. Hierarchical clustering analysis of time-course ANAS samples. Individual genes are represented by row; time points are represented in columns. Color changes indicate changes in gene expression. Genes are clustered on right based on expression analysis.

4.3 Investigation of the correlation between dechlorination and cobalamin riboswitches (This study is published in Johnson et al., FEMS Microbiol letters, 2009)

Another potential biomarker, the putative cobalamin riboswitches, were identified in the genome of strain 195. Since preliminary studies showed that ANAS was able to degrade TCE with an extreme low level of vitamin B_{12} (1 µg/l), the effects of exposure to limited, excess of cyanocobalamin, as well as ANAS spent medium on the whole genome expression of strain 195 were investigated. And the correlation between expression level of cobalamin-related biomarkers and cobalamin corrinoids availability will also be determined.

Materials and methods. Seventy-two parallel bottles containing basal medium were inoculated with strain 195 and amended with one 3-ml dose of TCE. Of these bottles, 36 contained 1 mg/l cyanocobalamin while the other 36 contained 100 mg/l cyanocobalamin. After complete conversion of trichloroethene to VC and ethene, an additional 3-mL dose of trichloroethene was added to each bottle, and the bottles were incubated for an additional 24 h to ensure that cultures were actively growing and dehalogenating. After that, bottles were sacrificed and RNA was extracted. ANAS spent medium was periodically removed from the ANAS batch-reactor and stored anaerobically under N_2/CO_2 (90 : 10, v/v) headspace at room temperature. For this investigation, 400ml of this spent medium was centrifuged (12, 000 g), and the supernatant was filter-sterilized using hydrophilic Durapore membrane filters (47mm diameter, 0.22 mm pore size) (Millipore, Billerica, MA) in an anaerobic chamber. For this experiment, 30 parallel bottles containing 80ml basal medium and 1 µg/l cyanocobalamin were inoculated with strain 195 and amended with one 7 µl dose of TCE. After complete conversion of TCE to VC and ethene, a second 7 µl dose of TCE was added to each bottle and the bottles were incubated for an additional 24 h to ensure that cultures were actively dehalogenating. Thereafter, the bottles were amended with 20ml of spent ANAS medium or 20ml of fresh ANAS medium to serve as

controls. After amendment, the bottles were incubated for a further 24 h and GC was used to verify that TCE was still present in the bottles. The cultures were then sacrificed and cells were collected by filtration. RNA was extracted from cell pellets. All RNA samples under different conditions were applied on whole genome microarray of strain 195 as described previously in "Objective 2".

Results and discussion.

Between limited and excess cyanocobalamin conditions, 20 genes were differentially regulated in total. Of these, 19 genes were downregulated (Table 4) while only one gene was up-regulated by excess cyanocobalamin (data not shown). Of the 19 genes that were down-regulated by excess cyanocobalamin, the most highly down-regulated genes were DET0125 (5.5-fold) and DET0126 (7.3-fold) (Table 4), which are co-transcribed and contain an upstream putative cobalamin riboswitch. Also down-regulated was a number of genes with annotated functions related to corrinoid metabolism and transport, including five genes located within a duplicated set of operons that encode a cobalamin ABC-type transport system and a CobD homologue (DET0650-0654/DET0684-0688). Another corrinoid-related gene that was down-regulated was DET0936, which encodes a CobQ homologue (Table 4). In other bacteria, CobQ is responsible for amidating several side chains of cobyric acid. The genes annotated to construct and attach the lower ligand base to cobyric acid (DET0657–0660/DET0691–0694) were not down-regulated by excess cyanocobalamin, even though these transcripts are predicted to have an upstream cobalamin-binding riboswitch.

Table 4. Summary of genes that were significantly down-regulated by excess cyanocobalamin (FDR<1%, >2-fold).

Locus tag*	Fold downregulation by excess cyanocobalamin (100 μg L^{-1})	Annotation*	Predicted operon unit	Detected riboswitch
DET0014	2.2	Hypothetical protein	DET0014	N
DET0125	5.5	Alcohol dehydrogenase, zinc-containing	DET0125–0126	Y
DET0126	7.3	Anthranilate phosphoribosyltransferase (TrpD)	DET0125–0126	Y
DET0297	2.9	Conserved hypothetical protein	DET0297	N
DET0298	5.0	Hypothetical protein	DET0298	N
DET0650/DET0684	2.3	ABC-type cobalamin Fe^{3+}-siderophores transport systems, periplasmic-binding protein	DET0650–0655/DET0684–0689	N
DET0651/DET0685	2.8	ABC-type cobalamin Fe^{3+}-siderophores transport systems, permease component	DET0650–0655/DET0684–0689	N
DET0652/DET0686	2.2	ABC-type cobalamin Fe^{3+}-siderophores transport system, ATP-binding protein	DET0650–0655/DET0684–0689	N
DET0653/DET0687	2.6	Conserved hypothetical protein	DET0650–0655/DET0684–0689	N
DET0654/DET0688	2.6	Cobalamin biosynthesis protein (CobD)	DET0650–0655/DET0684–0689	N
DET0936	2.7	Cobyric acid synthase (CobQ)	DET0936–0937	N
DET1049	3.3	HD domain protein	DET1049	N
DET1186	2.8	Conserved domain protein	DET1186	N
DET1303	2.7	Fasciclin domain protein	DET1303–1304	N
DET1307	4.8	Conserved hypothetical protein	DET1307	N
DET1324	2.5	Conserved hypothetical protein	DET1324	N
DET1379	2.3	Auxin-responsive GH3 protein homolog, putative	DET1378–1379	N
DET1515	2.0	Conserved hypothetical protein	DET1515	N
DET1580	2.0	Transcription regulator, TetR family	DET1580	N

*Annotations and locus tags were obtained from Seshadri *et al.* (2005).

For strain 195 exposed to ANAS spent medium, in total, 119 genes were differentially expressed, with 16 downregulated (Table 5) and 103 upregulated (data not shown). As observed

with excess cyanocobalamin, DET0125 and DET0126 were among the most highly down-regulated genes after exposure to ANAS spent medium (2.6- and 11.6-fold, respectively) (Table 5). In contrast to excess cyanocobalamin, three of the four genes encoding the predicted system for constructing and attaching the lower ligand base to cobyric acid (DET0657–0659/DET0691–0693) were also down-regulated by ANAS spent medium (2.4–5.8-fold). Moreover, the genes encoding the cobalamin ABC- type transport system and the co-transcribed CobD homologue (DET0650–0654/DET0684–0688) were not down-regulated by ANAS spent medium, nor was the gene encoding the CobQ (DET0936) homologue. The genes that were up-regulated by ANAS spent medium are not co-localized within operons or within metabolic pathways to any significant extent.

According to the results, DET0125-0126 were among the most highly down regulated genes when cells were grown with excess cyanocobalamin or exposed to ANAS spent medium (Table 4 and 5). These genes have an upstream putative cobalamin-binding riboswitch, although neither gene encodes functions with obvious roles in corrinoid metabolism or transport annotations. More experiments are needed to elucidate the role of these genes on cobalamin-salvaging system and thus validate the potential of these genes to be biomarkers indicative of cobalamin stress in a microbial community. The other set of genes that contains an upstream putative cobalamin riboswitch [DET0657–0660/DET0691–0694 (cobT, cobS, cobC, and cobU)] does have a clear function related to corrinoid metabolism. The products of these genes are predicted to construct the lower ligand base (5, 6-dimethylbenzimidazole) that is attached to cobyric acid. The differential expression of cobT, cobS, cobC, and cobU by ANAS spent medium but not by excess cyanocobalamin, along with the detection of an upstream putative cobalamin riboswitch, provides preliminary evidence that the corrinoids taken up from ANAS spent medium likely have different lower ligand bases than 5,6-dimethylbenzimidazole. So, these genes could also serve as biomarkers, which are correlated to the presence/absence of cobalamins with favorable lower ligands used by *Dehalococcoides* species

Further analysis of DET0126 suggests that its product likely performs a function different from TrpD. First, strain 195 contains a second divergent copy of trpD (DET1483) that is part of a putative eight-gene operon containing other trp genes (39), none of which were differentially regulated under any of the conditions tested in this study. Second, DET1483 shows close homology with proteins from other members of the Chloroflexi, while DET0126 does not. Instead, DET0126 shows close homology with proteins from Desulfotomaculum reducens of the Firmicutes. Finally, analyzing DET0126 with the Orthologue Neighborhood Viewer (http://img.jgi.doe.gov) indicates that nearly all homologues have an upstream gene predicted to encode an alcohol dehydrogenase. Additional experiments are now needed to elucidate the functions of DET0125–0126 and to examine whether they have a role in corrinoid metabolism.

The other set of genes that contains an upstream putative cobalamin riboswitch [DET0657–0660/DET0691–0694 (cobT, cobS, cobC, and cobU)] does have a clear function related to corrinoid metabolism. The products of these genes are predicted to construct the lower ligand base that is attached to cobyric acid (17, 29). Typically, this is 5,6-dimethylbenzimidazole in bacteria (including cyanocobalamin) and 5-hydroxybenzimidazole in methanogenic Archaea (13, 17). When cells were grown with excess or limiting cyanocobalamin, all four of these genes were expressed at stable (< 2-fold) but very high levels, having hybridization intensities among

the top 7% within the entire nonredundant protein-coding genome. The high expression of these genes independent of the cyanocobalamin concentration suggests two possibilities for their functional role. First, cobT, cobS, cobC, and cobU might be used to modify the lower 5, 6-dimethylbenzimidazole ligand present in cyanocobalamin. Alternatively, these genes might function in the reverse direction to cleave the lower ligand base from cyanocobalamin. In contrast with cyanocobalamin, exposure to ANAS spent medium resulted in the down regulation of cobT, cobS, and cobC (Table 5). cobU was also down regulated (1.8-fold), but did not meet our stringent criteria for differential expression (FDR < 1%; > 2-fold). Although their exact function in strain 195 is unclear, the differential expression of cobT, cobS, cobC, and cobU by ANAS spent medium but not by cyanocobalamin, along with the detection of an upstream putative cobalamin riboswitch, provides preliminary evidence that the corrinoids taken up from ANAS spent medium likely have different lower ligand bases than 5,6-dimethylbenzimidazole.

Together, the results presented here indicate that corrinoid-related genes respond differently to excess cyanocobalamin and to ANAS spent medium. This suggests that ANAS spent medium contains corrinoid forms different from cyanocobalamin and that strain 195 adjusts its metabolism according to the corrinoid forms available for uptake. It is noteworthy that only corrinoid-related processes were clearly identified as being differentially regulated in strain 195 after exposure to ANAS spent medium. No other metabolic systems were clearly identified, suggesting that corrinoid transfer is likely among the key interspecies interactions controlling the behavior of strain 195 within complex communities.

Table 5. Summary of genes that were significantly down regulated by ANAS spent medium (FDR < 1%, > 2-fold).

Locus tag*	Fold downregulation by exposure to ANAS spent medium	Annotation*	Predicted operon unit	Detected riboswitch
DET0125	2.6	Alcohol dehydrogenase, zinc-containing	DET0125–0126	Y
DET0126	11.6	Anthranilate phosphoribosyltransferase (TrpD)	DET0125–0126	Y
DET0467	2.2	3-Dehydroquinate synthase	DET0460–0468	N
DET0468	2.6	Phospho-2-dehydro-3-deoxyheptonate aldolase	DET0460–0468	N
DET0656	4.9	Hypothetical protein	DET0656	N
DET0657/ DET0691	5.8	Nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase (CobT)	DET0657–0660/ DET0691–0694	Y
DET0658/ DET0692	2.5	Cobalamin-5-phosphate synthase (CobS)	DET0657–0660/ DET0691–0694	Y
DET0659/ DET0693	2.4	α-Ribazole-5-phosphate phosphatase (CobC)	DET0657–0660/ DET0691–0694	Y
DET0748	2.1	Hypothetical protein	DET0748	N
DET0908	2.4	Arsenical pump membrane protein, putative	DET0908	N
DET0909	2.5	Membrane protein, putative	DET0909	N
DET1078	2.4	Tail tape measure protein, TP901 family	DET1077–1094	N
DET1088	2.0	Terminase, large subunit, putative	DET1077–1094	N
DET1094	2.0	HNH endonuclease domain protein	DET1077–1094	N
DET1125	2.4	Ammonium transporter	DET1122–1125	N
DET1296	2.0	Conserved hypothetical protein	DET1296	N

*Annotations and locus tags were obtained from Seshadri *et al*. (2005).

4.4 Identification of diverse corrinoid forms in Dehalococcoides-containing microbial communities (This manuscript is currently under preparation for submission)

As one of our extended objectives, we established a LC/MS/MS method to identify and quantify various corrinoid and lower ligand species (Figure 31), we further applied this method to identify diverse corrinoid forms present in *Dehalococcoides*-containing microbial communities, especially those without exogenous B_{12}.

Materials and methods. 13 corrinoid species targeted in this study are listed in Figure 31, together with the associated lower ligand bases. 5 of the benzimidozole lower ligands were also targeted, which are Bza, 5-MeBza, 5-OMeBza, 5-OHBza, and DMB. Cobalamin, Cbi, together with 4 lower ligands except 5-OHBza were purchased from Sigma Chemical Co. (St Louis, MO). Cobyric acid was abiologically synthesized using cobalamin, and all other corrinoid forms and 5-OHBza were biologically synthesized.

ANAS subcultures investigated in this study were seeded from ANAS. In 160-mL serum bottles, 95 mL defined basal medium was added, with 60 mL N_2/CO_2 (90:10) headspace. Lactate (25 mM final concentration) was amended as both the carbon source and electron donor, TCE (2 µL, ca. 0.2 mM final concentration) was supplied as the terminal electron acceptor, as well as 0.5 mL Wolin vitamin stock with a final concentration of 100 µg/L vitamin B_{12} as a co-factor. 5% of ANAS culture was inoculated into 160-mL serum bottles. After all TCE is dechlorinated to ethene, 5% of ANAS subculture was transferred into fresh medium using the same condition. In addition, another batch-culture was constructed by inoculating 5% ANAS subculture into 95 mL the same medium but without the addition of vitamin B_{12}. We use the term "ANAS" for ANAS subcultures and "ANAS*" for the subculture grown without vitamin B_{12}. Experiments were carried out after 5 times of subculturing ANAS and 3 times of subculturing ANAS*. Both ANAS and ANAS* were able to completely dechlorinate TCE to ethene. Groundwater enrichments were the same as described in "Objective 1".

Cells from 200-300 mL culture were collected by centrifugation at 15,000 × g for 10 min at 4 °C. Cells were stored in a 50 mL centrifuge tube at -80 °C before extraction. Supernatant was filtered by a 0.2 µm filter and loaded onto a Sep-Pak C_{18} cartridge (Waters, Milford, MA), which was activated by 3 mL methanol and 6 mL milliQ water. The cartridge was then washed with 50 mL milliQ water, and eluted with 3 mL 100% methanol. The eluate was collected in a 50 mL centrifuge tube and stored at -80 °C before extraction. For extraction, cells were resuspended in 20 mL 100% methanol. 20 mg KCN per gram of cells or per equivalent volume of supernatant was added into both cell and supernatant samples. All samples were then incubated in 60 °C water bath for 1.5 hours with thorough vortex every 20 min throughout the incubation. After the incubation, cell samples were centrifuged at 40,000 g, 4 °C for an hour to eliminate the cell debris. Supernatant collected after centrifugation of cell samples, together with supernatant samples, were dried by a rotary evaporator. The residue from each sample was redissolved in 20 mL milliQ water and then loaded onto a Sep-Pak C_{18} cartridge. The cartridge was washed with 10 mL milliQ water and eluted with 2 mL 100% methanol. The eluate was collected into two 1.5 mL micro-centrifuge tubes and vacuum-dried by an Eppendorf Vacufuge plus vacuum concentrator (Eppendorf, Hauppauge, NY). The residues in the two tubes were redissolved and combined in 200-300 µL milliQ water, resulting in a concentration factor of

1000 after the entire procedure. All samples were stored at -20 °C prior to LC/MS/MS measurement.

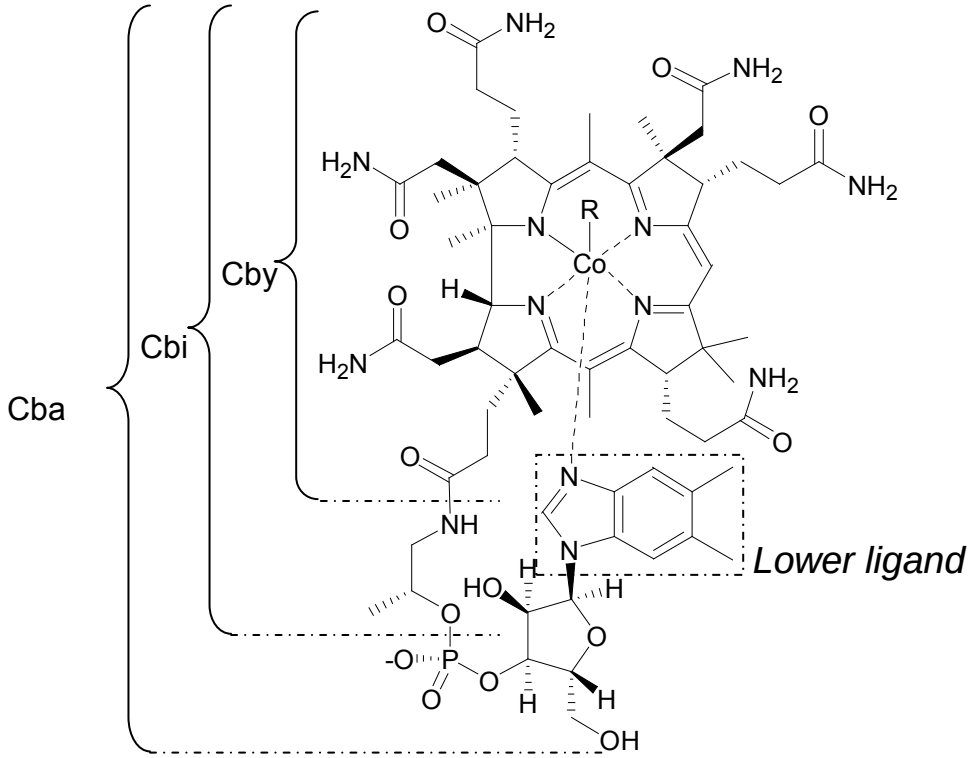
LC/MS/MS method was performed using Agilent 6410 triple quad LC/MS system (Agilent Technologies, Santa Clara, CA) and modified according to the methods described by Allen and Stabler (2). LC column was Agilent Eclipse Plus C_{18}, 1.8 µm, 3.0 × 50 mm (Agilent Technologies, Santa Clara, CA), with column temperature controlled at 40 °C. 0.5 mL/min gradient elution was used with initial solvent conditions of 82% miliQ water with 0.1% formic acid (solvent A) and 18% 100% methanol with 0.1% formic acid (solvent B) held for 3 min, increased to 21% B immediately and held for 2 min, increased to 100% B over 0.1min and held for 1 min, decreased back to 18% B over 0.1 min and held for 3.8 min. The injection volume was 10 µL. In order to identify the signature ions of each corrinoid and lower ligand species, MS2 scan was used with the fragmentor set at 135 volts, as well as product ion scan with the collision energy of 45 volts for each corrinoid and lower ligand species, except for [p-Cre]Cba and [Phe]Cba, which were quantified by method described by Allen and Stabler (2). Signature transitions were used for multiple reaction monitoring (MRM) on LC/MS/MS to fulfill quantitative analysis. Serial dilution of corrinoid and lower ligand standards in miliQ water with concentrations from 1 nM to 100 µM were used to establish calibration curves of each compound. The limit of detection (LOD) and the limit of quantification (LOQ) were also determined.

Results and discussion. With the LC/MS/MS method, we were able to detect 12 different corrinoid species and 5 lower ligand species with detection limits as low as 2 nM. The measurement of these compounds in ANAS and groundwater enrichment samples indicates that in ANAS subculture without the addition of vitamin B12, cobalamin was the most dominant corrinoid species, while in Meth and NoMeth, [p-Cre]Cba was about 10 times more than cobalamin (Table 6). [5-OHBza]Cba was the common corrinoid detected in methanogenic cultures, such as ANAS, ANAS* (ANAS subculture grown without B_{12}), MethB12 and Meth. [5-MeBza]Cba was detected in ANAS and ANAS*, but not in groundwater enrichments. [Ade]Cba and [2-MeAde]Cba were detected in ANAS*, as well as in groundwater enrichments. Cbi was also detected in all of the tested cultures, whereas Cby was not detected. Except for the dominant corrinoid species, the concentration of all other corrinoids were below the biologically significant level for *Dehalococcoides* (0.74 nM, c.a. 1µg/L), which was the experimental minimum B_{12} level for the growth and dechlorination of *Dehalococcoides* without negative effect.

Although *Dehalococcoides* spp. are not able to biosynthesize corrinoid *de novo*, their genomes possess genes involved in corrinoid uptake and salvage pathways, thus they are capable of corrinoid uptake, salvaging, and modification. So, in order to investigate the corrinoid forms originally produced by supportive microorganisms, as well as the way *Dehalococcoides* utilize them, we constructed a TCE-free enrichment (NoMeth_NT) from NoMeth by sub-culturing NoMeth into TCE-free medium, and all the other growth conditions are maintained the same as NoMeth. After 6 sub-culturing generations, corrinoid and lower ligand species were determined in both NoMeth and NoMeth_NT (Figure 32). A time-course result indicated little cobalamin production throughout the whole incubation (Figure 32B), indicating the production of cobalamin was associated with *Dehalococcoides*. For [p-Cre]Cba in NoMeth_NT, unlike a dramatic decrease in NoMeth, after 5 days it peaked and remained at ~ 22 nM, similar level as it

reached in NoMeth after 2 days. Moreover, about 20-50% of total [p-Cre]Cba was released into supernatant in NoMeth_NT, while almost all of the [p-Cre]Cba was kept in the cells in NoMeth (Figure 32A). Two other forms of complete corrinoid detected in NoMeth, [2-MeAde]Cba and [2-SMeAde]Cba, also increased in NoMeth_NT. Since the concentration of these two corrinoids were below biologically significant level during the whole incubation, they were considered of less importance than [p-Cre]Cba. More interestingly, although cobalamin was not produced when *Dehalococcoides* was absent, the associated lower ligand DMB was still detected in NoMeth_NT with a similar profile with that in NoMeth (Figure 32C), indicating an anaerobic production of DMB for other purposes rather than cobalamin biosynthesis in NoMeth_NT.

Corrinoid and lower ligand species, as well as microbial structure were further measured in NoMeth_NT+, in which NoMeth and TCE were re-amended. Not surprisingly, the growth of *Dehalococcoides* recovered to ~ 5×10^7 16S rRNA gene copies/mL after degrading a total of 4 µL TCE in NoMeth_NT+ (Figure 33B). In the meanwhile, the production of cobalamin also resumed to a total of 1.5 nM (Figure 33A), which corroborated that it was *Dehalococcoides* in the enrichment themselves that generated cobalamin for their own use in dechlorination process. However, no decreasing but slight increasing was observed for [p-Cre]Cba and DMB in NoMeth_NT+ and no change for the other less detected corrinoid forms (Figure 32A). Moreover, except for the recovered growth of *Dehalococcoides* and a stimulated growth of *Desulfovibrio*_GW, similar cell densities were observed for the other OTUs in NoMeth_NT, indicating little effect by the resume of dechlorination activity after all the other microorganisms were well grown already.



Benzimidazoles

Purines

Phenolics

5, 6-dimethylbenzimidazole
Cobalamin

Benzimidazole
[Bza]Cba

Adenine
[Ade]Cba

Phenol
[Phe]Cba

5-Methylbenzimidazole
[5-MeBza]Cba

5-hydroxybenzimidazole
[5-OHBza]Cba

2-Methyladenine
[2-MeAde]Cba

p-Cresol
[p-Cre]Cba

5-Methoxybenzimidazole
[5-OMeBza]Cba

5-Methoxy, 6-methylbenzimidazole
[5-OMe, 6-MeBza]Cba

2-Methylthioadenine
[2-SMeAde]Cba

Figure 31. Structures of corrinoid and lower ligand species together with abbreviated designation. Lower ligand name italized. Cby: cobyric acid; Cbi: cobinamide; Cba: cobamide.

Table 6. Corrinoid detection in different dechlorinating enrichment cultures.

Unit: nM	Cobalamin	[5-MeBza] Cba	[5-OHBza] Cba	[Ade] Cba	[2-MeAde] Cba	[2-SMeAde] Cba	[p-Cre] Cba	Cbi
Abiotic ctrl (with 74nM Cbl)	**83 ± 4**	0*	0	0	0.011 ± 0.005	0	0	0.34 ± 0.20
Abiotic ctrl (without Cbl)	0	0	0	0	0	0	0	0
ANAS sup	**66 ± 15**	0	0.052 ± 0.016	0	0	0	0	**0.78 ± 0.03**
ANAS cell	**2.3 ± 0.3**	0.012	0.24 ± 0.10	0.015 ± 0.003	0.015 ± 0.011	0	0.2 ± 0.1	0.068
ANAS* sup	0.073 ± 0.063	0	0	0	0.019 ± 0.010	0	0	0
ANAS* cell	**0.81 ± 0.02**	0.012 ± 0.007	0.27 ± 0.05	0.062 ± 0.059	0.59 ± 0.36	0.017	0.3 ± 0.1	0.039 ± 0.028
MethB12 sup	0.45 ± 0.17	0	0	0	0.0075 ± 0.0010	0	0.4 ± 0.2	0.18 ± 0.06
MethB12 cell	**9.0 ± 1.2**	0	0.10 ± 0.06	0.014 ± 0.0003	0.36 ± 0.27	0	**15.2 ± 1.3**	0.31 ± 0.18
Meth sup	0.17 ± 0.06	0	0.0065 ± 0.0037	0	0.012 ± 0.008	0	0.42 ± 0.01	0.040 ± 0.031
Meth cell	**3.2 ± 1.4**	0	0.22 ± 0.01	0.023 ± 0.019	0.13 ± 0.04	0	**5.89 ± 0.05**	0.13 ± 0.03
NoMethB12 sup	0.55 ± 0.14	0	0	0	0.0094 ± 0.0059	0	**0.80 ± 0.01**	0.26 ± 0.16
NoMethB12 cell	**1.4 ± 0.5**	0	0	0	0.032 ± 0.016	0	**6.0 ± 1.4**	0.64 ± 0.46
NoMeth sup	0.10 ± 0.04	0	0	0	0	0	0.38 ± 0.04	0.011 ± 0.001
NoMeth cell	**2.2 ± 1.5**	0	0	0.012 ± 0.006	0.026 ± 0.017	0	**5.5 ± 1.9**	0.56 ± 0.15

*: Below detection limit; Bolded figures indicate concentration above biologically significant level for *Dehalococcoides* spp. 0.74 nM (1 µg/L). The concentrations of [Bza]Cba, [5-OMeBza]Cba, [5-OMe, 6-MeBza]Cba, [Phe]Cba, and Cby in the listed cultures were all below detection limit, thus they were not shown in the table.

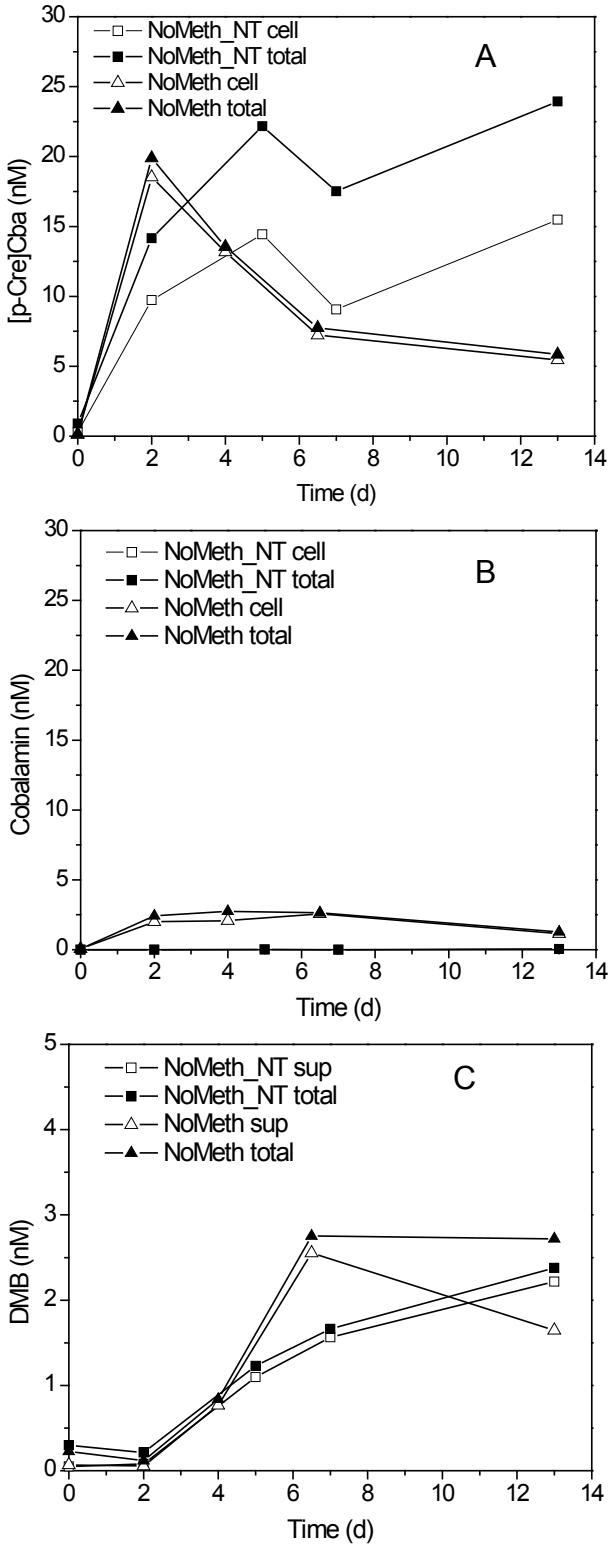


Figure 32. Temporal changes of [p-Cre]Cba, cobalamin and DMB in NoMeth and NoMeth_NT (A: [p-Cre]Cba; B: cobalamin; C: DMB).

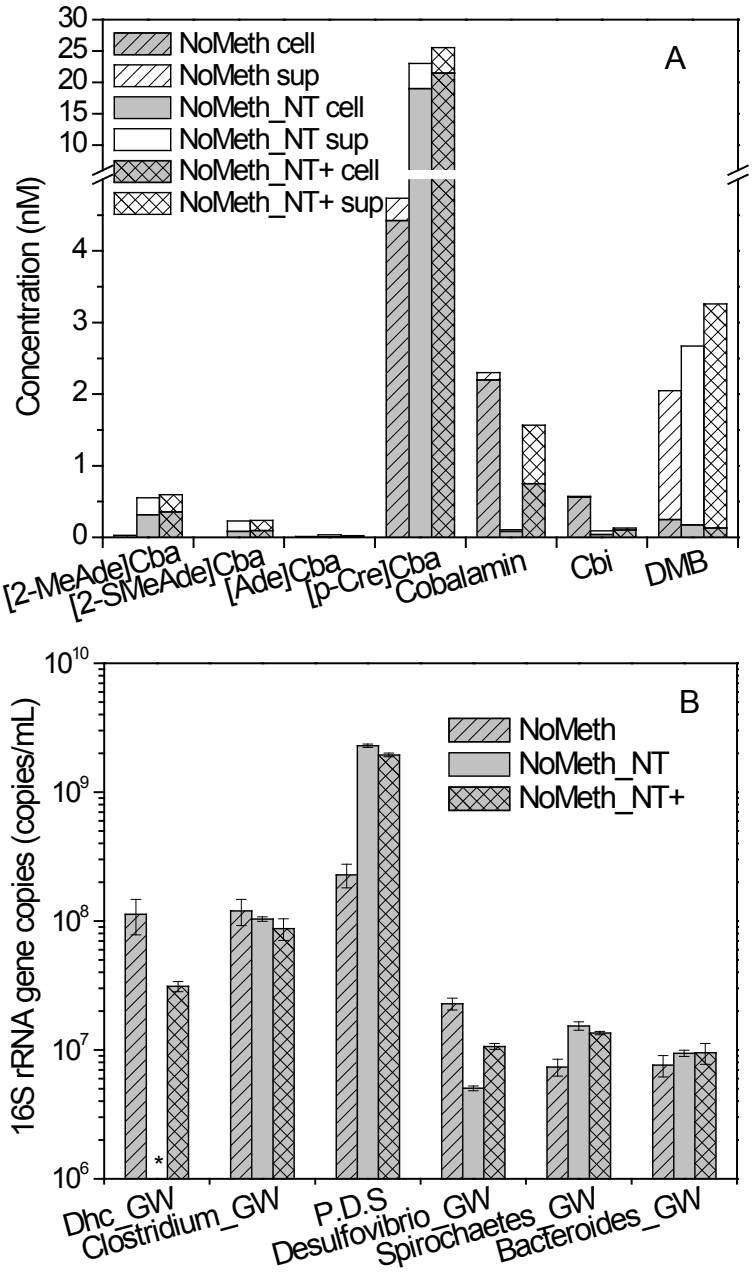


Figure 33. Comparison of corrinoid/lower ligand production and cell growth among NoMeth, NoMeth_NT and NoMeth_NT+ (A: Corrinoid/lower ligand production; B: numbers of OTUs, P.D.S represents *Pelosinus*_GW, *Dendrosporobacter*_GW and *Sporotalea*_GW (Men *et al*., submitted); *: < 2×10^3 copies/mL).

Overall, ANAS and groundwater enrichments have different corrinoid profiles. In ANAS sub-cultures with or without the addition of B_{12}, cobalamin was the only dominant corrinoid form, indicating that it was produced by supportive microorganisms, or even more likely by certain *Dehalococcoides* species, such as ANAS2 strain. While in groundwater enrichments, [p-Cre]Cba was produced in a large quantity, and was demonstrated to be the corrinoid produced by supportive microorganisms in those communities. *Dehalococcoides* in those communities were

able to modify other corrinoid forms by replacing the lower ligand base with the preferred DMB. DMB was produced in both ANAS and groundwater enrichments, even without the production of cobalamin, indicating that this compound might be commonly produced under anaerobic condition for other purposes rather than cobalamin biosynthesis. This study also provides a correlation between the presence of cobalamin and *Dehalococcoides*, which suggests that cobalamin could be used as an indicative and diagnostic biomarker for the bioremediation at a field site.

CONCLUSIONS AND IMPLICATION FOR FUTURE RESEARCH

In summary, we've been working on this project for 4 fiscal years since 2007 including a one-year non-cost extension. To date, we've successfully completed work on all of the four objectives and extended our work beyond the original scope.

We've applied 16S rRNA based molecular tools, such as clone libraries, quantitative PCR (qPCR), and PhyloChip, to different TCE dechlorinating communities, from laboratory enrichment cultures to environmental microcosms and TCE-contaminated field sites, from semi-batch reactor to a continuous chemostat. Methanogens were the most commonly identified Archaea, while Firmicutes (i.e. *Clostridium*), Proteobacteria (i.e. *Desulfovibrio*), Bacteroidetes (i.e. *Bacteroides*), and Spirochaetes were the common Bacterial phyla detected in those *Dehalococcoides*-containing communities. Most of the supporting organisms are fermentative bacteria capable of fermenting organic substrates, such as lactate and whey, to acetate and hydrogen, the carbon and energy sources for *Dehalococcoides*. The combination of PhyloChip analyses of DNA and RNA, together with clone library construction provided insights into the in situ microbial ecology and population dynamics at the TCE-contaminated field site undergoing biostimulation and bioaugmentation.

We've constructed more robust and stable defined consortium by growing *Dehalococcoides mccartyi* strain 195 with other commonly found bacteria in the communities, such as *Desulfovibrio, Acetobacterium, Syntrophomonas*, as well as Methanogens. *Desulfovibrio vulgaris* Hildenborough and *Syntrophomonas wofei* turned out to be more supportive for the growth of *Dehalococcoides* than other tested species. In the co-culture of strain 195 and *D. vulgaris* Hildenborough, the latter one might assist the dechlorination and growth of *Dehalococcoides* with enhanced availability to hydrogen, cobalamin corrinoid, as well as amino acids.

Besides the successful design and application of whole-genome microarray targeting strain 195 genome, we've further designed, validated and applied the genus-wide microarray targeting all four known *Dehalococcoides* genomes (strain CBDB1, BAV1, 195 and VS). The validation of this array using gDNA of strain 195 and BAV1 showed that it was able to detect genes and gene expression levels qualitatively and quantitatively. gDNA of ANAS, four groundwater enrichment cultures as well as 2 isolated strains (ANAS1 and ANAS2) were queried using this array. Although these cultures have different dechlorination capabilities, they all possessed a 195-like core genome. Genes within HPR/IE regions were missing, the percentage of which were different for different cultures. The tryptophan operon of strain 195

was missing in all of the tested cultures, instead, that of strain VS was detected. Horizontal gene transfer might be one reason for this observation. Compared with the whole genome, functional RDase genes were more reliable as a biomarker indicative of dechlorination capabilities.

Finally, we've investigated the correlation among certain biomarkers, growth phases, dechlorination performance, and corrinoid availability. Differential expression of genes with functions of translation, energy production, as well as amino acids metabolism reflect the nutrient level of dechlorination. In addition, the down-regulation of two putative riboswitches (DET0125 and DET0126) could correlate with an excess cobalamin condition, and another set of genes that contains an upstream putative cobalamin riboswitch [DET0657–0660/DET0691–0694 (cobT, cobS, cobC, and cobU)] could indicate an environment with unfavorable corrinoids. Other forms of corrinoids, such as [p-Cre]Cba, [5-OHBza]Cba, [5-MeBza]Cba, [2-MeAde]Cba, and [Ade]Cba were present in TCE-dechlorinating communities. Cobalamin is the preferred form for *Dehalococcoides*. In the groundwater enrichments, supportive microorganisms do not produce cobalamin directly, but other corrinoid forms (i.e. [p-Cre]Cba) instead. *Dehalococcoides* were able to uptake and modify those unfavorable corrinoid species to cobalamin with the presence of the right lower ligand base (DMB). The production of DMB was detected in two TCE-dechlorinating communities without exogenous vitamin B_{12}, even when the growth of *Dehalococcoides* and the production of cobalamin were eliminated.

Based on the current findings and established methods, we summarize the suggested biomarkers that could be applied in bioremediation processes with the developed methods in this project in Table 7. Future research includes applying FACS-WGA method along with genus-wide microarray to target *Dehalococcoides* species in newly constructed enrichments and chemostats; analyzing metagenomic and metatranscriptomic information of well-characterized communities to identify supportive functions. Moreover, we will also apply another promising approach, stable isotope probing (SIP) of nucleic acids to identify specific functional guilds and trophic interactions within syntrophic dechlorinating communities. Target communities include methanogenic and non-methanogenic cultures that were maintained in batch, semi-fed batch or continuous-chemostat conditions, using lactate as electron donor and TCE as electron acceptor. SIP can provide a fundamental understanding of syntrophic lifestyles governing anaerobic dechlorinating communities and can identify specific functional groups that support *Dehalococcoides* activity. These studies could provide us a guide in monitoring, diagnosing, and optimizing the *in situ* bioremediation of chlorinated solvents in the future.

Table 7. Summary of suggested Dhc community dechlorination biomarkers

Suggested biomarker	Advantage	Limitation	Skills required	Cost
RDase genes	Direct indication of dechlorination capabilities	Not providing other information than dechlorination	Moderate-High	Moderate
Hydrogenase (Hup, Fdh, Hym) genes	Diagnostic of hydrogen utilization	Indirect indicator	Moderate-High	Moderate
Corrinoid biosynthesis genes (*cobT*, *btuF*) and riboswitches	Diagnostic of corrinoid synthesis	Indirect indicator	Moderate-High	Moderate
16S rRNA of Dhc	Indication of dechlorination potential	Not well associated with dechlorination patterns	Moderate	Low
16S rRNA of *Desulfovibrio* spp.	Hydrogen producing microorganisms	Indirect and may be substrate-selected	Moderate	Low
16S rRNA of *Syntrophomonas* spp.	Hydrogen producing microorganisms	Indirect and may be substrate-selected	Moderate	Low
16S rRNA of *Pelosinus* spp.	Possible corrinoid producing bacteria	Indirect and may be substrate-selected	Moderate	Low
16S rRNA of *Clostridium* spp.	Possible hydrogen and corrinoid producing bacteria	Indirect and may be substrate-selected	Moderate	Low
Corrinoids and DMB	Indication of the need of external source of cobalamin	Sensitivity may be low for *in situ* monitoring, concentration step is required	Moderate	Low

LITERATURE CITED

1. **Ahmann, G. J., W. J. Chng, K. J. Henderson, T. L. Price-Troska, R. W. DeGoey, M. M. Timm, *et al.*** 2008. Effect of tissue shipping on plasma cell isolation, viability, and RNA integrity in the context of a centralized good laboratory practice - Certified tissue banking facility. *Cancer Epidemiol. Biomarkers Prev.* 17:666-673.
2. **Allen, R. H.,S. P. Stabler.** 2008. Identification and quantitation of cobalamin and cobalamin analogues in human feces. *Am. J. Clin. Nutr.* 87:1324-1335.
3. **Amann, R., W. Ludwig,K. Schleifer.** 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59:143-169.
4. **Amend, A. S., K. A. Seifert,T. D. Bruns.** 2010. Quantifying microbial communities with 454 pyrosequencing: does read abundance count? *Mol. Ecol.* 19:5555-5565.
5. **Banerjee, R.,S. W. Ragsdale.** 2003. The many faces of vitamin B_{12}: Catalysis by cobalamin-dependent enzymes. *Annual Review of Biochemistry* 72:209-247.
6. **Benjamini, Y.,Y. Hochberg.** 1995. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J. Roy. Stat. Soc. B* 57:289-300.
7. **Boga, H. I., R. Ji, W. Ludwig,A. Brune.** 2007. (lactate fermentation product) Sporotalea propionica gen. nov sp nov., a hydrogen-oxidizing, oxygen-reducing, propionigenic firmicute from the intestinal tract of a soil-feeding termite. *Arch. Microbiol.* 187:15-27.
8. **Brodie, E. L., T. Z. DeSantis, D. C. Joyner, S. M. Baek, J. T. Larsen, G. L. Andersen, *et al.*** 2006. Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. *Appl. Environ. Microbiol.* 72:6288-6298.
9. **Campbell, B. J.,S. C. Cary.** 2001. Characterization of a novel spirochete associated with the hydrothermal vent polychaete annelid, Alvinella pompejana. *Appl. Environ. Microbiol.* 67:110-117.
10. **Chen, M.,M. J. Wolin.** 1981. Influence of heme and vitamin B_{12} on growth and fermentations of *Bacteroides* Species. *J. Bacteriol.* 145:466-471.
11. **Chourey, K., M. R. Thompson, J. Morrell-Falvey, N. C. VerBerkmoes, S. D. Brown, M. Shah, *et al.*** 2006. Global molecular and morphological effects of 24-hour chromium(VI) exposure on Shewanella oneidensis MR-1. *Appl. Environ. Microbiol.* 72:6331-6344.
12. **Croft, M. T., A. D. Lawrence, E. Raux-Deery, M. J. Warren,A. G. Smith.** 2005. Algae acquire vitamin B-12 through a symbiotic relationship with bacteria. *Nature* 438:90-93.
13. **Daas, P. J. H., J. T. Keltjens, W. R. Hagen,C. Vanderdrift.** 1995. The Electrochemistry of 5-Hydroxybenzimidazolylcobamide. *Arch. Biochem. Biophys.* 319:244-249.
14. **DeSantis, T. Z., E. L. Brodie, J. P. Moberg, I. X. Zubieta, Y. M. Piceno,G. L. Andersen.** 2007. High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. *Microb. Ecol.* 53:371-383.
15. **Dick, G. J., A. F. Andersson, B. J. Baker, S. L. Simmons, A. P. Yelton,J. F. Banfield.** 2009. Community-wide analysis of microbial genome sequence signatures. *Genome Biol.* 10.
16. **Eng, J. K., A. L. McCormack,J. R. Yates.** 1994. An Approach to Correlate Tandem Mass-Spectral Data of Peptides with Amino-Acid-Sequences in a Protein Database. *J. Am. Soc. Mass Spectrom.* 5:976-989.
17. **Escalante-Semerena, J. C.** 2007. Conversion of cobinarnide into adenosylcobamide in bacteria and archaea. *J. Bacteriol.* 189:4555-4560.

18. **Gentleman, R. C., V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, *et al.*** 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 5.
19. **Guimaraes, D. H., A. Weber, I. Klaiber, B. Vogler,P. Renz.** 1994. Guanylcobamide and hypoxanthylcobamide - corrinoids formed by *Desulfovibrio vulgaris*. *Arch. Microbiol.* 162:272-276.
20. **He, J. Z., V. F. Holmes, P. K. H. Lee,L. Alvarez-Cohen.** 2007. Influence of vitamin B_{12} and cocultures on the growth of *Dehalococcoides* isolates in defined medium. *Appl. Environ. Microbiol.* 73:2847-2853.
21. **Holodniy, M., L. Rainen, S. Herman,B. Yen-Lieberman.** 2000. Stability of plasma human immunodeficiency virus load in VACUTAINER PPT plasma preparation tubes during overnight shipment. *J. Clin. Microbiol.* 38:323-326.
22. **Hongoh, Y., M. Ohkuma,T. Kudo.** 2003. Molecular analysis of bacterial microbiota in the gut of the termite Reticulitermes speratus (Isoptera; Rhinotermitidae). *FEMS Microbiol. Ecol.* 44:231-242.
23. **Johnson, D. R., E. L. Brodie, A. E. Hubbard, G. L. Andersen, S. H. Zinder,L. Alvarez-Cohen.** 2008. Temporal transcriptomic microarray analysis of "*Dehalococcoides ethenogenes*" strain 195 during the transition into stationary phase. *Appl. Environ. Microbiol.* 74:2864-2872.
24. **Johnson, D. R., A. Nemir, G. L. Andersen, S. H. Zinder,L. Alvarez-Cohen.** 2009. Transcriptomic microarray analysis of corrinoid responsive genes in *Dehalococcoides ethenogenes* strain 195. *Fems Microbiology Letters* 294:198-206.
25. **Kube, M., A. Beck, S. H. Zinder, H. Kuhl, R. Reinhardt,L. Adrian.** 2005. Genome sequence of the chlorinated compound respiring bacterium *Dehalococcoides* species strain CBDB1. *Nature Biotechnology* 23:1269-1273.
26. **Kunin, V., A. Copeland, A. Lapidus, K. Mavromatis,P. Hugenholtz.** 2008. A Bioinformatician's Guide to Metagenomics. *Microbiol. Mol. Biol. Rev.* 72:557-578.
27. **Lee, P. K. H., D. R. Johnson, V. F. Holmes, J. Z. He,L. Alvarez-Cohen.** 2006. Reductive dehalogenase gene expression as a biomarker for physiological activity of *Dehalococcoides* spp. *Appl. Environ. Microbiol.* 72:6161-6168.
28. **Macbeth, T. W., D. E. Cummings, S. Spring, L. M. Petzke,K. S. Sorenson.** 2004. Molecular characterization of a dechlorinating community resulting from in situ biostimulation in a trichloroethene-contaminated deep, fractured basalt aquifer and comparison to a derivative laboratory culture. *Appl. Environ. Microbiol.* 70:7329-7341.
29. **Martens, J. H., H. Barg, M. J. Warren,D. Jahn.** 2002. Microbial production of vitamin B-12. *Appl. Microbiol. Biotechnol.* 58:275-285.
30. **Mazumder, T. K., N. Nishio, S. Fukuzaki,S. Nagai.** 1987. *(B12)Production of Extracellular Vitamin-B12 Compounds from Methanol by Methanosarcina-Barkeri. *Appl. Microbiol. Biotechnol.* 26:511-516.
31. **Moe, W. M., R. E. Stebbing, J. U. Rao, K. S. Bowman, M. F. Nobre, M. S. da Costa, *et al.*** 2011. *Pelosinus defluvii* sp. nov., isolated from chlorinated solvent contaminated groundwater, emended description of the genus *Pelosinus*, and transfer of *Sporotalea propionica* to *Pelosinus propionicus* comb. nov. *IJSEM* 62:1369-1376.
32. **Morgan, J. L., A. E. Darling,J. A. Eisen.** 2010. Metagenomic Sequencing of an In Vitro-Simulated Microbial Community. *Plos One* 5.

33. **Ray, A. E., J. R. Bargar, V. Sivaswamy, A. C. Dohnalkova, Y. Fujita, B. M. Peyton, *et al.*** 2011. Evidence for multiple modes of uranium immobilization by an anaerobic bacterium. *Geochim. Cosmochim. Acta* 75:2684-2695.
34. **Renz, P.** 1999. Biosynthesis of the 5, 6- dimethylbenzimidazole moiety of cobalamin and of the other bases found in natural corrinoids. . *In: Banerjee, R (ed). Chemistry and biochemistry of B_{12}. John Wiley & Sons, Inc.: New York.*:pp 557-575.
35. **Reva, O. N.,B. Tummler.** 2005. Differentiation of regions with atypical oligonucleotide composition in bacterial genomes. *BMC Bioinf.* 6.
36. **Richardson, R. E., V. K. Bhupathiraju, D. L. Song, T. A. Goulet,L. Alvarez-Cohen.** 2002. Phylogenetic characterization of microbial communities that reductively dechlorinate TCE based upon a combination of molecular techniques. *Environ. Sci. Technol.* 36:2652-2662.
37. **Rodionov, D. A., A. G. Vitreschak, A. A. Mironov,M. S. Gelfand.** 2003. (B12_un) Comparative Genomics of the vitamin B-12 metabolism and regulation in prokaryotes. *J. Biol. Chem.* 278:41148-41159.
38. **Rowe, A. R., B. J. Lazar, R. M. Morris,R. E. Richardson.** 2008. Characterization of the community structure of a dechlorinating mixed culture and comparisons of gene expression in planktonic and biofloc-associated "*Dehalococcoides*" and *Methanospirillum* species. *Appl. Environ. Microbiol.* 74:6709-6719.
39. **Seshadri, R., L. Adrian, D. E. Fouts, J. A. Eisen, A. M. Phillippy, B. A. Methe, *et al.*** 2005. Genome sequence of the PCE-dechlorinating bacterium *Dehalococcoides ethenogenes*. *Science* 307:105-108.
40. **Shelobolina, E. S., K. P. Nevin, J. D. Blakeney-Hayward, C. V. Johnsen, T. W. Plaia, P. Krader, *et al.*** 2007. *Geobacter pickeringii* sp nov., *Geobacter argillaceus* sp nov and *Pelosinus fermentans* gen. nov., sp nov., isolated from subsurface kaolin lenses. *IJSEM* 57:126-135.
41. **Tabb, D. L., W. H. McDonald,J. R. Yates.** 2002. DTASelect and contrast: Tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* 1:21-26.
42. **Verberkmoes, N. C., A. L. Russell, M. Shah, A. Godzik, M. Rosenquist, J. Halfvarson, *et al.*** 2009. Shotgun metaproteomics of the human distal gut microbiota. *ISME J* 3:179-189.
43. **von Wintzingerode, F., U. B. Gobel,E. Stackebrandt.** 1997. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol. Rev.* 21:213-229.
44. **Walker, C. B., Z. L. He, Z. K. Yang, J. A. Ringbauer, Q. He, J. H. Zhou, *et al.*** 2009. The electron transfer system of syntrophically grown *Desulfovibrio vulgaris*. *J. Bacteriol.* 191:5793-5801.
45. **West, K. A., D. R. Johnson, P. Hu, T. Z. DeSantis, E. L. Brodie, P. K. H. Lee, *et al.*** 2008. Comparative genomics of "*Dehalococcoides ethenogenes*" 195 and an enrichment culture containing unsequenced "*Dehalococcoides*" strains. *Appl. Environ. Microbiol.* 74:3533-3540.
46. **Yang, Y.,J. Zeyer.** 2003. Specific detection of *Dehalococcoides* species by fluorescence in situ hybridization with 16S rRNA-targeted oligonucleotide probes. *Appl. Environ. Microbiol.* 69:2879-2883.

47. **Yang, Y. R.,P. L. McCarty.** 1998. Competition for hydrogen within a chlorinated solvent dehalogenating anaerobic mixed culture. *Environ. Sci. Technol.* 32:3591-3597.
48. **Zhang, Z., S. Schwartz, L. Wagner,W. Miller.** 2000. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7:203-214.

APPENDIX

Peer reviewed papers (published and submitted):

Men, Y., P. K. H. Lee, K. C. Harding, L. Alvarez-Cohen. Characterization of four TCE-dechlorinating microbial enrichments grown with different cobalamin stress and methanogenic conditions. 2011. Submitted.

Brisson, V. L., K. A. West, P. K. H. Lee, S. G. Tringe, E. L. Brodie, and L. Alvarez-Cohen. 2012. Metagenomic Analysis of a Stable Trichloroethene Degrading Microbial Community. *ISME J* accepted.

Lee, P. K. H., B. D. Dill, T. S. Louie, M. Shah, N. C. Verberkmoes, G. L. Andersen, S. H. Zinder, and L. Alvarez-Cohen. 2012. Global Transcriptomic and Proteomic Responses of *Dehalococcoides ethenogenes* Strain 195 to Fixed Nitrogen Limitation. *Appl. Environ. Microbiol.* PMID: 22179257.

Lee, P. K. H., F. Warnecke, E. L. Brodie, T. W. Macbeth, M. E. Conrad, G. L. Andersen, and L. Alvarez-Cohen. 2012. Phylogenetic microarray analysis of a microbial community performing reductive dechlorination at a TCE-contaminated site. *Environ. Sci. Technol.* **46**: 1044-1054.

Men, Y. H. Feil, N. C. Verberkmoes, M. B. Shah, D. R. Johnson, P. K. H. Lee, K. A. West, S. H. Zinder, G. L. Andersen, and L. Alvarez-Cohen. 2011. Sustainable syntrophic growth of *Dehalococcoides ethenogenes* strain 195 with *Desulfovibrio vulgaris* Hildenborough and *Methanobacterium congolense*: global transcriptomic and proteomic analyses. *ISME J* **6**: 410-421.

Zhuang, W. Q., S. Yi, X. Feng, S. H. Zinder, Y. J. Tang, L. Alvarez-Cohen. 2011. Selective utilization of exogenous amino acids by *Dehalococcoides ethenogenes* strain 195 and its effects on growth and dechlorination activity. *Appl. Environ. Microbiol.* **77**: 7797-7803.

Lee, P. K. H., D. Cheng, P. Hu, K. A. West, G. J. Dick, E. L. Brodie, G. L. Andersen, S. H. Zinder, J. He, and L. Alvarez-Cohen. 2011. Comparative genomics of two newly isolated *Dehalococcoides* strains and an enrichment using a genus microarray. *ISME J* **5**: 1014-1024.

Lee, P. K. H., J. Z. He, S. H. Zinder, and L. Alvarez-Cohen. 2009. "Evidence for Nitrogen fixation by "*Dehalococcoides ethenogenes"* strain 195", *Appl. Environ. Microbiol.* **75** (23):7551-7555.

Tang, Y.J., S. Yi, W.Q. Zhuang, S. H. Zinder, J.D. Keasling, and L. Alvarez-Cohen, 2009. "Investigation of Carbon Metabolism in "*Dehalococcoides ethenogenes*" strain 195 Using Isotopomer and Transcriptomic Analysis", *J. Bacteriol.* **191** (16): 5224-5231. PMID: 19525347

Johnson, D. R., A. Nemir, **G. L. Andersen, S. H. Zinder and L. Alvarez-Cohen.** 2009. "Transcriptomic Microarray Analysis of Corrinoid Responsive Genes in *Dehalococcoides ethenogenes* strain 195", *FEMS Microbial Letters*. **294:** 198-206.

West, K. A., D. R. Johnson, P. Hu, T. Z. DeSantis, E. L. Brodie, P. K. H. Lee, H. Feil, G. L. Andersen, S. H. Zinder, and Lisa Alvarez-Cohen. 2008. Comparative genomics of "*Dehalococcoides ethenogenes*" 195 and an enrichment culture containing unsequenced "*Dehalococcoides*" strains. *Appl. Environ. Microbiol*. **74:** 3533-3540.

Johnson, D. R., E. L. Brodie, A. E. Hubbard, G. L. Andersen, S. H. Zinder, and L. Alvarez-Cohen. 2008. Temporal transcriptomic microarray analysis of "*Dehalococcoides ethenogenes*" strain 195 during the transition into stationary phase. *Appl. Environ. Microbiol*. **74:** 2864-2872.

Lee, P. K. H., T. W. Macbeth, K. S. Sorenson, Jr., R. A. Deeb, and L. Alvarez-Cohen. 2008. Quantifying genes and transcripts to assess the in situ physiology of "*Dehalococcoides*" spp. in a trichloroethene-contaminated groundwater site. *Appl. Environ. Microbiol.,* **74:** 2728-2739.

He, J, V. F. Holmes, P. K. H. Lee and L. Alvarez-Cohen. 2007. "Influence of Vitamin B_{12} and Co-cultures on the Growth of *Dehalococcoides* Isolates in Defined Medium", *Appl. Environ. Microbiol.*, **73** (9): 2847-2853.

Selected Technical Abstracts:

Comparison of Microarray and Metagenomic Sequencing for Detecting *Dehalococcoides* Genes **Mao, X.W., Men, Y., K. and L. Alvarez-Cohen,** 2010. "Development of a chemostat reactor to study TCE dechlorination by Dehalococcoides", Partners in Environmental Technology Symposium and Workshop, Washington DC.

Men, Y., K. Harding, Y. Shan, W. Zhuang, G. L. Andersen, S. H. Zinder, M. E. Taga, Lisa Alvarez-Cohen. 2010. **"**Identification of corrinoid-providing supportive microorganisms for *Dehalococcoides* in TCE-dechlorinating enrichment cultures by analytical and molecular tools". 110th general meeting of American Society for Microbiology, San Diego, CA.

West, K. A., P. K. H. Lee, P. Hu, S. H. Zinder, G. Andersen, L. Alvarez-Cohen. 2010. "Genome and Transcriptome analyses of unsequenced Dehalococcoides strains in a dechlorinating enrichment culture using a genus-wide Microarray". 110th general meeting of American Society for Microbiology, San Diego, CA.

Lee, P. K. H., D. Cheng, P. Hu, K. A. West, E. L. Brodie, G. L. Andersen, S. H. Zinder, J. He, and Lisa Alvarez-Cohen. 2010. "Querying the genomes of unsequenced Dehalococcoides strains via a
genus-wide microarray". 13th International Symposium on Microbial Ecology, Seattle, WA.

Yi, S., W. Q. Zhuang, X. Feng, S. H. Zinder, Y. J. Tang and L. Alvarez-Cohen, 2010. "Exogenous Amino Acid Utilization by *Dehalococcoides ethenogenes* strain 195", 110th general meeting of American Society for Microbiology, San Diego, CA.

Men, Y., K. Harding, Y. Shan, W. Zhuang, H. Feil, G. L. Andersen, S. H. Zinder, J. He, and L. Alvarez-Cohen. 2009. "Application of molecular and analytic tools to track enrichment of reductive dechlorination cultures from a TCE contaminated groundwater site". Partners in Environmental Technology Symposium and Workshop, Washington DC.

Men, Y., H. Feil, S. H. Zinder, G. L. Andersen, and L. Alvarez-Cohen. 2008. "Differential expression by *Dehalococcoides ethenogenes* strain 195 growing syntrophically with *Desulfovibrio vulgaris* Hildenborough and in tri-culture with *D. vulgaris* Hildenborough and *Methanobacterium congolense*" Partners in Environmental Technology Symposium and Workshop, Washington DC.

Johnson, D. R., N. C. VerBerkmoes, S. H. Zinder and L. Alvarez-Cohen, 2008. "Absolute comparison of transcriptomic and proteomic data using normalized spectrum counts for protein abundance estimates" 108th general meeting of American Society for Microbiology, Boston, MA.

Lee, P. K. H., E. L. Brodie, G. L. Andersen, T. W. Macbeth, K. S. Sorenson, R.L. Deeb and L. Alvarez-Cohen, 2008. "From Individuals to Community: Molecular Tools Shed Lights into the Environmentally Relevant Organisms". Battelle Annual Conference on Remediation of Chlorinated Compounds, Monterey, CA. Invited Speaker.

West, K. A., D. R. Johnson, E. L. Brodie, H. Feil, G. L. Andersen, S. H. Zinder and L. Alvarez-Cohen, 2008. "Molecular Tools Provide New Genomic and Transcriptomic Insights into Dehalococcoides spp". Battelle Annual Conference on Remediation of Chlorinated Compounds, Monterey, CA. Invited Speaker.