

# **Chasing Organohalide Respirers:**

## **Ecogenomics Approaches to Assess the Bioremediation Capacity of Soils**

**Farai Maphosa**

## **Thesis Committee**

### **Thesis supervisor**

Prof. Dr. W. M. de Vos  
Professor of Microbiology  
Wageningen University

### **Thesis Co-supervisor**

Dr. H. Smidt  
Associate professor, Laboratory of Microbiology  
Wageningen University

### **Other Members**

Prof. Dr. Ir. Huub H.M. Rijnaarts  
Wageningen University

Prof. Dr. Dirk Springael  
Catholic University of Leuven, Belgium

Prof. Dr. Lorenz Adrian  
Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany

Dr. Inez Dinkla  
Bioclear B.V., Groningen, The Netherlands

This research was conducted under the auspices of the Graduate School  
SENSE (Socio-Economic and Natural Sciences of the Environment)

# Chasing Organohalide Respirers:

## Ecogenomics Approaches to Assess the Bioremediation Capacity of Soils

Farai Maphosa

### **Thesis**

submitted in fulfillment of the requirements for the degree of doctor  
at Wageningen University

by the authority of the Rector Magnificus

Prof. dr. M.J. Kropff,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Monday 31 May 2010

at 1:30 p.m. in the Aula

Farai Maphosa

Chasing Organohalide Respirers: Ecogenomics Approaches to Assess the  
Bioremediation Capacity of Soils

PHD Thesis Wageningen University, Wageningen, The Netherlands 2010

170 pp.

With references and summaries in English and Dutch

ISBN : 978-90-8585-656-6

...dedicated to my lovely wife Busi and our sons, Anotida and Hillel



## Abstract

Organohalide respiring bacteria (OHRB) are efficient degraders of chlorinated ethenes, chlorophenols, and other halogenated aliphatic and aromatic hydrocarbons. Nevertheless, these organohalides appear to persist at various locations, and this lack of degradation can be attributed to the absence of OHRB in sufficient numbers or improper physico-chemical conditions for their growth and activity. Hence, there is an urgent need for fast, robust and sensitive methods that allow for predicting and monitoring the bioremediation potential and activity of OHRB. The ecogenomic toolbox that capitalizes on genomics and other high throughput approaches offers completely new avenues to study OHRB in natural environments. To this end, we studied the genomic repertoire, as well as functional gene expression patterns, in two representative Gram-positive OHRB. Firstly, we elucidated the genome of *Dehalobacter* sp. strain E1, which dehalogenates hexachlorocyclohexene in co-culture with a *Sedimentibacter* sp., and identified 10 putative reductive dehalogenase (Rdh)-encoding gene clusters, including the *pceABCT* operon previously described in *D. restrictus*. This suggests that strain E1 has a greater organohalide respiration potential than previously observed and we confirmed this experimentally by showing dechlorination of trichloroethene by the co-culture. Secondly, *Desulfitobacterium hafniense* DCB-2, capable of dehalogenating different chloroaromatic compounds, has in its genome seven chlorophenol Rdh (*cprA*) -like genes, five of which (*cprA1-cprA5*) are predicted to encode functional proteins. Different specificities and transcriptional patterns in response to meta- and ortho- chlorophenols were observed for these five genes using as substrates 2,3-dichlorophenol (2,3-DCP) for the meta-dechlorination and 2,4-DCP for the ortho-dechlorination positions suggesting distinct control at the transcriptional level. Using molecular monitoring and model simulations we determined in-situ degradation performance of an on-site dechlorinating bioreactor and its influence on a chloroethene contamination plume. *Dehalococcoides* was the dominant OHRB as revealed by quantitative PCR (qPCR) targeting 16S rRNA- and chloroethene Rdh genes (*tceA*, *vcrA*, *bvcA*). Saturation analysis showed that while performance of the bioreactor improved due to especially the addition of molasses, OHRB were fast-growing. Lastly we developed an assay that combines multiplex capabilities and specificity of ligation based circularization probes with high-throughput qPCR in the Biotrove OpenArray™, resulting in a flexible, quantitative multiplex detection system for OHRB in contaminated soils. Moreover, an outlook is provided on how the ecogenomics approaches described here and new tools such as proteomics and metabolomics currently being developed can be applied to advance both the understanding and the exploitation of soil bioremediation by OHRB.

**Key Words:** Bioremediation, organohalide respiration, *Dehalobacter*, *Desulfitobacterium*, *Dehalococcoides*, *Sedimentibacter*; reductive dehalogenase, quantitative PCR, chloroethenes, chlorophenols, ligation probes





# CONTENTS

	Aim and Outline	6
Chapter 1	General Introduction : Exploiting the Ecogenomics Toolbox for Environmental Diagnostics of Organohalide Respiring Bacteria	15
Chapter 2	Advanced Insight in a Hexachlorocyclohexane-Dechlorinating Coculture of <i>Dehalobacter</i> sp. and <i>Sedimentibacter</i> sp. by Metagenome Analysis	33
Chapter 3	Differential Expression of the <i>Desulfitobacterium hafniense</i> DCB-2 Chlorophenol Reductive Dehalogenase Genes in Response to Meta- or Ortho-Chlorophenols	61
Chapter 4	Microbial Community- and Metabolite Dynamics of an Anoxic Dechlorinating Bioreactor	77
Chapter 5	Chasing Organohalide Respirers: Quantitative Multiplex Detection of Organohalide Respiring Bacteria using a Ligation Probe-Based System Coupled to OpenArrays	101
Chapter 6	General Discussion and Future Perspectives	125
	References	143
	Summary	159
	Nederlandse Samenvatting	163
	About the Author	168
	Acknowledgements	170
	Overview of completed training activities	173



# Aim and Outline

## Aim and Outline

The living components of soil, including complex microbial communities, play an essential role in so-called life-support functions that sustain the biological activity of ecosystems. These functions allow the soils to maintain their ecological integrity such as by the recycling of nutrients and inactivation of pollutants. Although the importance of this latter self-attenuating capacity has been recognised, the underlying biological and biochemical processes are still neither fully understood nor exploited to the full extent.

The aim of the research described in this thesis was to unveil and explore the genetic potential related to the self-attenuating capabilities of soils with organohalide contamination using an integrated ecogenomics approach. Organohalide degradation was used as a model system to study the attenuation capabilities because ecophysiological and molecular ecological studies of the past two decades have indicated that a variety of anaerobic bacteria are able to efficiently degrade organohalides, including organohalide-respiring bacteria (OHRB). Nevertheless, organohalides such as chlorinated ethenes, chlorophenols, and other halogenated aliphatic and aromatic hydrocarbons appear to persist in deeper soil subsurfaces at various locations. The reason for this obvious lack of degradation can be the absence of the specific ORHB in sufficient numbers, or improper physico-chemical conditions for their growth and activity. To advance the application of OHRB, there is an urgent need for fast, robust and sensitive methods for predicting and monitoring bioremediation capacity and activity in-situ.

The genomic potential, taxonomic diversity, organohalide degradation activity and gene expression were explored in OHRB growing in laboratory pure and enrichment cultures, contaminated soils and groundwater. Furthermore, we describe the development of a high throughput monitoring tool, using multiplex Biotrove OpenArray<sup>TM</sup> technology, to assess the bioremediation capacities for organohalide respiration, enabling the comprehensive and simultaneous measurement of multiple variables.

**Chapter 1** introduces the present ecogenomics approaches, and defines how the ability to probe the environment at the molecular level with a range of 'omics' methods has created an important new paradigm in bioremediation design and management. The potential of ecogenomics approaches in developing high-throughput-techniques for detecting and monitoring OHRB, and providing improvements to selection, specificity and sensitivity of target biomarkers, is discussed, including the application of such approaches to evaluate bioremediation strategies.

Genomic studies have expanded our knowledge of genetic diversity among cultured isolates from different species. Moreover, they have been instrumental in our understanding of the metabolic diversity of OHRB, providing clues about the adaptation of these bacteria

to organohalide respiration. Presently, the complete genome sequences of eight OHRB have been determined. However, no genomic information is yet available for *Dehalobacter* spp., a widely spread member of Gram-positive OHRB. To this end, **Chapter 2** describes the first insights into the genome sequence of the b-HCH dechlorinating *Dehalobacter* sp. strain E1 that grows in strict coculture with *Sedimentibacter* sp. strain B4. Metagenome sequence analysis of the consortium followed by advanced bioinformatic analysis indicated that strain E1 (2.6Mbp) has the capacity to utilize a broader range of organohalides than previously observed. Its strict dependence on strain B4 (4.2 Mbp) can be attributed to its own limited metabolism, making it dependent on the wide metabolic capability of *Sedimentibacter* sp. that includes the catabolism of a large spectrum of carbohydrates and the biosynthesis of amino acids and vitamins.

**Chapter 3** describes the transcriptional analysis for 5 reductive dehalogenase genes (termed *cprA1* to *cprA5*) in *Desulfitobacterium hafniense* DCB-2 cells cultured in the presence of 2,3-dichlorophenol and 2,4-dichlorophenol as model substrates for the meta- and ortho- dechlorination positions. Specific expression of these *cprA* genes was observed during growth on different halogenated substrates, indicating their control by distinct regulatory systems.

Beyond the laboratory study of OHRB, their in-situ monitoring and quantification is essential for optimization of in-situ bioremediation of anoxic subsurface sites contaminated with organohalogenes. To this end molecular monitoring and mass balance model simulations were applied to determine the chloroethene-degradation performance of an in-situ dechlorinating bioreactor and its influence on a contamination plume (**Chapter 4**). The combination of molecular diagnostics with mass-balancing and kinetic modeling allowed for detailed insight into the complex interplay of OHRB and metabolite dynamics in the bioreactor, giving crucial information for optimizing bioremediation.

**Chapter 5** describes the development of a ligation-based, quantitative detection assay for the high-throughput characterization of the dehalogenating capacity of soils in contaminated environments. This probe ligation-based assay uses a real-time PCR methodology combined with OpenArray™ technology, enabling accurate target quantification in a highly multiplex format. Its extensive validation and dynamic range is described as well as its application to monitor the bioremediation of contaminated soils.

Finally **Chapter 6** provides a summary and general discussion of the results presented in this thesis. It also provides a perspective on the impact of ecogenomics approaches to improve and expand our knowledge of biomarkers that can be effectively used for monitoring and optimizing bioremediation strategies at sites with single and multiple contaminations.



# CHAPTER 1

## General Introduction

### **Exploiting the Ecogenomics Toolbox for Environmental Diagnostics of Organohalide Respiring Bacteria**

Farai Maphosa, Willem M. de Vos, and Hauke Smidt

*A modified version of this chapter, with parts of chapter 6, is in press in Trends in Biotechnology*

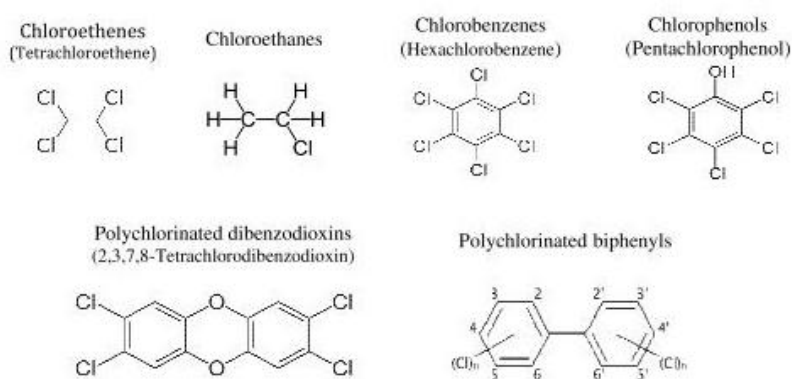
## **ABSTRACT**

The ability to probe the environment at the molecular level with a range of 'omics' methods has created an important new paradigm in bioremediation design and management. Ecogenomics, the application of genomics to ecological and environmental sciences, defines phylogenetic and functional biodiversity at the DNA, RNA and protein levels, to elucidate functions and interactions of organisms in relation to ecological and evolutionary processes. Effective bioremediation of wide-spread halo-organic pollutants in anaerobic environments requires knowledge of catabolic potential and in-situ dynamics of organohalide-respiring and co-metabolizing microorganisms. Here, we discuss the potential of ecogenomics approaches in developing high-throughput-techniques for detection and monitoring organohalide-respirers, and for providing improvements to selection, specificity and sensitivity of target biomarkers, including their application to evaluate bioremediation strategies.



## Introduction

Bioremediation technologies are increasingly being applied in the clean-up of organohalide contaminants, such as solvents, polychlorinated biphenyls (PCBs) and polybrominated diphenyl ethers (PBDEs) (Fig. 1) [1-5]. Most contaminated sites, including aquatic sediments, submerged soils and groundwater aquifers, are oxygen-depleted, thus anaerobic bacteria capable of organohalide respiration are promising candidates for bioremediation in these environments. The incorporation of genomic tools now allows to further open the ‘black box’ of processes underlying bioremediation, such as organohalide respiration, and to better understand mechanisms of microbial adaptations to toxic organohalide compounds and other relevant environmental factors [6].



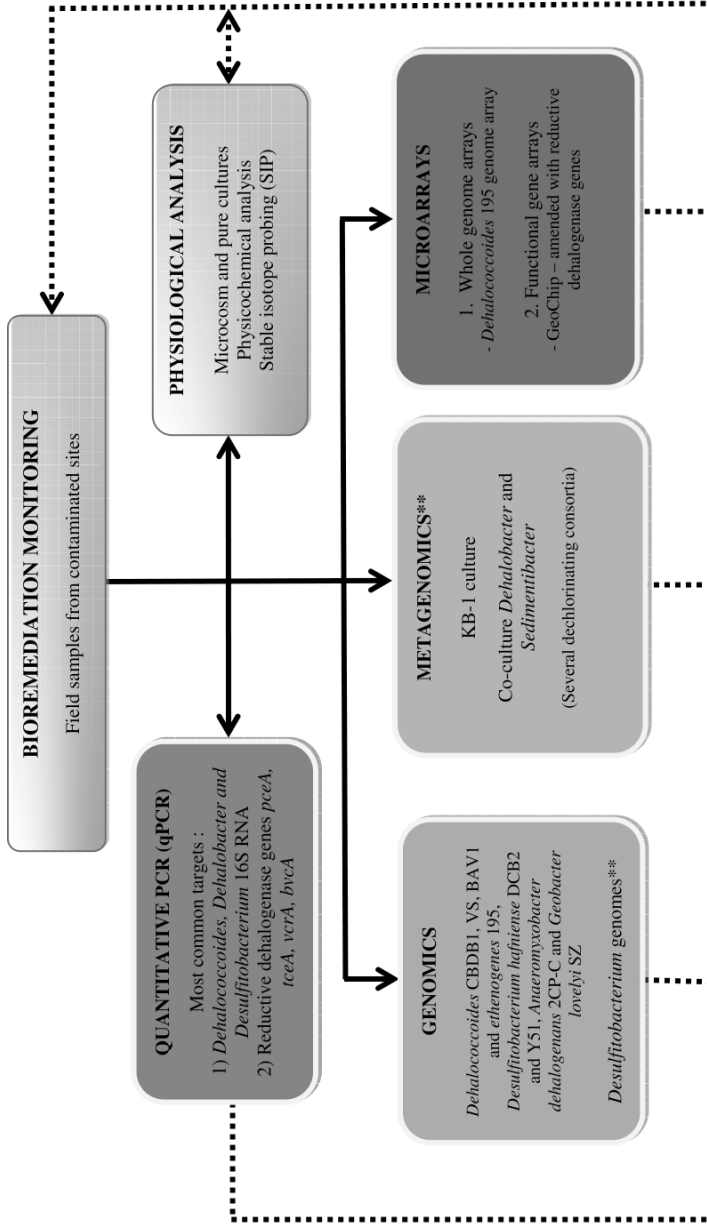
**Fig. 1.** Structures and examples of typical organohalide substrates.

Ecogenomics, the application of genomics to answer environmental questions, is a relatively new field. Ecogenomics defines phylogenetic and functional biodiversity at the level of DNA, RNA, and proteins [7-10], using this knowledge to quantify microorganisms, and to elucidate their functions and interactions within an ecosystem in relation to ecological and evolutionary processes [11,12]. Ecogenomics seeks to understand how genes determining community activity and function respond to environmental stimuli, such as organohalide pollutants. Ecogenomics has capitalized on recent advances in high-throughput and -output technologies to address complex problems at a system level. These genomics technologies are

considerably more unbiased, sensitive and responsive compared to traditional methods such as culturing, enabling to look beyond traditional biological endpoints of microbial growth. The available ecogenomics toolbox has emerged as an important means in the detection and monitoring of organohalide-respiring bacteria (OHRB) - dedicated degraders of halogenated aliphatic and aromatic pollutants in anoxic ecosystems. This toolbox comprises microbial techniques that yield information on: (i) DNA sequence diversity; (ii) expression of genes; and (iii) linking genes to organohalide respiration activity (Fig. 2). A detailed description of all existing utensils of the ecogenomics toolbox is beyond the scope of this review, and excellent overviews can be found elsewhere [8,9,13]. Here, we discuss the development of the ecogenomics toolbox and its application for improving the exploitation of OHRB for in situ bioremediation.

### **Challenges in the search for organohalide respiring bacteria**

Ever since the first discovery of a bacterial species capable of coupling the reductive dehalogenation of 3-chlorobenzoate to energy conservation and microbial growth [14], the search for OHRB as potential bioremediation candidates has resulted in isolates, which belong to different genera, including *Desulfitobacterium*, *Sulfurospirillum*, *Anaeromyxobacter*, *Desulfomonile*, *Desulfovibrio*, *Desulfuromonas*, *Geobacter*, *Dehalogenimonas*, *Dehalococcoides*, and *Dehalobacter* [15-18]. In contrast to other microorganisms, in which dehalogenation is fortuitously catalyzed in an unspecific fashion at metal-ion-containing cofactors and corresponding enzymes, reductive dehalogenation by OHRB is catalyzed by dedicated enzyme systems linked to an anaerobic respiratory chain [15]. It should be noted that there is no direct correlation of phylogeny and substrate specificity of isolates. Chloroethene- and haloaryl-degrading strains could be isolated from all phylogenetic branches mentioned above. In contrast, dehalogenation capacities of the isolates appear to be strain dependent. Isolates belonging to the genera *Anaeromyxobacter*, *Desulfitobacterium*, *Sulfurospirillum*, *Desulfomonile*, *Desulfuromonas*, *Desulfovibrio*, and *Geobacter* are metabolically versatile with respect to their spectrum of electron donors and acceptors and are able to dehalogenate a wide range of halogenated aromatic and aliphatic compounds. *Dehalococcoides* and *Dehalobacter* spp. isolates appear



**Fig. 2.** The ‘ecogenomics toolbox’: current set of techniques in use (i.e. genomics, microarrays, metagenomics) in monitoring bioremediation. \*\*Denotes genomes and metagenomes expected in 2010.

as highly specialized bacteria that strictly depend on organohalide respiration for growth, in most cases coupled to hydrogen as the sole electron donor.

The fact that many OHRB are notoriously difficult to culture and do not grow to high density poses a great challenge for their detailed physiological and genetic characterization. For example, *Dehalococcoides* spp. are difficult to maintain in pure culture, but are more easily grown in a microbial community on which they depend for H<sub>2</sub> supply [16,19]. Similarly, several *Dehalobacter* spp. have now been reported to grow only in strict co-culture or consortia [20-22], raising intriguing questions about the possible role of syntrophic interactions that could act as drivers for dehalogenation. The whole spectrum of Ecogenomics techniques will be useful in shedding light to the lifestyles of OHRB.

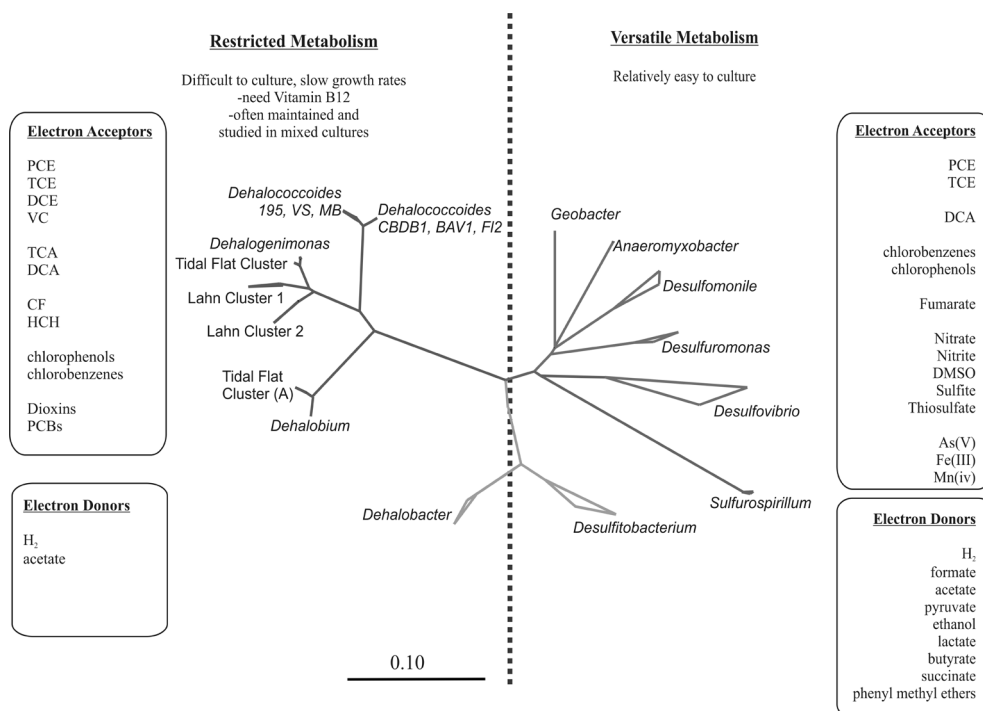
### **Genomic basis of organohalide respiration**

Genomic studies have further expanded our knowledge of genetic diversity among cultured isolates from different species, building on the wealth of findings obtained through cultivation and physiological characterization [3,7,23,24] (Fig. 3). The genomes of eight OHRB have now been fully sequenced (Table 1), and this number is expected to at least double in the next year (GOLD Genomes Online Database; <http://www.genomesonline.org>). Consequently, there is now an unprecedented number of gene homologues that have been retrieved from these sequenced genomes in addition to those obtained from environmental samples, which are available for future studies aimed at identifying and extrapolating gene functions in dehalogenation processes occurring at bioremediation sites [10,23,25]. Physiological studies to unravel the functions of these multiple gene homologues can now be developed and conducted in systems biology approaches involving proteomics and metabolomics.

#### *Versatile vs. restricted metabolism of different OHRB*

Genomics has been instrumental to understanding the differences in the metabolism of OHRBs, giving us clues on the adaptation and dedication of these bacteria to organohalide respiration. The larger genomes of *Anaeromyxobacter dehalogenans* [23] and of two strains of *Desulfitobacterium hafniense* [24] confirm the versatile metabolism of these bacteria as compared to “*Dehalococcoides*” spp., which have

smaller genomes and are restricted to organohalide respiration (Fig. 3). The *A. dehalogenans* 2CP-C genome has up to 68 putative c-type cytochrome genes [23], most likely enabling this organism to occupy environments with variable redox conditions.



**Fig. 3.** Phylogenetic tree of OHRB based on bacterial 16S rRNA sequences. Alignment and phylogenetic analysis were performed with the ARB software and tree was constructed using the neighbor-joining (NJ) method [26]. The reference bar at the bottom center indicates the branch length that represents 10% sequence divergence. Electron donors and acceptors are listed in the flanking text boxes. Halogenated electron acceptors listed here are restricted to chlorinated compounds, however, there are many other organohalide compounds that these OHRBs can respire. Color key: Chloroflexi (red), Deltaproteobacteria (blue), Epsilonproteobacteria (purple), Firmicutes (green). Abbreviations: DCA, dichloroethane; DCE, dichloroethene; DMSO, Dimethyl sulfoxide; HCH, hexachlorocyclohexane; PCB, polychlorinated biphenyls; PCE, tetrachloroethene; TCE, trichloroethene; TCA, trichloroethane; VC, vinyl chloride; CF, chloroform.

Table 1: Whole genome sequence statistics of OHRB and key genes involved in organohalide respiration processes<sup>1</sup>

Genome Name	Bases (Mbp)	GC (%)	CDs	RdhA genes	Formate dehydrogenases	Hup-type hydrogenases	Other Ni-Fe-type hydrogenases
<i>Anaeromyxobacter dehalogenans</i> 2CP-C	5.01	75	4361	2	8	5	2
<i>Geobacter lovleyi</i> SZ	3.87	55	3476	2	2	1	2
<i>Desulfotobacterium hafniense</i> DCB-2	5.28	48	4712	7	5	2	3
<i>Desulfotobacterium hafniense</i> Y51	5.73	47	5060	4	2	3	1
<i>Dehalococcoides ethenogenes</i> strain 195	1.47	49	1590	17	1	1	4
<i>Dehalococcoides</i> sp. CBDB1	1.4	47	1385	32	1	1	4
<i>Dehalococcoides</i> sp. BAV1	1.34	47.2	1371	11	1	1	4
<i>Dehalococcoides</i> sp. VS	1.41	47.3	2096	36	1	1	4

<sup>1</sup> Data taken from IMG Database [76], as of September 2009

Abbreviations: Genes, total gene count; CDS, coding sequences; RdhA, reductive dehalogenase catalytic subunit, Hup, hydrogen uptake hydrogenase.

In contrast, *Dehalococcoides* spp., which are lacking c-type cytochromes, are restricted to anaerobic zones where reductive dechlorination is a feasible terminal electron accepting process [27].

Hydrogenases, which are involved in the transfer of electrons from electron donors such as  $H_2$  in many OHRB [15,18], have also been identified in the *A. dehalogenans* strain 2CP-C genome. The two different predicted Ni-Fe-type hydrogenases might confer respiratory versatility under distinct environmental conditions such as high versus low  $H_2$  partial pressures. One of the two Ni-Fe hydrogenases encoded in the *A. dehalogenans* genome is most similar to the putative F420-reducing cytoplasmic hydrogenase VhuA [28] encoded in *Dehalococcoides* genomes [23]. *Dehalococcoides* also carry a large number of genes (over 40 in strain 195) potentially devoted to  $H_2$  interconversions with protons and electrons, including those encoding five different multi-subunit hydrogenases (four [Ni-Fe]-hydrogenases and one [Fe]-hydrogenase) [7,10]. These are typically involved in energy transduction, highlighting the importance of this reaction to their lifestyle.

Unlike in *Dehalococcoides* and *A. dehalogenans*, the use of hydrogen as an electron donor has not been proven to support growth in *D. hafniense* strains DCB-2 and Y51, even though their genomes encode Hup-type Ni-Fe periplasmic hydrogenases and Ni-Fe cytoplasmic hydrogenases (Table 1). Furthermore, hydrogenotrophic growth has been shown for the closely related *D. hafniense* strains TCE-1, TCPA, DP7 and G2, as well as *D. dehalogenans*, *D. chlororespirans* and *D. dichloroeliminans* [25]. Knowledge of the encoded diversity of hydrogenases in different OHRB can, in future studies, facilitate the elucidation of kinetic aspects of hydrogen metabolism. Such information needs to be more closely considered in order to optimize the growth of OHRB and regulate the rates of dehalogenation in contaminated environments where different hydrogen concentrations may stimulate or inhibit growth of OHRB.

### Reductive dehalogenases

The key enzymes in organohalide respiration, reductive dehalogenases (Rdh), consist of a catalytic subunit RdhA and a small putative membrane anchor, RdhB [15,18]. Several Rdhs have been purified and characterized. With few exceptions the

catalytic subunit contains a corrinoid and two iron-sulphur centres [18,29,30] and has been localized at the outer side of the cytoplasmic membrane [18,25]. From the 8 sequenced OHRB genomes, 111 *rdhA*-like genes have been annotated (Table 1), and a large number of additional gene sequences have been retrieved from other isolates and mixed cultures, as well as from pristine and polluted environmental samples.

Phylogenetic analysis shows various clusters of Rdhs with similar functions (Fig. 4), like chlorophenol reductases (CprA), trichloroethene reductases (TceA) and vinyl chloride reductases (VcrA, BvcA). However, RdhAs with an identical function, e.g. tetrachloroethene reductases (PceA) are also present in different phylogenetic branches, suggesting a convergent evolution in different phylogenetic lineages of OHRBs. Generally, *Dehalococcoides* RdhAs form a separate clade compared to those from other OHRBs with a few exceptions like CbrA of strain CBDB1 [3]. Observed clades and clusters may be used in further characterization of representative homologues, and to design comprehensive sets of primers and probes for monitoring OHRBs using for example functional gene arrays or multiplex qPCR detection platforms.

Furthermore, full genome sequences coupled with biokinetic assays have allowed characterization of several *rdh* genes, for example *Dehalococcoides ethenogenes* 195 DET0318 as a *pceA* gene, and cbdbA84 from *Dehalococcoides* sp. strain CBDB1 was recently identified as a chlorobenzene Rdh-encoding gene (*cbrA*) [31]. A key finding was the identification of VC reductase genes from *Dehalococcoides* spp. strains BAV1 (*bvcA*) [32] and VS (*vcrA*) [33], enabling their use as strain- and function-specific biomarkers in monitoring bioremediation. Similarly in *Desulfitobacterium hafniense* several chlorophenol Rdh-encoding (*cprA*) genes have been identified and characterized, like *crdA* and *cprA5* in *D. hafniense* strain PCP-1 [25,34] and *cprA1-cprA5* in *D. hafniense* strain DCB-2 [35,36]. Overall this suggests OHRB have the ability to degrade a wider than previously anticipated range of halogenated compounds. Therefore, exploring the functional diversity of the *rdh* gene-family can not only serve the identification of useful biomarkers in monitoring of different organohalide transformation potential in contaminated environments, but also provide novel leads towards yet-unknown degradation activities.

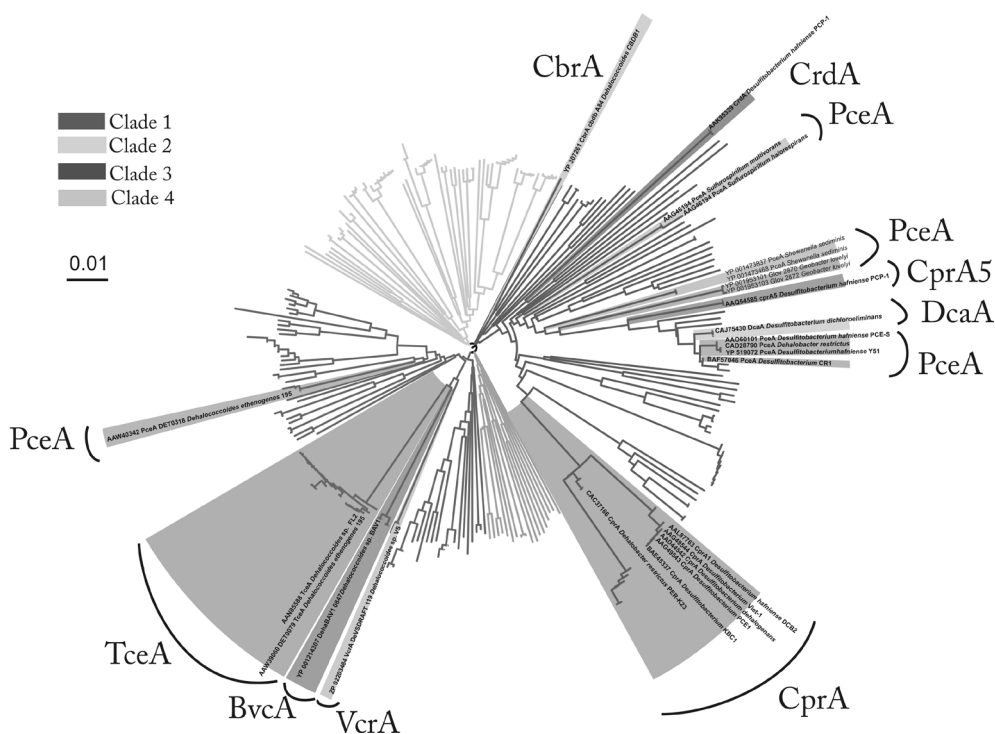


### *Transcriptional regulation of organohalide respiration*

Rdh gene clusters in OHRB are often associated with genes encoding putative regulatory proteins, suggesting that expression of *rdhAB* genes is in many cases tightly regulated [7]. In *Dehalococcoides* spp. nearly all *rdhAB* gene pairs have adjacent genes encoding two-component regulatory systems (RdhCD) or MarR-type regulators (RdhR). For example in the presence of trichlorinated benzenes *cbrA* expression is strongly induced by such a 2-component regulatory system [38]. Similarly, in *Desulfitobacterium* spp., expression of organohalide respiration-related genes is transcriptionally regulated, but this is mediated by FNR/CRP-type regulators (CprK) rather than 2-component regulatory systems. CprK regulatory proteins recognize specific promoter sequences (dehalo-boxes) in the presence of the corresponding halogenated substrate, initiating transcription of specific *rdhA* genes [39-41]. It has also been shown that a single chlorinated compound can induce transcription of multiple *rdhA* genes, albeit at different levels [38,42-44]. Genomics in combination with enzyme assays will help to elucidate the apparent complexity of the regulatory networks affecting transcription of *rdh* genes and their responses towards a wide spectrum of organohalides and other environmental stimuli such as concentration gradients of oxygen or other alternative electron acceptors and availability of different electron donors. Such information will help design bioremediation plans conducive for the quick biostimulation of the required OHRB at passive sites and/or speedy adaptation of bioaugmenting cultures.

### *From genomes to metagenomes*

Processes underlying bioremediation most often rely on intricate multispecies interactive networks. OHRB have been found to thrive in consortia rather than in pure culture, reinforcing the need to sequence the communities in which OHRBs exist. To this end the metagenomes of defined mixed cultures and entire microbial communities are currently being unraveled, allowing determination of metabolic and sensory interactions within consortia. Ongoing sequencing projects include (i) KB-1 culture, containing populations of *Dehalococcoides*, *Geobacter*, *Methanosarcina*, *Spirochaeta*, and *Sporomusa*; (ii) PCE-dechlorinating mixed community containing, *Dehalococcoides ethenogenes*, *Methanosaeta*, *Methanospirillum*, *Bacteroidetes*,



**Fig. 4:** Phylogenetic tree based on the deduced amino acid sequences of *rdhA* gene products. Specifically, these include for Clade 1 - *Dehalococcoides* RdhA, including PceA, TceA, BvcA and VcrA; Clade 2 - RdhA from *Dehalococcoides* and uncultured marine bacteria (no characterised representative); Clade 3 - RdhAs from other OHRBs as well as CbrA from *Dehalococcoides*; Clade 4 - Rhds from uncultured marine bacteria (no characterized representative). The tree was constructed by a neighbor-joining (NJ) method based on a ClustalW alignment of almost-complete amino acid sequences deduced from the respective gene sequence. The resulting tree was displayed using Interactive Tree Of Life (iTOL) [37]. Names are only displayed for characterized reductive dehalogenases. The scale bar represents 0.01 substitutions per amino acid position. Definitions: CbrA, chlorobenzene reductive dehalogenase; CrdA, 2,4,6-trichlorophenol reductive dehalogenase; CprA, chlorophenol reductive dehalogenase; DcaA, 1,2-dichloroethane reductive dehalogenase; PceA, tetrachloroethene reductive dehalogenase; TceA, trichloroethene reductive dehalogenase; VcrA and BvcA, vinyl chloride reductive dehalogenases.

and *Clostridiales* [45]; (iii) a highly stable and efficient dechlorinating bioreactor community with *Dehalococcoides* (<http://www.jgi.doe.gov/sequencing/allinoneseqplans.php>), and (iv) a defined *Dehalobacter-Sedimentibacter* coculture [22]. Several commercial dehalogenating bioreactor communities are in common use, and having their metagenomes will allow for modeling and optimization of the growth conditions at specific sites by supply of limiting nutrients and/or adjustment of operating parameters. It can be anticipated that with decreasing sequencing costs it will soon be feasible to routinely sequence directly the environmental metagenomes and transcriptomes to gain holistic impressions of the genetic and functional potential at a given site, allowing design of tailor-made interventions.

### Quantifying organohalide respirers

Although not intrinsically dependent on genomic information, quantitative PCR (qPCR) is an important element in the ecogenomics toolbox and is currently the most routinely used molecular tool in quantifying OHRB and their *rdh* genes. Although the use of qPCR in microbial ecology in general, and more specifically its application for quantifying *Dehalococcoides* populations, have been reviewed extensively detailing its prospects and limitations [46,47], it is important to note that this tool continues to provide valuable information for contaminated site management. The primer sets currently used for the detection of *Dehalobacter*, *Desulfitobacterium* [48] and *Dehalococcoides* [46] have become standard in field studies [49-53]. With the increasing genomic information these primer sets need to be further developed to target the increasing number of orthologs and also extended to additional gene functions in order to give a more comprehensive picture of the reductive dechlorination potential within a site. Such additional functions could e.g. include critical determinants of stress response to adverse environmental conditions, indicators of specific limitations, and pivotal factors in the syntrophic interactions with other members of degrading consortia. Also assays to routinely target new genera like *Dehalogenimonas* [54] and “Lahn” and “Tidal flat” Chlorofexi clusters [55] may be needed to broaden our understanding of their environmental distribution. Based on the gene copy numbers obtained for targeted dechlorinating bacteria and *rdh*'s, strategies for on-site remediation have sought to optimize growth and activity

of these OHRB. In cases, for which no, or low very numbers of, OHRB could be detected by qPCR, biostimulation via addition of nutrients or electron donors and bioaugmentation with dechlorinating consortia have been implemented [50,52,56].

### *Detection of new isolates*

Beyond routine environmental diagnostic applications, qPCR has been instrumental in characterization of new isolates or strains from enrichments and environmental samples. Analysis of copy numbers of *Dehalococcoides* 16S rRNA- and chloroethene *rdh* genes (i.e., *vcrA*, *tceA*, and *bvcA*) revealed that an enrichment culture from a contaminated aquifer consisted of multiple, distinct *Dehalococcoides* populations. Subsequent transfers, along with qPCR monitoring, yielded isolate *Dehalococcoides* GT that contained only *vcrA* [57]. Furthermore, using specific qPCR primers targeting *Dehalobacter* it was shown that cells of a new *Dehalobacter* isolate did not grow in the absence of a chlorinated acceptor, but only in the presence of 1,2-DCA [21]. Moreover, qPCR was used to show that two novel isolates within the Chloroflexi were able to couple reductive dechlorination of 1,2,3-trichloropropane to growth [58].

In view of the extreme diversity of *rdh* genes within and across genomes and metagenomes, it has become clear that specific qPCR assays are a particularly powerful tool in combination with other, more generic molecular profiling techniques. Combining Terminal-Restriction-Fragment-Length-Polymorphism and qPCR showed expression of 29 of 32 *rdhA* genes in *Dehalococcoides* sp. CBDB1 growing with 1,2,3- or 1,2,4-trichlorobenzene [38]. Recently qPCR in combination with functional gene microarrays has been used to track functional guilds involved in hexachlorobenzene degradation in European rivers [53].

### *Advances in qPCR*

Due to the ease and relatively low cost of qPCR compared to microarrays, it remains the most favorable technique for fast analysis of microbial numbers and indicators of bioremediation potential from in-situ samples. Emerging technological developments will further increase the throughput capacity of qPCR. Recent examples include microfluidic cards that contain 384 miniaturized qPCR assays

(<http://www.appliedbiosystems.com>), microfluidic-dynamic-array systems allowing 2,304 qPCR gene expression measurements in a single chip [59], Openarray™ accommodating 3072 33nL qPCR reactions [60], and the SmartChip™ Real-time PCR system equipped with high-density chips containing 5,000–30,000 nano-wells [61]. However, it has to be realized that qPCR as a highly sensitive technique is prone to errors and validity of resulting data sets should be considered with regard to specificity of primers, efficiency in DNA extraction methods and errors arising from PCR methodology and instrumentation [46].

### **Comparative genome hybridization and transcriptomics**

The availability of annotated genome sequences has enabled genome-wide analysis of strain-to-strain variation, as well as gene expression (transcriptomics), largely through the application of microarrays [9]. Two major classes of microarrays have been applied: (i) whole-genome arrays (WGAs), targeting all genes within the genome of a given isolate, and (ii) functional-gene arrays (FGAs), typically target genes that are key to microbially mediated biogeochemical processes, such as C, N and S cycling, phosphorus utilization, organic contaminant degradation and metal resistance, and redox cycling [62].

#### *Characterization of new strains*

Application of microarrays to analyze genomes of pure cultures has provided in-depth understanding of putative gene functions, not only of targeted organisms, for which the array was designed, but also of non-sequenced strains. *Dehalococcoides* comparative genome hybridizations (CGH) and transcriptome studies have been performed using a WGA with probes designed to cover >99% of the predicted protein-coding sequences of the *Dehalococcoides* strain 195 genome showing that the analysis of expression trends are as important as genetic characterizations [63–65]. Transcript profiling provided a systematic description of central metabolic pathways and also may assist in the assignment of specific functions to *rdhA* clusters [63,65]. Application of this array in CGH studies provided insight into the newly isolated *Dehalococcoides* sp. strain MB's complex nutrient requirements and its commitment to organohalide respiration [17]. Interestingly, the inability of *Dehalococcoides*

strain MB to dechlorinate dichloroethenes (DCEs) might be due to the absence of three genes: DET0079 (*tceA* gene), DET0876, and DET1559, which are thought to define the ability of strain 195 to dechlorinate TCE past DCE intermediates [17]. The potential of WGAs to speed up the characterization of novel OHRB isolates also needs to be exploited in the study of e.g. *Anaeromyxobacter*, *Desulfitobacterium* and *Geobacter* for which genome sequences are available.

#### *Distinguishing between subgroups*

WGA can be used to distinguish between *Dehalococcoides* subgroups in the environment, which can be key to determine the strains responsible for in situ dechlorinating activity. For example two strains of the Pinellas subgroup, *Dehalococcoides* sp. BAV1 and CBDB1, could be distinguished from strain 195 based on the hybridization of *rdh* genes on the *D. ethenogenes* microarray [17]. Similarly WGA-analysis of an *Dehalococcoides*-containing TCE enrichment from Alameda Naval Air Station (ANAS) revealed a difference in the *rdh* repertoire as 13 of the 19 *rdh*-like genes of strain 195 were not detected in ANAS [63]. In monitoring sites or bioreactors used in bioaugmentation, WGA can give a detailed overview of dehalogenation potential compared to available strains. Additionally WGA has shown that 88% of genes in predicted integrated genetic elements of strain 195 are not detected in ANAS, consistent with these elements being genetically mobile [64,65]. There is evidence for horizontal gene transfer of *rdh*'s in *Dehalococcoides* and *Desulfitobacterium* genomes [3,24]. As such WGA provides important leads for studies towards elucidating the extent and effects of these evolutionary events for spread and stability of reductive dehalogenation potential at polluted sites.

#### **Functional gene arrays**

FGAs are highly useful tools for comprehensive high throughput analysis of genomic potential and activities of microbial communities. For example, the GeoChip [62] can detect more than 10,000 catabolic genes involved in a broad variety of biogeochemical processes including pollutant degradation, and has recently been applied in the search for OHRB [53]. Probes targeting 153 *rdh* genes were added to the GeoChip and sediment samples taken from several locations within the Ebro river

basin with divergent histories of organohalogen pollution were analyzed, showing that diversity and abundance of *rdh*-like genes was affected by the sampling location [53]. In samples from a location with high hexachlorobenzene pollution, *rdh* genes of *Dehalococcoides* spp. strains CBDB1 and 195 dominated, whereas in samples from a location with a broader range of contaminants, a wide spectrum of *rdh* genes from *Dehalococcoides* and other OHRB were detected, indicating a higher diversity of potential degraders, albeit at lower relative abundance as revealed by genus-specific qPCR [53]. Thus bioremediation strategies for sites with multiple contaminants need to take into account optimizing for more organohalide respirers and not just *Dehalococcoides* in order to harness the full potential at these sites. Regardless of the challenges involved in microarray technology, such as suitable probe design, specificity, sensitivity, hybridization behavior, and quantification of target populations [9], it is anticipated that FGAs in combination with other techniques, like high throughput non-gel based proteomics [66] and metatranscriptome sequencing, will considerably enhance our understanding of microbial pollutant degradation.

### Concluding Remarks

Finally, the ecogenomic toolbox offers completely new ways to design and manipulate experiments geared at studying OHRB. For more in-depth understanding of organohalide respiration, the microorganisms and their respective communities collaboration between microbiologists, molecular ecologists and bioinformaticians, complemented with input from geneticists and evolutionary biologists is needed. Improvements in knowledge introduced by ecogenomics techniques has assisted in target selection for assessing and monitoring potential, specificity in detecting biomarkers, understanding spatial and temporal variability of key targets. Sensitivity, normalization and standardization of assays has also been improved and established. The mounting omics data on many environmental microbes, in particular OHRB, and the modeling of their individual and joint biological activities will guide interventions for stimulating the performance of desired biodegradation processes [11].

**Acknowledgments:**

We would like to acknowledge support by EU-FP7 Project Aquarehab (Project no. FP7-ENV-2008-1 226565), and the Netherlands Genomics Initiative (NGI) Innovative Cluster Ecogenomics (BSIK03011).



# CHAPTER 2

Chapter 2

## **Advanced Insight in a Hexachlorocyclohexane- Dechlorinating Coculture of *Dehalobacter* sp. and *Sedimentibacter* sp. Based on Metagenome Analysis**

Farai Maphosa, Mark W.J van Passel, Willem M. de Vos and Hauke Smidt

*Manuscript in preparation*

## ABSTRACT

The importance of organohalide-respiring *Dehalobacter* species in the bioremediation of chlorinated organics has been well recognized as is their limited range of metabolic modes. However, little is known about the full genomic repertoires of these dedicated degraders. Here we report detailed insight into the genome sequence of *Dehalobacter* sp. E1 that grows in strict coculture with *Sedimentibacter* sp. B4. The coculture metagenome and the genome of strain B4 were sequenced using 454 GS-FLX technology. Using the genome of strain B4 as a reference, all contigs belonging to the genome of strain E1 were filtered, binned separately and annotated. The resulting draft genome of strain E1 is approximately 2.6 Mbp in size and harbors 10 putative reductive dehalogenase (Rdh)-encoding gene clusters. Seven of these are most closely related to chlorophenol Rdh genes and 3 to tetrachloroethene Rdh genes, including the *pceABCT* operon in *D. restrictus*. This predicts that the dehalogenating potential of strain E1 is greater than previously anticipated. This was confirmed in degradation assays showing effective dechlorination of trichloroethene by the coculture, which was coupled to growth. Nearly all Rdh genes are located adjacent to genes encoding transcription regulators of the Crp/Fnr family or the MarR-type. Transcriptional analysis by RT-qPCR showed specific induction of several Rdh genes during growth on  $\beta$ -hexachlorocyclohexane. Advanced genome annotation indicated that *Dehalobacter* sp. is highly evolved to utilize chlorinated compounds. The strict dependence of strain E1 on *Sedimentibacter* sp. B4 is likely due to the fact that the larger *Sedimentibacter* genome of approximately 4.2 Mbp encodes a wide range of metabolic capabilities, including the catabolism of a large spectrum of carbohydrates, the biosynthesis of amino acids, and the production of several vitamins.

## Introduction

Subsurface microbiota comprises a large number of functional guilds, many of which play a central role in the bioremediation of polluted ecosystems. Organohalide respiring bacteria (OHRB) constitute one of those functional guilds, and are pivotal for detoxification of halogenated organics. OHRB thrive via anaerobic respiration with chlorinated pollutants as terminal electron-acceptors, a form of reductive dehalogenation that involves dedicated enzyme systems (reductive dehalogenases) to conserve energy [15,67]. OHRB have been isolated from a variety of contaminated environments around the world and include a phylogenetically wide range of 11 genera belonging to the Chloroflexi, Firmicutes, Delta Proteobacteria and Epsilon-Proteobacteria [68]. Following the genomic characterization of *Dehalococcoides* strain 195 [10], various other genomes have been fully characterized and presently the genomes are available of 8 strains belonging to the genera *Dehalococcoides*, *Desulfitobacterium*, *Geobacter* and *Anaeromyxobacter* [3,7,19,23,24]. The genomes of the four sequenced *Dehalococcoides* isolates are relatively small (1.34 -1.47 Mbp) and have a contextually conserved core that is interrupted by high plasticity regions near the origin of replication, which contain genomic islands and strain-specific genes [3,7]. Multiple copies of reductive dehalogenase (Rdh) genes are located within these genomic islands, suggesting they have initially been acquired by horizontal transfer [3,7]. In contrast, the genomes of the studied *Desulfitobacterium*, *Anaeromyxobacter* and *Geobacter* spp. are relatively large (3.87 – 5.73 Mbp), reflecting the versatile metabolic lifestyle of these OHRB that are able to utilize a wide range of electron donors and acceptors [19,23,25]. Remarkably, no genomic information is yet available for *Dehalobacter* spp, a widely spread group of Gram-positive OHRB. Three strains belonging to the genus *Dehalobacter* have been isolated and physiologically characterized, including *D. restrictus* strains PER-K23 [69] and TEA [70], which use tetrachloroethene and trichloroethene as terminal electron acceptors, and *Dehalobacter* sp. strain TCA1 [71], which grows by the dechlorination of 1,1,2-trichloroethane and 1,1-dichloroethane. All *Dehalobacter* spp. isolates are restricted in their metabolism to respiratory reductive dehalogenation. This feature is shared with *Dehalococcoides* spp. that, however, are phylogenetically unrelated and belong to the Gram-negative Chloroflexi (Chapter 1)[68]. Moreover,

the isolated *Dehalobacter* strains have a relatively low growth rate in pure culture, indicative of syntrophic interactions and it is not surprising that several other *Dehalobacter* strains have been obtained in defined co- and enrichment cultures with fermenting microorganisms, including a  $\beta$ -hexachlorocyclohexane ( $\beta$ -HCH) degrading culture of *Dehalobacter* sp. strain E1 and *Sedimentibacter* sp. strain B4 [22], 4,5,6,7-tetrachlorophthalide-dehalogenating consortia of *Dehalobacter* sp. FTH1 and FTH2 with *Clostridium* sp. and *Sedimentibacter* sp. [20], and a coculture of *Dehalobacter* sp. WL and *Acetobacterium* dechlorinating dichloroethanes [21,72]. *Dehalocobacter* spp. have also been detected in a 2,3,4,5-tetrachlorobiphenyl (2,3,4,5-TeCB)-dechlorinating sediment culture dominated by populations related to *Dehalococcoides* spp. [73]. The importance of *Dehalobacter* spp. in the bioremediation of chlorinated environments and their limited repertoire of other metabolic modes have been recognized at sites contaminated with chlorinated ethanes and chloroform [21,72,74]. *Dehalobacter* spp. are restricted in their catabolism to energy conservation via organohalide respiration using hydrogen as an electron donor in syntrophic interactions with for example *Sedimentibacter* spp. [20-22]. However, as genome sequences are lacking, we only have limited knowledge of their genes needed for the breakdown of chlorinated compounds, let alone their full metabolic capabilities. Therefore, we determined the draft genome sequence of *Dehalobacter* sp. strain E1, and its coculture companion *Sedimentibacter* sp. strain B4, by means of 454 FLX pyrosequencing. This coculture was the first known culture capable of metabolic dechlorination of  $\beta$ -HCH under anaerobic conditions [22]. Moreover, this coculture grows and dechlorinates more efficiently than the available pure cultures of *Dehalobacter*. It is very well possible that the presently isolated pure cultures of *Dehalobacter* spp. have been adapted to axenic growth and lost their efficient dechlorination capacity that is present in their natural habitats. Hence, the coculture is a better model for gaining insight into the in situ situation than the pure culture isolates. Based on the annotation of the draft genome that included 10 reductive dehalogenase (Rdh) genes, we predicted an extended dechlorination capacity of the coculture that was experimentally verified. Moreover, we showed the induction of a specific set of the Rdh genes, demonstrating the functionality of these genes coding for the key enzymes in the dechlorination of  $\beta$ -HCH.

## Materials and Methods

### Chemicals

The HCH isomers (>99% purity) were obtained from C.N. Schmidt B.V. (Amsterdam, The Netherlands). Peptone was purchased at Oxoid (Basingstoke, UK). Gas mixtures ( $N_2/CO_2$ ,  $H_2/CO_2$ ) were obtained at Hoekloos (Schiedam, The Netherlands). All other chemicals used were of the highest available purity and purchased from Sigma (Zwijndrecht, The Netherlands) and Merck (Amsterdam, The Netherlands).

### Cultivation of the Co- and Pure Cultures

The *Dehalobacter*–*Sedimentibacter* coculture and pure *Sedimentibacter* cultures were grown in 1 L serum bottles containing 500 mL of methanogenic mineral medium under  $H_2/CO_2$  (80:20 (v/v); 140 kPa) headspace with a final concentration HCH of approximately 170  $\mu M$  in the media as previously described [22]. Triplicate tests were performed for the chlorinated compound degradation experiments in 120 mL serum bottles containing 20 mL of methanogenic mineral medium. Estimation of chlorinated compounds utilization was done by measuring the release of chloride on a DX-600 IC system (Dionex Corporation, Salt Lake City, USA). Final concentration of tested chlorinated compounds in media were 10 mM of 3-chloro-4-hydroxyphenylacetic acid (CHPA) and 500  $\mu M$  each for trichloroethene (TCE), cis-dichloroethene (cis-DCE), 2,3-dichlorophenol (2,3-DCP), 2,4-dichlorophenol 92,4-DCP) and 2,4,6-trichlorophenol (2,4,6-TCP).

### Genome Sequence Determination

Coculture and *Sedimentibacter* sp. cells were harvested by centrifugation and used for high molecular weight DNA isolation using the standard Bacterial genomic DNA isolation using Cetyl trimethylammonium bromide (CTAB) method recommended by JGI (Joint Genome Institute, Walnut Creek, CA) with minor modifications. In short, cells were resuspended in 14.8 ml modified TE (10 mM Tris; 20 mM EDTA, pH 8.0). Subsequently, cells were lysed using lysozyme and proteinase K, and DNA was extracted and purified using CTAB and phenol:chloroform:isoamylalcohol extractions. After precipitation in isopropanol and washing

in 70% ethanol, DNA was resuspended in 400 µl TE containing 40 µg RNase A. Quality and quantity of the DNA were checked using agarose gel electrophoresis and spectrophotometric measurement using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE), respectively, before the DNA was shipped for sequencing.

DNA was prepared for sequencing using the GS FLX General Library Preparation Kit (Roche, Germany) and sequenced at GATC Biotech (Konstanz, Germany) on the GS FLX using the manufacturer's protocols. The sequencing data were assembled using the GS De Novo Assembler (Roche).

Filtering of the *Dehalobacter* genome sequence data from the metagenomic sequence of the coculture was done based on BLAST hits and G+C content. The *Sedimentibacter* genome sequence was used to query the metagenomic coculture sequence. Contigs of more than 150 bp in length, with at least 2 reads, and having no hit with the *Sedimentibacter* contigs of more than 90% identity for at least 100 bp were binned as *Dehalobacter*. BLAST and data parsing were done using in-house Perl scripts. Secondly contigs were sorted based on G+C percentage. Dinucleotide frequencies were used to check assignment of sequence data as either *Sedimentibacter* or *Dehalobacter*. The Compare Genomic Islands application [75,76] allowed for the compositional comparison between contigs.

### Sequence Analysis, Annotation and Comparative Genomics

Genome annotation was done using the Rapid Annotation using Subsystem Technology (RAST) [77] and the metagenome pipeline MG-RAST [77]. Metabolic pathways were examined by using the SEED [78] and KEGG databases [79]. The tRNA genes were further confirmed by tRNAscan-SE [80]. Artemis v11 was used to organize data and further facilitate annotation [81]. Orthologs were defined using ORTHOMCL [82]. The phylogeny was reconstructed using orthologous gene sets identified from other bacterial genomes using ORTHOMCL [82], which were aligned with MUSCLE [83] for maximum likelihood trees. Furthermore, putative functions for reductive dehalogenase genes and associated genes related to organohalide respiration processes were inferred using BLAST against the National Center for Biotechnology Information databases [84]. Phylogenetic tree reconstruction

was performed with reductive dehalogenase catalytic subunit (RdhA) sequences. Deduced amino acid sequences were aligned with ClustalW [85]. Phylogenetic trees were calculated using the neighbour joining and maximum likelihood methods in Bioedit [86]. Stability of the tree topology was further refined by bootstrapping (1,000 replications). Only RdhAs  $\geq 400$  amino acids were included and trees were displayed using ITOL [37].

### Identification of Repeated Elements and IS Elements

Repeated elements greater than or equal to 18 bp in length were identified using the repeat-match algorithm in MUMmer3 [87]. IS elements and IS-transposases were detected by BLAST and BLASTP searches and by manual inspection of the genomic context surrounding significant search hits to IS elements and transposases in the ISFinder database [88].

### RNA Isolation

The coculture and *Sedimentibacter* cultures were grown in 500 ml of media in the presence of 170  $\mu\text{M}$   $\beta\text{-HCH}$ . After 10 days of growth, when the culture media appeared colourless showing depletion of  $\beta\text{-HCH}$ , another pulse of  $\beta\text{-HCH}$  (170  $\mu\text{M}$  final concentration) was added to induce the cells. Batches were sacrificed to obtain samples one day after induction and after 30 days (20 days after induction). Microbial cells were harvested from 500 ml cultures and RNA was extracted by bead-beating with Macaloid clay [89], sodiumdodecylsulfate (SDS) and acidic phenol. Nucleic acids in the supernatant were purified by consecutive extraction with phenol/chloroform/isoamylalcohol (25 : 24 : 1) and chloroform/isoamylalcohol (24 : 1). Contaminating genomic DNA was removed with the RNeasy Mini Kit (Qiagen, Venlo, The Netherlands). In order to eliminate traces of genomic DNA that co-purified with the RNA samples, two consecutive DNase I (Roche, Almere, The Netherlands) treatments were done according to supplier protocols. After elution, RNA integrity was checked by standard agarose gel electrophoresis and quantified with a NanoDrop spectrophotometer (NanoDrop Technologies). Synthesis of cDNA was performed on 500 ng total RNA (double DNaseI-treated), using random hexamers and the SuperScript<sup>TM</sup> III First-Strand Synthesis System (Invitrogen Life

Technologies, Breda, The Netherlands), according to the manufacturer's instructions. Removal of contaminating genomic DNA was confirmed by PCR on non-reverse transcribed RNA samples with primers Dre441F and Dre645R [48] targeting the *Dehalobacter* 16S rRNA gene.

### **Quantitative PCR assays**

Gene expression was measured for 9 target genes using the iQ5 iCycler (BioRad). Quantitative PCR (qPCR) was performed on 25 ng of cDNA sample using the iQ SYBR Green Supermix kit (BioRad) in triplicate 25 µl reactions. Primers Dre441F and Dre645R [48] were selected for qPCR of the 16S rRNA gene of *Dehalobacter*. Primers for qPCR targeting the reductive dehalogenase genes (Table 1) were designed using Primer3 software [90], (<http://frodo.wi.mit.edu/primer3/input.htm>). Quantitative PCR amplification parameters were: 10 min at 95.0°C, followed by 40 amplification cycles of 15 sec at 95.0°C, 30 sec at 60.0°C, 30 sec at 72.0°C. The transcript levels were normalized to 16S rRNA transcripts.



**Table 1:** Primers used in this study

	<b>Forward primer</b>	<b>Reverse primer</b>	
Dhb_965F	aagatagcaaatgctggctgt	ccattccagccaccacttat	Dhb_965R
Dhb_968F	gggagtgattggagcgataa	ctgtacatctccggccact	Dhb_968R
Dhb_985F	acattgtggctcccgtaaag	ttcagcttgccatttgagtg	Dhb_985R
Dhb_950F	ccttctcatgaagacgcaca	ttctgtccctccgaagctaa	Dhb_950R
Dhb_968F	aaccaagccgcaatatcaag	tatgaagctctggccattcc	Dhb_968R
Dhb_490F	tccatacaaggtggatgcaa	gctacatcgtggttcaggt	Dhb_490R
Dhb_1133F	gcactggcagtagggattgt	ccattccagccaccatttat	Dhb_1133R
Dhb_1222F	tggcagattcttataaaagtaaaaa	caaactgtttacggcgatg	Dhb_1222R
Dhb_1238F	tacgatgatgggatgggaat	ccaccaaggcacgtctttat	Dhb_1238R
Dhb_1153F	tgggaggaagaaaatgaaaaa	ttctcttgcttcgtaagcttct	Dhb_1153R

<b>Primers for quantitative PCR</b>			
Dhb_965_QF	gatttctgccgggtttgtaa	cactgttcagcgcagataa	Dhb_965_QR
Dhb_968_QF	atgaagtgcgtatcggcaag	accatcacgacgatgacgta	Dhb_968_QR
Dhb_985_QF	tttctgcacgatgacgttc	ccagtttacgccatccaact	Dhb_985_QR
Dhb_950_QF	cgtacggtagagcatgaca	atccttcaggcggagtact	Dhb_950_QR
Dhb_84_QF	acggttattacggcacgaag	cgcattgactggcttacaga	Dhb_84_QR
Dhb_490_QF	gaagctctacgcacctcacc	cggccataggaacactcatt	Dhb_490_QR
Dhb_1133_QF	ccttcgcaggatgattgtt	atcgtttcaaatgctggtc	Dhb_1133_QR
Dhb_1222_QF	atgggatttgcgttagttgc	tttgctcgggatcaggtatc	Dhb_1222_QR
Dhb_1238_QF	taagccaaaaaggcgaagaa	gttgcaatccccacatc	Dhb_1238_QR
Dhb_1153_QF	atcagaggcgaacaaaagga	gtcaatcgatgccacaaatg	Dhb_1153_QR

## Results and Discussion

### Reconstructing the *Dehalobacter* Genome Based on the Coculture Metagenome and the *Sedimentibacter* Genome

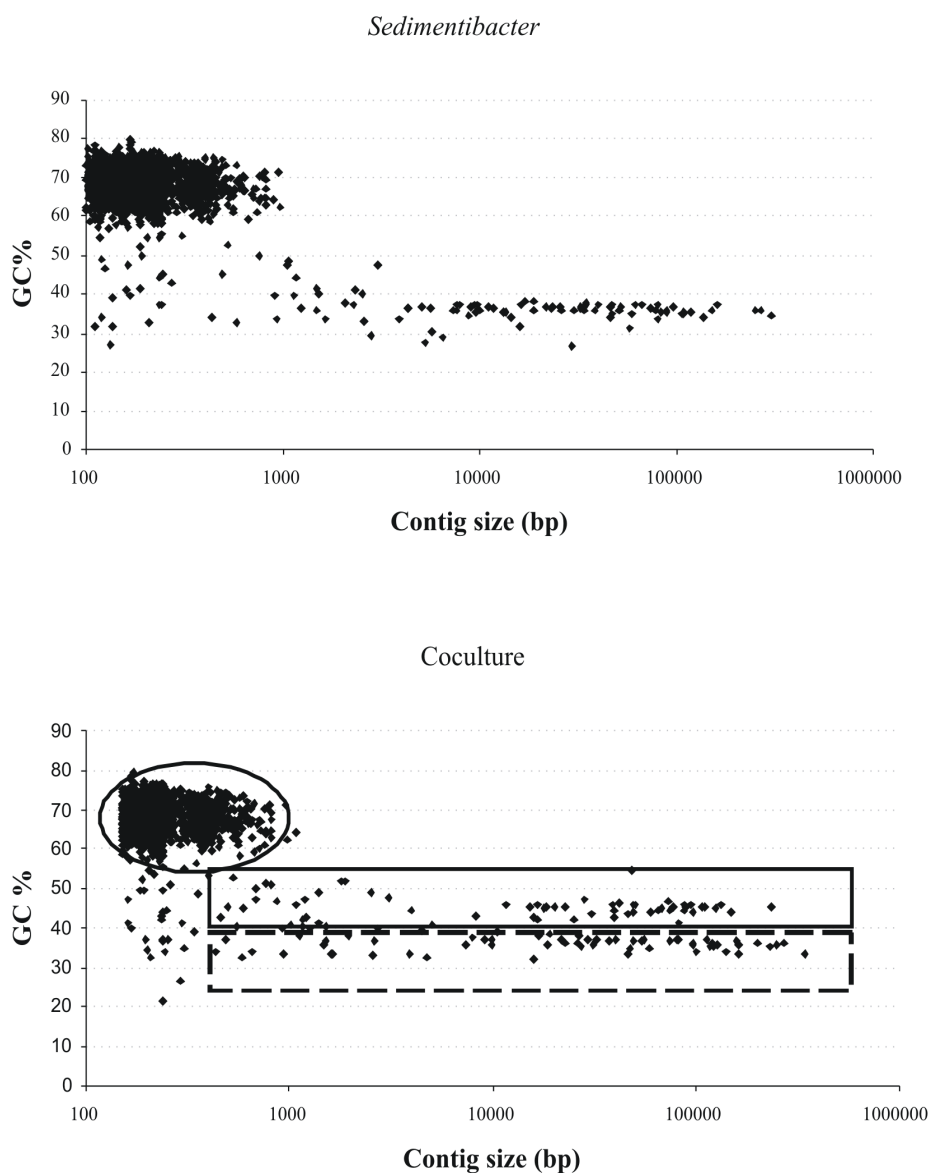
To reconstruct the *Dehalobacter* sp. strain E1 genome we compared and contrasted the coculture metagenome and the genome of *Sedimentibacter* sp. strain B4 grown in pure culture. Analysis of the sequencing data output (Table 2) for the coculture metagenome resulted in 3711 contigs greater than 100 bp, whereas that of the *Sedimentibacter* sp. strain B4 preassembled into 2400 contigs greater than 100 bp. Firstly, based on G+C composition analysis, two distinct clusters were obtained from the coculture sequence data comprising of *Sedimentibacter* sp. strain B4 sequences with a relatively low G+C percentage and a second cluster with higher G+C composition from the *Dehalobacter* sp. strain E1 sequence (Fig. 1). A third cluster of short sequences, less than 1000 bp, with a very high G+C composition (60-80 %) was also obtained, with the majority of these sequences belonging to the *Sedimentibacter* genome.

**Table 2:** 454 GS-Flx –sequencing output

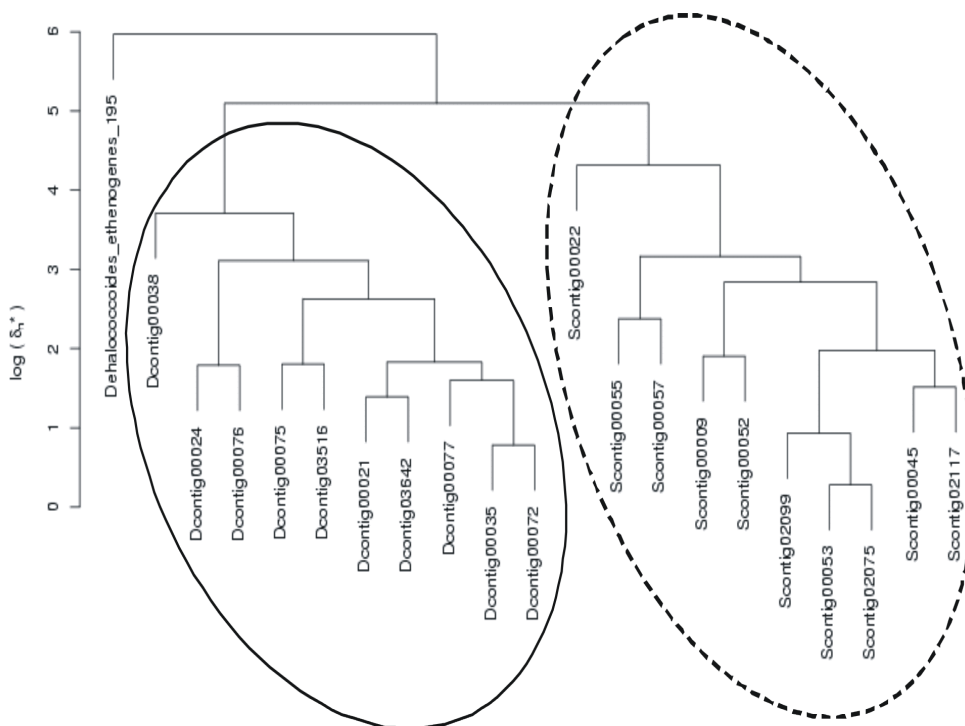
	Coculture metagenome	<i>Sedimentibacter</i> sp. B4 genome
Total Number Of Reads	245560	251870
Total Number of Bases	58487358	58278990
Number of Contigs >100bp	3711	2400
Total bases in contigs >100bp	5606204	4258799
Largest contig size (bp)	233646	306841

Secondly, using composition similarity analyses of the contigs enabled us to confirm and bin the *Dehalobacter* sp. strain E1 sequence data. Each genome has typical dinucleotide frequencies and therefore related species or individual sequences within a genome have a similar genome signature, whereas unrelated species have dissimilar genome signatures [91,92]. Analysis of large contigs (> 1000 bp) binned to the *Dehalobacter* genome sequence showed a different genome signature than those

of *Sedimentibacter* (Fig. 2).



**Fig. 1:** Separation of contigs based on G+C percentage. (A) *Sedimentibacter* sp. strain B4 contigs and (B) – a combined assembly of: Solid-line rectangle – *Dehalobacter*, dashed-line rectangle – *Sedimentibacter* and short contigs with very high G+C content mostly from the *Sedimentibacter* genome.



**Fig. 2:** Hierarchical clustering of the normalized genome dissimilarity scores for the contigs of each of *Dehalobacter* sp. strain E1 and *Sedimentibacter* sp. strain B4 in comparison with the complete genome of *Dehalococcoides ethenogenes* 195 [10]. Depicted in solid and dashed circles are the *Dehalobacter* and *Sedimentibacter* contigs respectively, indicating they have similar genome signatures.

### General Description of *Dehalobacter* sp. strain E1 and *Sedimentibacter* sp. strain B4 genomes

The *Dehalobacter* sp. strain E1 genome sequence reads were assembled into 111 contigs greater than 150 bp with at least two reads, resulting in a draft genome size of 2,607,109 bp. The *Sedimentibacter* sp. strain B4 sequence assembly yielded 1711 contigs greater than 150 bp, giving rise to a draft genome of 4,202,354 bp. The total estimate from all the *Sedimentibacter* sequences from contigs with at least two reads is 4,255,211 bp. Due to the large number of small contigs with very high (75%) G+C content (Fig. 1), further annotation of the *Sedimentibacter* sequence was done for 91

**Table 3:** Overview of genome characteristics of *Dehalobacter* sp. strain E1 and comparison to genomes of other OHRB<sup>1</sup>.

Genome Name	Bases (Mbp)	G+C (%)	Genes	CDS	tRNA + other RNAs	16S rRNA genes	rdhA genes	Formate dehydrogenases	Hup-type hydrogenases	Other Ni-Fe-type hydrogenases
<i>Anaeromyxobacter dehalogenans</i> 2CP-C	5.01	75	4419	4361	56	2	2	8	5	2
<i>Geobacter lovleyi</i> SZ	3.87	55	3514	3476	37	1	2	2	1	2
<i>Desulfotobacterium hafniense</i> DCB-2	5.28	48	4801	4712	84	5	7	5	2	3
<i>Desulfotobacterium hafniense</i> Y51	5.73	47	5137	5060	71	6	4	2	3	1
<i>Dehalobacter</i> sp. strain E1	2.6	45	2587	2532	55	3	10	2	3	4
<i>Dehalococcoides ethe- nogenes</i> strain 195	1.47	49	1641	1591	50	1	17	1	1	4
<i>Dehalococcoides</i> sp. CBDB1	1.34	47	1436	1385	57	1	32	1	1	4
<i>Dehalococcoides</i> sp. BAV1	1.39	47	1516	1458	50	1	10	1	1	4
<i>Dehalococcoides</i> sp. VS	1.41	55	2160	2096	63	1	36	1	1	4

<sup>1</sup>Genes: total gene features; CDS: coding sequences; RNA: number of rRNA, tRNA and other RNA genes; 16S: number of 16S rRNA gene copies; rdhA genes: confirmed and predicted reductive dehalogenase-encoding genes

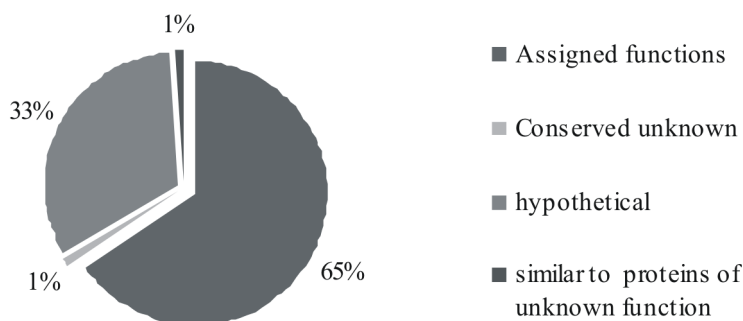
contigs with at least 1000 bp (total sequence 3,800,057 bp) on the RAST Server. All other contigs (<1000 bp) were analyzed on MG-RAST as metagenomic dataset. The G+C content of the *Dehalobacter* sp. strain E1 genome is estimated to be 45.1 % while that of the *Sedimentibacter* sp. strain B4 is 38%.

The deduced genome sequence of *Dehalobacter* strain E1 was compared to the genomes of other OHRB (Table 3). *Desulfitobacterium hafniense* strains Y51 and DCB2 are members of the Firmicutes and closely related to *Dehalobacter*, whereas *Dehalococcoides* species are phylogenetically distantly related but share the restricted lifestyle and dedication to reductive dechlorination with *Dehalobacter* spp.. While the predicted genome of *Dehalobacter* sp. E1 has a similar, although somewhat larger, size than the four *Dehalococcoides* spp. genomes, it is significantly and more than 2 Mbp smaller than currently known *Desulfitobacterium* genomes (Table 3). It is also smaller than the genomes of *Geobacter lovleyi* and *Anaeromyxobacter dehalogenans*, which belong to the  $\delta$ -Proteobacteria. The relatively small genome size of strain E1 is in line with the restricted metabolism observed for *Dehalobacter* spp. strains. In contrast, *Desulfitobacterium*, *Anaeromyxobacter* and *Geobacter* spp. are more versatile in their metabolism [23,25] and are found in more diverse environments compared to *Dehalobacter* and *Dehalococcoides*. The small genome size of strain E1 could be a result of gene loss over time as its growth depends on syntrophic relationships with fermentative microorganisms like *Sedimentibacter* and *Clostridia* that contain large genomes and may encode functions that compensate for the restricted metabolism of *Dehalobacter* spp.

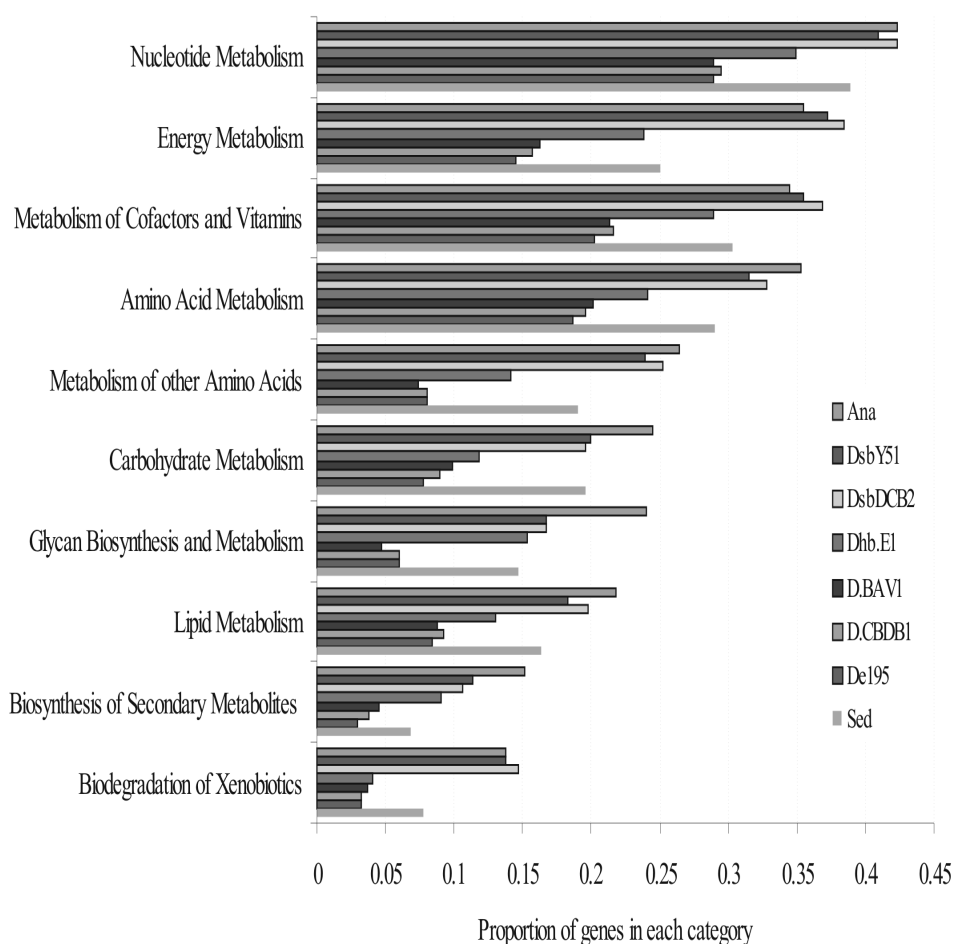
There are 55 tRNA genes distributed throughout the *Dehalobacter* genome and there are 3 predicted rRNA operons. In agreement with the presence of multiple rRNA copies, three non-identical 16S rRNA gene sequences have previously been identified via sequencing and DGGE analysis of the coculture [22]. Similarly, it was also observed by qPCR in *Dehalobacter* WL cultures that there were approximately four rRNA operons in the *Dehalobacter* strains [21]. *Desulfitobacterium hafniense* Y51 and DCB2, which are phylogenetically related to the genus *Dehalobacter* have six and five rRNA operons, respectively [19,24].

Of the 2587 predicted CDS 65% have an assigned function while 33% are annotated as hypothetical proteins (Fig. 3). To get an indication of the metabolic

profile of *Dehalobacter* sp. E1, we examined the metabolic pathways represented in the KEGG database in comparison to those of other OHRB (Fig. 4). This type of analysis shows an overall view of the metabolic capacity based on the percentage of genes within a genome that give hits with representative genes from 10 main KEGG categories. Overall, and in line with its slightly larger genome size, strain E1 has a higher proportion of gene representatives in each KEGG category compared to *Dehalococcoides*, and lower gene proportions compared to the larger genomes of *Desulfitobacterium* and *Anaeromyxobacter*. All genomes harbor a high proportion of genes for pathways involved in nucleotide, energy, amino acid, and cofactor and vitamin metabolism. The four complete *Dehalococcoides* genomes have similar proportions of genes represented in KEGG categories in line with the observation that there is a conserved core genome accounting for 68-77% genes of each individual genome [3]. The proportion of genes for biodegradation of xenobiotics is similar for *Dehalococcoides* and *Dehalobacter*. No particular KEGG category was found to be overly represented in *Dehalobacter* when compared to other OHRB.



**Fig. 3:** General protein coding features of the *Dehalobacter* genome. Hypothetical – no significant similarity to any other sequenced gene, conserved unknown – conserved hypothetical protein with sequence similarity to a translation of an ORF in another organism, however, there is no experimental evidence for its protein expression. Unknown function – significant sequence similarity to a named protein for which no specific function is currently assigned.



**Fig. 4:** Comparative analysis of genes by functional KEGG categories, from the genomes of *Dehalobacter* sp. strain E1 (Dhb.E1), *Anaeromyxobacter dehalogenans* 2CP-C (Ana), *Desulfotobacterium hafniense* strains Y51 (DsbY51) and DCB2 (DsbDCB2), *Dehalococcoides ethenogenes* 195 (De195), *Dehalococcoides* sp. strain CBDB1 (D.CBDB1), BAV1 (D.BAV1) and *Sedimentibacter* sp. strain B4 (Sed). Proportion of total genes in each KEGG category is shown for each genome as a fraction of the total possible number of gene currently in that KEGG category.



## Reductive dehalogenases in *Dehalobacter* sp. strain E1

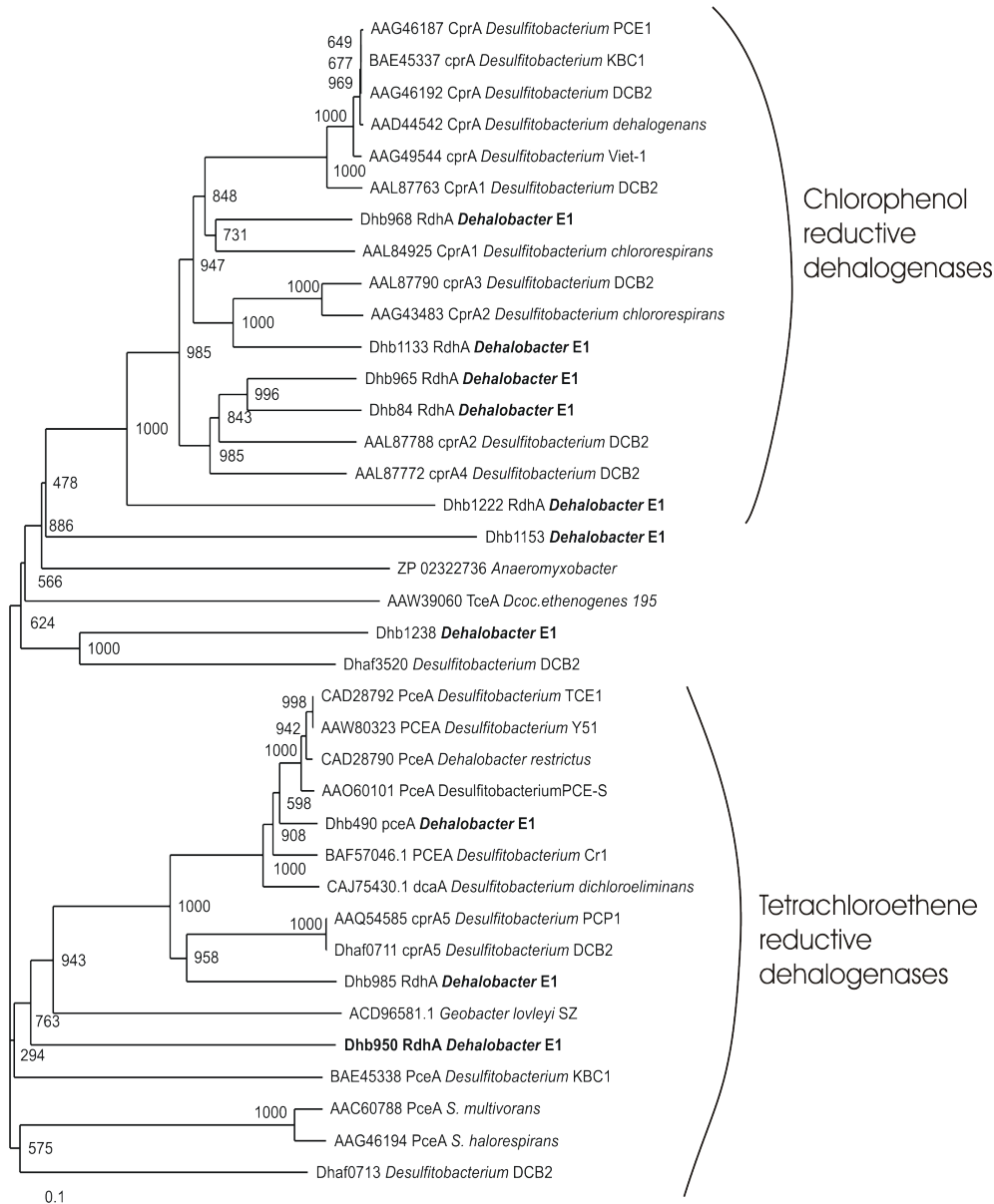
Studies in *Dehalobacter restrictus* have yielded important information on the biochemical properties of the PceA reductive dehalogenase [93-95]. Recently it was furthermore shown that a 1,2-DCA dechlorinating *Dehalobacter* has multiple Rdh enzymes [21]. Because of its dedication to organohalide respiration, we particularly examined the genome of strain E1 for the presence of Rdh-encoding genes. Based on the RAST analysis, there are 10 open reading frames annotated as Rdh in the *Dehalobacter* genome (Fig. 5). No Rdh-encoding genes were found in the *Sedimentibacter* genome, further confirming that strain E1 is responsible for all  $\beta$ -HCH dechlorination activity observed in the coculture [22]. Phylogenetic analysis showed that seven of the putative Rdh's clustered with characterized chlorophenol Rdh's (CprA) while three formed a cluster with known tetrachloroethene Rdh's (PceA) (Fig. 5). The genes encoding Rdh catalytic subunits (RdhA) are all organized in gene clusters with predicted membrane-docking protein genes *cprB* or *pceB*, except for one of the *cprA*-like genes (DhbA1153), which is annotated as an ortho-chlorophenol Rdh gene (Table 4). This gene is 1074 bp in size and does not have an associated *cprB* gene. All annotated *rdhA* genes have characteristic nucleotide sequence lengths between 1245-1845 bp and their predicted Rdh proteins all contained the consensus sequence CXXCXXCXXXCP and GXXCXXCXXXCS, which is characteristic for  $\text{Fe}_4\text{-S}_4$  centers [18].

With respect to the individual predicted Rdh gene clusters, contig00022 (107206 bp) harbors the composite equivalent to the *Dehalobacter restrictus* strain PER-K23 *pceABCT* operon [95], and the transposon Tn-Dha1 and *pce* operon of *Desulfitobacterium hafniense* strain TCE1 [96] (Fig. 6). The predicted *Dehalobacter pceA* gene product shows 90% identity to that of *Dehalobacter restrictus* and *Desulfitobacterium hafniense* PCE-S and TCE1. The gene cluster found in contig00022 and the *pceABCT* genes of strain PER-K23 share 97% identity over a region of 8625 bp, including a transposase (IS3/IS911 family protein), an uncharacterized Archaeon Conserved Region (COG2078) gene involved in tRNA modification in Archaea and a gene for a putative pyruvate-formate lyase-activating enzyme (Fig. 7). The conservation of the complete *pceABCT* cluster as well as the transposase-

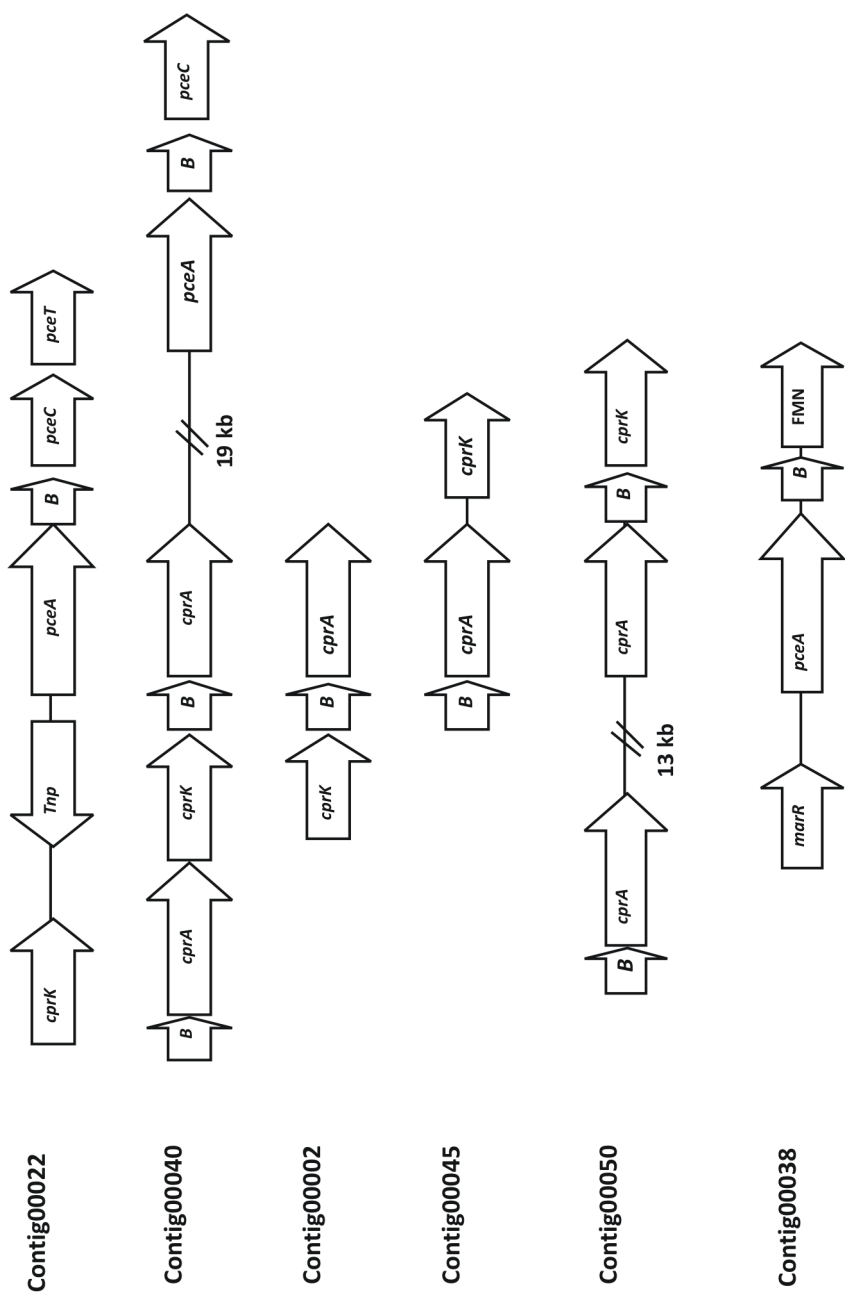
encoding gene suggests that this has been inherited from a common ancestor of these two strains. In contrast, the *Desulfitobacterium hafniense* TCE1 operon is almost identical (97%) over a stretch of 4593bp that includes the *Dehalobacter pceABC* genes only but not the transposase-encoding gene. Previously it was reported that the structure of the Tn-Dha1 transposon associated with *Desulfitobacterium hafniense* TCE1 *pceABCT* was not conserved in *Dehalobacter restrictus* [96].

**Table 4 :** Characteristics of the Rdh genes identified in the *Dehalobacter* genome.

Feature ID	Contig Number	Strand	Size (nt)	Protein sequence/s with highest identity	% identity
Dhb84	Contig 00002	-	1383	AAL87772 CprA4 <i>Desulfitobacterium</i> DCB2	64.8
Dhb490	Contig 00022	+	1686	AAO60101 PceA <i>Desulfitobacterium</i> PCE-S CAD28790 PceA <i>Dehalobacter restrictus</i> CAD28792 PceA <i>Desulfitobacterium</i> TCE1	90
Dhb985	Contig 00040	+	1647	Dhaf0711 CprA5 <i>Desulfitobacterium</i> DCB2	66.7
Dhb968	Contig 00040	+	1389	AAL84925 CprA1 <i>Desulfitobacterium chlororespirans</i>	66.7
Dhb965	Contig 00040	+	1395	AAL87772 CprA4 <i>Desulfitobacterium</i> DCB2 AAG43483 CprA2 <i>Desulfitobacterium chlororespirans</i>	69
Dhb1133	Contig 00045	-	1341	AAL87790 CprA3 <i>Desulfitobacterium</i> DCB2	68.9
Dhb1222	Contig 00050	+	1245	AAL87772 CprA4 <i>Desulfitobacterium</i> DCB2	37.8
Dhb1238	Contig 00050	+	1584	Dhaf3520 ZP01373225 <i>Desulfitobacterium</i> DCB2	25.4



**Fig. 5:** Phylogenetic analysis of the predicted Rdh enzymes. Deduced amino acid sequences were aligned with ClustalW [85]. Phylogenetic trees were calculated using the neighbour joining and maximum likelihood methods in Bioedit [86]. Stability of the tree topology was further refined by bootstrapping (1,000 replications). Tree was displayed using ITOL [37]. *Dehalobacter* RdhA are highlighted in bold face.



**Fig. 6:** Gene arrangements of reductive dehalogenase related loci. Chlorophenol reductive dehalogenase (*cprA*), tetrachloroethene reductive dehalogenase (*pceA*) and membrane anchor protein gene (*B*), Transcriptional regulator genes *marR*, *pceC* and *cprK*, FMN-binding, *Tnp* transposition related sequence.

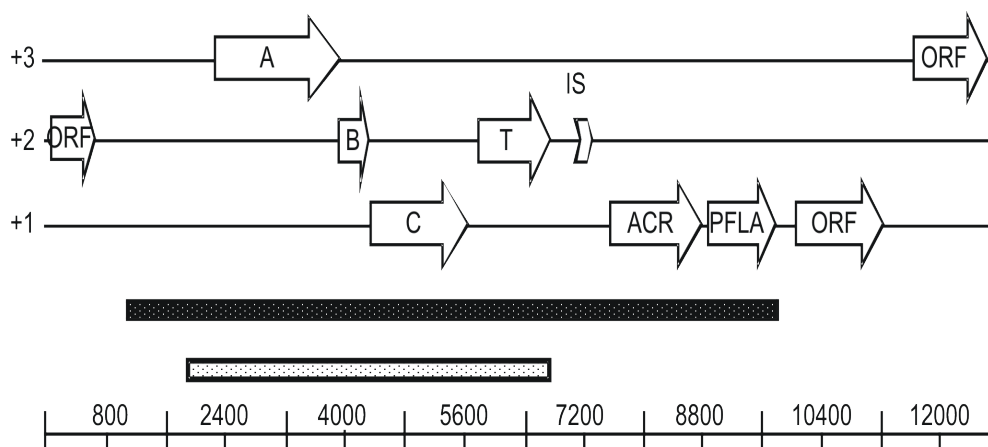


Fig. 7: The *pceABCT* operon on Contig00022 of the *Dehalobacter* sp. strain E1 genome sequence. Genes are shown in their reading frame (indicated on the left) bottom axis represents bp position on contig00022 (size of the contig 107026 bp), thick black line show alignment of *pceABCT* operon from *Dehalobacter restrictus* and dotted line *Desulfitobacterium hafniense* TCE1. IS represents an insertion sequence element of the IS3/IS911 family, ACR an uncharacterized Archeon Conserved Region (COG2078) and PFLA a putative pyruvate-formate lyase-activating enzyme-encoding gene.

Furthermore, contig00040 harbors three Rdh genes. The first cluster of these genes is most similar to the *cprBA2* - *cprBA3* region of *Desulfitobacterium hafniense* DCB2, while the second cluster is located 19 kb downstream and resembles the *cprBA5* operon. Dhb985 on contig00040 was annotated by the RAST pipeline as *pceA*, however, the amino acid sequence shares 66 % identity with CprA5 of *D. hafniense* strains PCP-1 and DCB-2. CprA5 of strain PCP-1 has been characterized as a 3,5-dichlorophenol reductase, and expression of the corresponding gene in *D. hafniense* DCB-2 was found to be induced by ortho- and meta-substituted chlorophenols (Chapter 3 of this thesis). Hence, it would be interesting to determine whether Dhb985 has similar function (see below).

Finally, four of the annotated Rdh genes most likely code for yet uncharacterized Rdh's having less than 40% identity to any of the known proteins. Similarly, a novel Rdh has recently been identified in *Dehalobacter* WL mixed cultures [21].

The sequences upstream of the Rdh genes were examined for promoter binding sites that contain a specific DNA sequence termed the dehalobox (TTAAT-N4-ATTAA). Dehalobox consensus sequence motifs have been identified in *Desulfitobacterium hafniense* DCB2 [36,40] as the DNA binding sites for the transcriptional regulators (CprK). Sequence motif search did not give any significant results. No dehaloboxes similar to those in *Desulfitobacterium* could be identified suggesting there should be different binding sites for these transcriptional regulators in the genome of *Dehalobacter* sp. strain E1.

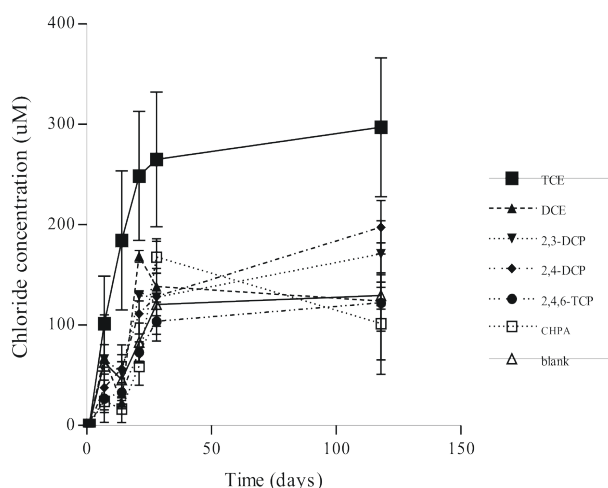
### **Rdh- associated genes**

The Rdh enzymes all are predicted to have a characteristic putative membrane-anchor protein encoded by the *rdhB* gene, which forms a bicistronic operon with the catalytic subunit-encoding gene (*rdhA*) [15,18]. Secondly the operons are all associated with transcriptional regulator-encoding genes. Five of the *Dehalobacter cprA* genes are located in close proximity to genes predicted that code for CRP/FNR superfamily (*cprK*) transcriptional regulators similar to those found in *Desulfitobacterium* spp. (Fig. 6) [97]. The *pceABCT* cluster contains *pceC*, which is thought to encode a membrane bound regulatory protein [98]. Interestingly there is also a *cprK* gene located 1kb upstream from this operon, and its gene product might also exert a regulatory effect on the *pceABCT* operon. One operon located on contig00038 contains a MarR-type regulator (*rdhR*) encoding gene. These regulators have been described to be present in *Dehalococcoides* spp. genomes associated with Rdh genes. The *Dehalococcoides* sp. CBDB1 genome contains 16 *rdhR* genes, whereas seven were reported for *Dehalococcoides ethenogenes* 195 [7,10]. The *Dehalobacter* MarR-type regulator is associated with one of the unique Rdh's (Dhb950), which has highest identity at only 18.9% to *Geobacter lovleyi* PceA. The MarR-type regulator gene is positioned 463 bp upstream to Dhb950 *rdhA*. Similarly this novel cluster of *rdhR*'s in *Dehalococcoides* genomes have been observed to be mostly located 200-320 bp upstream of the corresponding *rdhA* [7].

### **Dehalogenation of Chlorinated Compounds**

A degradation experiment was conducted following observations that

there were 10 putative Rdh's encoded in the *Dehalobacter* genome, giving rise to the possibility of a wider dechlorination potential than previously observed. The coculture has previously been shown to actively dechlorinate  $\beta$ -HCH, but not PCE or 1, 2-DCA [22]. We further tested for degradation of CHPA, TCE, *cis*-DCE, 2,3-DCP, 2,4-DCP and 2,4,6-TCP. Remarkably, release of free chloride was observed for TCE but not for the other compounds (Fig. 8), suggesting that the coculture is able to dechlorinate TCE. This is the first indication of this potential and more comprehensive experiments will be needed in the future to determine degradation rates. The observed TCE-dechlorination activity is in agreement with the presence of a complete *pceABCT* operon and 2 other genes annotated as tetrachloroethene reductive dehalogenases in strain E1. No degradation was observed for *cis*-DCE after 118 days, in line with the fact that the PCE reductive dehalogenase (PceA) of *Dehalobacter restrictus* is only able to degrade PCE to *cis*-DCE as the end product [95]. None of the chlorophenols tested were degraded by the coculture, suggesting that this *Dehalobacter* sp. strain is not capable of utilizing these compounds in spite of the presence of genes most closely related to chlorophenol Rdh genes of *Desulfotobacterium hafniense* DCB2 [25,36], which has been shown to dechlorinate all compounds tested here.



**Fig. 8:** Chloride formation as an indication of chlorinated compound reduction by the coculture.

### Expression of reductive dehalogenase genes in $\beta$ -HCH grown cultures

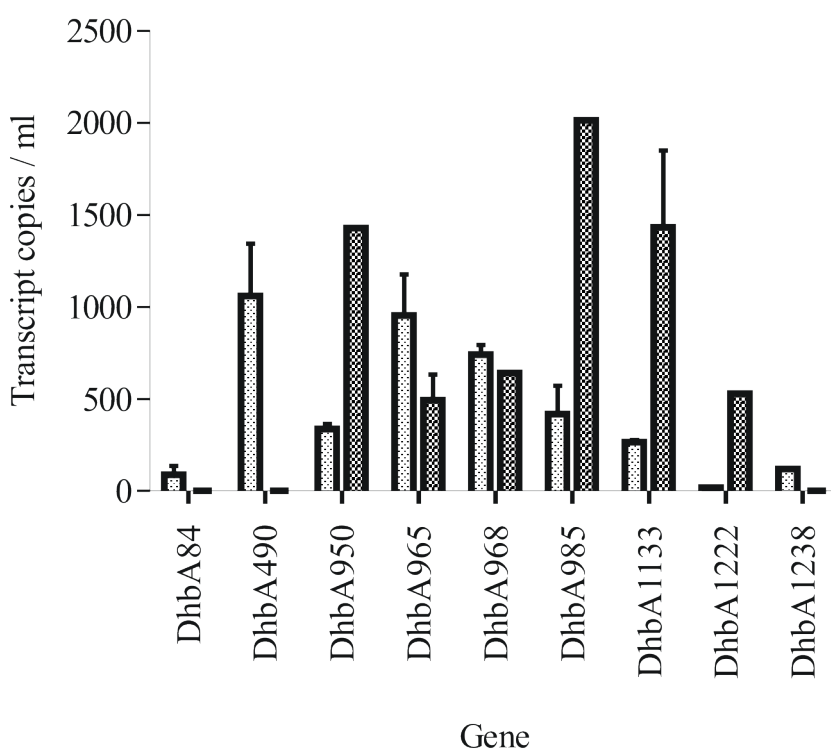
The expression of the different predicted *rdhA* genes in  $\beta$ -HCH grown cocultures was analyzed by RT-qPCR to determine which of the genes could encode the protein responsible for degradation of  $\beta$ -HCH. The expression levels of individual genes were normalized to 16S rRNA transcript copy numbers / mL culture (Fig. 9). Eight of the 10 *rdhA*-like genes were found to be transcribed in the cultures after 11 days of growth. Of these the *pceA*-like genes Dhb490 and Dhb985, and the *cprA*-like Dhb965 and Dhb968 had high expression levels in these cultures. Dhb490 is identical to the *pceA* of *Dehalobacter restrictus*, the product of which has been shown to have a wide dechlorination spectrum in enzyme assays [95]. The expression pattern of *rdhA*-like genes observed after 30 days of incubation was markedly different.

More specifically, there were no transcripts detected for 3 of the genes, which had the highest expression level in 11 day cultures, including Dhb490. It is tempting to speculate that the protein encoded by Dhb490 is involved in the initial dechlorination of the  $\beta$ -HCH and not the breakdown products tetrachlorocyclohexene and dichlorocyclohexadiene. Degradation of  $\beta$ -HCH is known to proceed via two steps of dichloroelimination followed by either a third dichloroelimination to benzene, or a dehydrodehalogenation to monochlorobenzene [22]. Dhb950 and Dhb985 (*pceA*-like), Dhb1133 (*cprA*-like) and Dhb1222 (*rdhA*) have higher transcript levels at 30 days, suggesting that these genes may be involved in dechlorinating the breakdown products tetrachlorocyclohexene and dichlorocyclohexadiene in the degradation pathway. Dhb985 had the highest expression after 30 days suggesting it could be a key gene in the degradation of these breakdown products. The two *cprA*-like genes Dhb965 and Dhb968 were detected at both time points, albeit with lower expression levels at 30 days, suggesting that these genes could be constitutively expressed.

The expression of multiple Rdh enzymes has previously been observed in *Dehalobacter* cocultures [21] and *Dehalococcoides* spp. [38,42,44]. Of specific interest in the coculture studied here is the combination of expressed genes (both *cprA*- and *pceA*-like). Further postgenomic studies will be needed to determine the substrate spectrum of the proteins encoded by these genes, and which of these genes are responsible for the different steps of dichloroelimination and/or dehydrodehalogenation during the sequential dehalogenation of  $\beta$ -HCH [22].



PceA's have been observed as being bifunctional, able to degrade chlorinated ethenes and chlorophenols [42], whereas this has so far not been described for chlorophenol reductive dehalogenases. Although it could not be determined exactly which gene products are responsible for  $\beta$ -HCH degradation, the present results are an essential basis for further genetic analysis of the degradation pathway. Moreover, the draft *Dehalobacter* genome will be of great importance for further transcriptomic and proteomic studies aimed to understand the  $\beta$ -HCH degradation pathway.



**Fig. 9:** Transcriptional analysis of *Dehalobacter* sp. strain E4 reductive dehalogenase genes in cocultures grown in the presence of  $\beta$ -HCH after 11 (dotted bars) and 30 days (filled bars). Coculture batches (500 mL) were used to extract total RNA and. Expressions level were normalized to 16S rRNA transcripts and error bars represent triplicate cultures.

### Role of *Sedimentibacter* sp. strain B4

*Dehalobacter* sp. strain E1 is only able to grow and reductively dechlorinate  $\beta$ -HCH in the presence of *Sedimentibacter* suggesting that there might be essential growth factors supplied by the *Sedimentibacter*. So far its exact role in the coculture has not been unraveled in spite of the fact that *Sedimentibacter* spp. continue to be found in consortia with dechlorinating strains [20,21,99,100]. The draft genome of *Sedimentibacter* sp. strain B4 isolated from the coculture described here is much larger (4.2Mbp) than that of *Dehalobacter* sp. strain E1, and has 3591 predicted protein-encoding genes and 40 tRNA genes. It is possible that *Sedimentibacter* stimulates the transformation of  $\beta$ -HCH via the excretion of growth factors like vitamins, amino acids or other compounds [22]. The *Sedimentibacter* has a higher KEGG proportion of nucleotide metabolism genes compared to OHRB (Fig. 8). It is therefore tempting to speculate that the *Dehalobacter* benefits from the nucleotide biosynthesis by *Sedimentibacter* and its capacity to produce amino acids. Two other categories that have a remarkable gene representation include the biodegradation of xenobiotics and the carbohydrate metabolism. The *Sedimentibacter* may be involved in detoxification of xenobiotics that may hinder growth and activity of the *Dehalobacter* sp. Lastly the carbohydrate metabolism spectrum for *Sedimentibacter* is much wider than that of *Dehalobacter*. In addition, *Sedimentibacter* is able to grow fermentatively on peptone as a substrate [22]. Within its genome there are genes encoding for proteins involved in monosaccharide metabolism and for the fermentative conversion of acetyl-CoA to butyrate and lactate. The products of these reactions may be of benefit to the *Dehalobacter* sp. which does not have the ability of monosaccharide fermentation.

The *Sedimentibacter* genome also revealed a high number of contigs with high G+C content. Their annotation using the metagenome analysis pipeline MG-RAST revealed that 27 % of these contigs harbor genes involved in the metabolism of various carbohydrates, including mono, di- and oligosaccharides. Examples of the di- and oligosaccharides that are predicted to be used include fructooligosaccharides, maltose and maltodextrins and trehalose. Predicted monosaccharides genes include those for arabinose, mannitol, xylose, D-ribose and rhamnose utilization. Lactate and mixed acid fermentation genes were also identified. These short contigs make

over 40 kb of the expected *Sedimentibacter* genome. The high G+C content of these contigs may be indicative of recent horizontal gene transfer but is not a prerequisite for the degradation of the identified carbohydrates.

Lastly, *Sedimentibacter* shows genes of several vitamin biosynthesis pathways including those for folate, lipoic acid riboflavin coenzyme A and cobalamin synthesis. The identified genes for cobalamin (vitamin B12) biosynthesis include those encoding the enzymes CobADUTCS but not those coding for the cobalt transport and insertion, CobG and CobF. Interestingly, CobG and CobF are the variable parts of cobalamin synthesis in prokaryotes and thus are often missing [101]. Moreover, some of the short contigs that have not yet been assembled were found to contain additional genes involved in cobalamin biosynthesis (data not shown). Cobalamin is important for growth of *Dehalobacter* sp. as it is the cofactor of the Rdh's and the growth media are amended with excess amount of this vitamin to achieve appreciable levels of growth in cultures [22]. It is conceivable that it is the capacity of *Sedimentibacter* to continuously produce cobalamin in moderate concentrations that stimulates growth and dechlorination of the *Dehalobacter* in the coculture. *Dehalococcoides* spp. have also been found to have higher growth rates and dechlorination activity when grown in media with excess vitamin B12 or in consortia with cobalamin-producing strains [102]. The above metabolic features are the first indications that point to a role of *Sedimentibacter* as syntrophic partner in the coculture. Additional efforts towards refined annotation and pathway analysis will be required to provide a more complete picture of the genomes and its functional implications for bioremediation

## Conclusion

The draft genome assembly of *Dehalobacter* sp. strain E1 comprises 2.6 Mbp of sequence while that of *Sedimentibacter* sp. strain B4 is 4.2 Mbp. Within the *Dehalobacter* genome are 10 putative Rdh-encoding gene clusters, the products of which show homology to known chlorophenol reductases and tetrachloroethene reductive dehalogenases. Among these Rdh operons is the *pceABCT* operon described in *Dehalobacter restrictus* PER-K23 and *Desulfotobacterium hafniense* strains TCE1 and PCE-S. Based on RT-qPCR the *pceA* gene (Dhb490) of this operon, Dhb985

(pceA-like gene), as well as Dhb965 and Dhb968 cprA-like genes are suggested to be involved in the initial stages of the dechlorination of  $\beta$ -HCH. We also observed that the coculture was capable of dechlorinating TCE but did not dechlorinate cis-DCE or any of the chlorophenols used in growth experiments. The availability of the first draft genome sequences of strains of Dehalobacter and Sedimentibacter provides an opportunity to investigate complementary gene expression in order to understand the mechanistic basis of the syntrophy of these two bacteria that form, a partnership dedicated to organohalide respiration. Future proteomics studies may be done in combination with transcriptomics to examine the molecular interactions involved in this obligate relationship. The genomic characterization presented here provides the basis for applying the ecogenomics toolbox [68] to elucidate the physiology of the  $\beta$ -HCH degrading coculture and to provide necessary (kinetic) data towards a metagenome-scale metabolic model of this syntrophic partnership.

### Acknowledgements

This work was supported by the Ecogenomics Project of the Netherlands Genomics Initiative. MWJvP is funded by the Netherlands Organization for Scientific Research (NWO) via a VENI grant.

# CHAPTER 3

Chapter 3

## **Differential Expression of *Desulfitobacterium hafniense* DCB-2 Chlorophenol Reductive Dehalogenase Genes in Response to *meta*- or *ortho*-Chlorophenols**

Farai Maphosa, Krisztina Gabor, Mark H.J. Sturme, Tian Ye, Willem M. de Vos and Hauke Smidt

*Manuscript in preparation*

**ABSTRACT**

Although *Desulfitobacterium hafniense* DCB-2 is able to dehalogenate a diverse range of chloroaromatic compounds, little is known about the expression of the several chlorophenol reductive dehalogenase genes (*cprA*) its genome harbors and their specific responses to the presence of either *ortho*- or *meta*-chlorophenols. The *D. hafniense* DCB-2 genome sequence contains seven *cprA*-like genes, five of which (*cprA1-cprA5*) are predicted to encode putative proteins that share 18-91% amino acid sequence identity with the *ortho*-chlorophenol reductive dehalogenase (CprA) from *Desulfitobacterium dehalogenans*. We investigated the transcriptional patterns for *cprA1* to *cprA5* using as substrates 2,3-DCP for the *meta*-dechlorination and 2,4-DCP for the *ortho*-dechlorination positions. Clear differences in the induction patterns were observed. In the presence of 2,4-DCP, *cprA1* and *cprA5* were strongly induced, while 2,3-DCP induced *cprA3* and *cprA5*. Our results suggest that the gene product of *cprA5* has a broad range substrate specificity while those for *cprA1* and *cprA3* have specific activity towards *ortho*- and *meta*-chlorophenols, respectively. The genes *cprA2* and *cprA4* for which transcription had been reported to be induced by both *ortho*- and *meta*-chlorophenols were not induced by these substrates. The results indicate that several different regulatory proteins might be involved in the transcriptional regulation and expression of the five *cprAs*, resulting in their complementary specificities and responses to *meta*- and *ortho*- chlorophenols.

## Introduction

Organohalide respiration is a process catalyzed by a broad diversity of anaerobic bacteria, and the low-G+C Gram-positive genus *Desulfitobacterium* comprises a major group of isolates capable of reductive dechlorination of various chlorinated hydrocarbons, including chlorinated phenols and/or alkenes [103]. Reductive dehalogenation associated with organohalide respiration is catalyzed by reductive dehalogenases (Rdh) and the organohalide respiratory system has been well characterized in *Desulfitobacterium dehalogenans* [98].

The *D. dehalogenans* genome contains a 11.5-kb chlorophenol reduction (*cpr*) gene cluster, harboring eight genes designated *cprTKZEBACD*, that are involved in expression and regulation of organohalide respiration [39,98]. The cluster includes genes encoding the reductive dehalogenase (*cprA*), its membrane anchor (*cprB*), several proposed chaperonins (*cprT*, *-E*, and *-D*), a putative membrane-bound regulatory protein (*cprC*) and transcriptional activator (*cprK*) that are organized into four transcriptional units: *cprT*, *cprBA*, *cprZE*, and *cprBACD* [98]. The *cprTKZEBACD* gene cluster was also observed in *Desulfitobacterium hafniense* DCB-2, suggesting a similar gene regulation [35]. Five predicted Rdh's, designated CprA1 to CprA5, are encoded in the genome of *D. hafniense* DCB-2. Nevertheless, only the *cprA1* gene of the above-mentioned *cprTKZEBACD* gene cluster has been characterized and confirmed to encode a functional Rdh, with *ortho*-dechlorination activity towards 3-chloro-4-hydroxyphenylacetic acid (CHPA) [35,104]. The genes *cprA2-cprA5* are not yet functionally characterized but they contain all the conserved regions characteristic of Rdh's: a twin arginine signal motif and two iron-sulfur cluster centres [15,98]. The *cprA2-cprA5* genes have also been identified in *D. hafniense* strains PCP-1 and TCP-A, and shown to be expressed in the presence of chlorophenols [105]. Transcription of *cprA5* was only detected when strains PCP-1, TCP-A and DCB-2 were cultured in the presence of 2,4,6-trichlorophenol or 3,5-dichlorophenol (3,5-DCP), but not in the absence of chlorophenol, showing that *cprA5* is inducible [105].

Transcriptional induction by a specific organohalide was also observed for the *cprTKZEBACD* locus of *D. dehalogenans*; where bicistronic transcription of *cprBA* was induced 15-fold upon addition of the *ortho*-chlorophenolic substrate CHPA [98].

Chlorophenylacetate, which lacks the *o*-hydroxy group, and hydroxyphenylacetate (lacking the chlorine group) did not activate transcription of *cprA* [39]. *D. hafniense* DCB2 shows the highest diversity of proteins belonging to the CRP-FNR (cAMP-binding protein / fumerate nitrate reduction regulatory protein) family of transcriptional regulators encoded in a single organism [39,98,106]. The CRP-FNR-type regulator CprK1 is responsible for activating transcription of *cprA1* in strain DCB-2 [40]. Binding of CHPA to CprK1 was shown to result in an active DNA-binding conformation, which enables the regulator to activate transcription from promoters that contain a specific DNA sequence termed the dehalobox (TTAAT-N4-ATTAA). Four other CprK-encoding genes (*cprK2-cprK5*) were also identified in *D. hafniense* DCB-2 that cluster with genes encoding potential organohalide respiration proteins [35,39]. CprK1 and CprK2 have a partially overlapping effector specificity, with preference for *ortho*-chlorophenols (CHPA or 2,4-DCP), while *meta*-chlorophenols (like 2,3-DCP) proved to be effectors for CprK4 [36]. These regulators interacted *in vivo* and *in vitro* with *cprA* promoters and also with a promoter of *macA*, a putative methyl-accepting chemotaxis protein-encoding gene. The ability of *D. hafniense* DCB-2 to respire with a wide diversity of halogenated compounds as terminal electron acceptors could be due to this unusually high occurrence of CprK like regulators, where each regulator corresponds to a specific group of halogenated compounds [36,107]. To further expand our understanding of differential *cprA* gene expression and their specific responses in the presence of different chlorophenols, we investigated the induction of the full set of *cprA* paralogues of *D. hafniense* DCB-2 in the presence of two different effector compounds; an *ortho*- and *meta*-chlorophenol.

## Methods

### Bacterial strains and growth conditions

*D. hafniense* strain DCB2 was cultured in a phosphate-bicarbonate buffered medium with a low chloride concentration [108] under N<sub>2</sub>/CO<sub>2</sub> (80:20) headspace at 37 °C. Anaerobic stock solutions (5mM) of 2,3-DCP and 2,4-DCP and 10 mM of CHPA were prepared in water, pH was adjusted to 7.0 by the addition of 1N NaOH, and



filter sterilized. CHPA was obtained from Sigma-Aldrich (The Netherlands) and dichlorophenols from Janssen Chimica (Germany).

### DNA isolation

Chromosomal DNA of *D. hafniense* DCB-2 was isolated by first mechanically disrupting the cells by bead-beating followed by purification using the Bio101 FASTDNA isolation kit according to manufacturers' protocols (MP Biomedicals, Solon, OH, USA).

### RNA isolation

For gene expression analysis *D. hafniense* DCB-2 cells were grown fermentatively in medium containing 40 mM pyruvate as carbon source. When cells reached mid-exponential growth phase ( $OD_{600} \sim 0.1$ , in 3 days), 2,3-DCP or 2,4-DCP was added to the cultures to a final concentration of 40  $\mu$ M. Experiments were performed in duplicate. RNA was extracted from 80 to 100 ml culture at the time of the induction, and after 4 and 44 hours, using the Macaloid method [89]. In short: cells were collected by centrifugation ( $13,000 \times g$ , 4 °C) and subsequently resuspended in 0.5 ml TE (pH 7). The cell suspension was transferred to a bead-beating tube containing 0.6 g glass beads, 0.17 ml 4% macaloid [89], 0.5 ml acid phenol (pH  $\sim 3.5$ ) and 0.05 ml 10% SDS. All aqueous solutions used for RNA isolation and downstream analyses were prepared with diethylpyrocarbonate (DEPC) treated water. Cells were disrupted by  $3 \times 1$  min of bead-beating, with 1 min cooling intervals on ice. Cell debris were separated from the RNA fraction by centrifugation ( $13,000 \times g$ , 4 °C, 15 min). The supernatant was extracted with an equal volume of phenol:chloroform:isoamylalcohol (25:24:1) using an Eppendorf Phase-Lock Heavy gel (Eppendorf AG, Hamburg, Germany), followed by extraction with an equal volume of chloroform:isoamyl alcohol (24:1). RNA was precipitated by the addition of 1/10 volume 3 M Na-acetate (pH 5.2) and 1 volume of isopropanol and incubation at -80°C for 2 hours. RNA was washed with 70% ethanol and resuspended in 50  $\mu$ l water.

Co-purified traces of genomic DNA were removed by two consecutive DNase I treatments. First, RNA samples were applied to an RNeasy column (Qiagen, Venlo, The Netherlands), and an on-column DNase I treatment was performed according

to supplier protocols. Secondly, the eluted RNA samples were treated with DNase I. using Promega RQ1 DNase mix (Promega, Leiden, The Netherlands). Total nucleic acid concentrations and purity were estimated using absorbance readings (260 nm/280 nm) on a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE).

Synthesis of cDNA was performed on 500 ng total RNA (double DNaseI-treated), using random hexamers and the SuperScript™ III First-Strand Synthesis System (Invitrogen Life Technologies, Breda, The Netherlands), according to the manufacturer's instructions. Removal of contaminating genomic DNA was confirmed by PCR on non-reverse transcribed RNA samples with primers Dsb406F and Dsb619R [48] targeting the *Desulfitobacterium* 16S rRNA gene.

### Quantitative PCR assays

Gene expression was measured for 6 target genes (*cprA1-cprA5* and *macA*) using the iQ5 iCycler (BioRad). Quantitative PCR (qPCR) was performed on 25 ng of cDNA sample using the iQ SYBR Green Supermix kit (BioRad) in triplicate 25 µl reactions including No-Template Controls. Primers Dsb406F and Dsb619R [48] were selected for qPCR of the 16S rRNA gene of *Desulfitobacterium*. Primers for qPCR of Rdh genes *cprA1* to *cprA5* and *macA* (Table 1) were designed using Primer3 software [90] (<http://frodo.wi.mit.edu/primer3/input.htm>). Primers were tested *in silico* using OligoAnalyzer 3.1 (<http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/>). Primer specificity and performance were tested by temperature-gradient PCRs on the iQ5 iCycler, using plasmid DNA containing target gene inserts as PCR-standards. Primers were used at a final concentration of 200 nM for SYBR Green assays. Quantitative PCR amplification parameters for the 16S rRNA-targeted assay were: 10 min at 95 °C, followed by 40 amplification cycles of 15 s at 95 °C, 30 s at 60 °C, 30 s at 72 °C. Parameters for amplification of other target genes were the same as described above, expect for the annealing temperatures: *cprA1*, *cprA3* and *cprA4* at 66 °C; *cprA2* at 63 °C; *cprA5* and *macA* at 65 °C. Melt curve analysis was performed from T<sub>annealing</sub> to 95 °C in steps of 0.5 °C and 10 s at each step. The transcript levels were normalized to the 16S rRNA transcripts.

### Bioinformatic analysis

The genome sequence of *D. hafniense* DCB-2 was screened for organohalide respiration genes using the Integrated Microbial Genome tool (Joint Genome Institute) [109] and basic local alignment search tool (BLAST) [110]. Accession numbers of proteins predicted to be encoded by *cprA* genes (based on *D. hafniense* genome sequence NC\_011830) are shown in Table 2. Multiple sequence alignment was produced by ClustalW [85].

### Statistical analysis

The significance of the observed induction levels was verified using the Student's *t*-test.

**Table 1:** *D. hafniense* DCB-2 gene-targeted real-time PCR primers used in this study

Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')
<i>cprA1</i>	CAC GTA CCG GAG ACA GTG C	AAG GAA ATC GCT TCG TTG G
<i>cprA2</i>	TCG GGA ACA TTG AAA AGA CC	TGA TTT CCT GAG TGG GAA GC
<i>cprA3</i>	GAA ACT GCC CAT ACC TGA CC	CGC CTT GTG GGA ATA GTA GG
<i>cprA4</i>	GCA ACG ATG AAG GAG TCA GC	ACA ATC TCC GTG CCA TAA CC
<i>cprA5</i>	AGG TGG AAG TTT ACG GAT CG	ATC CAT CTC GAA GGT CAT GG
<i>macA</i>	CAA TCA TCG TCT TCC ACA GC	AAC CTG ACC GTT TCA ACA CC

## Results

### Genome analysis

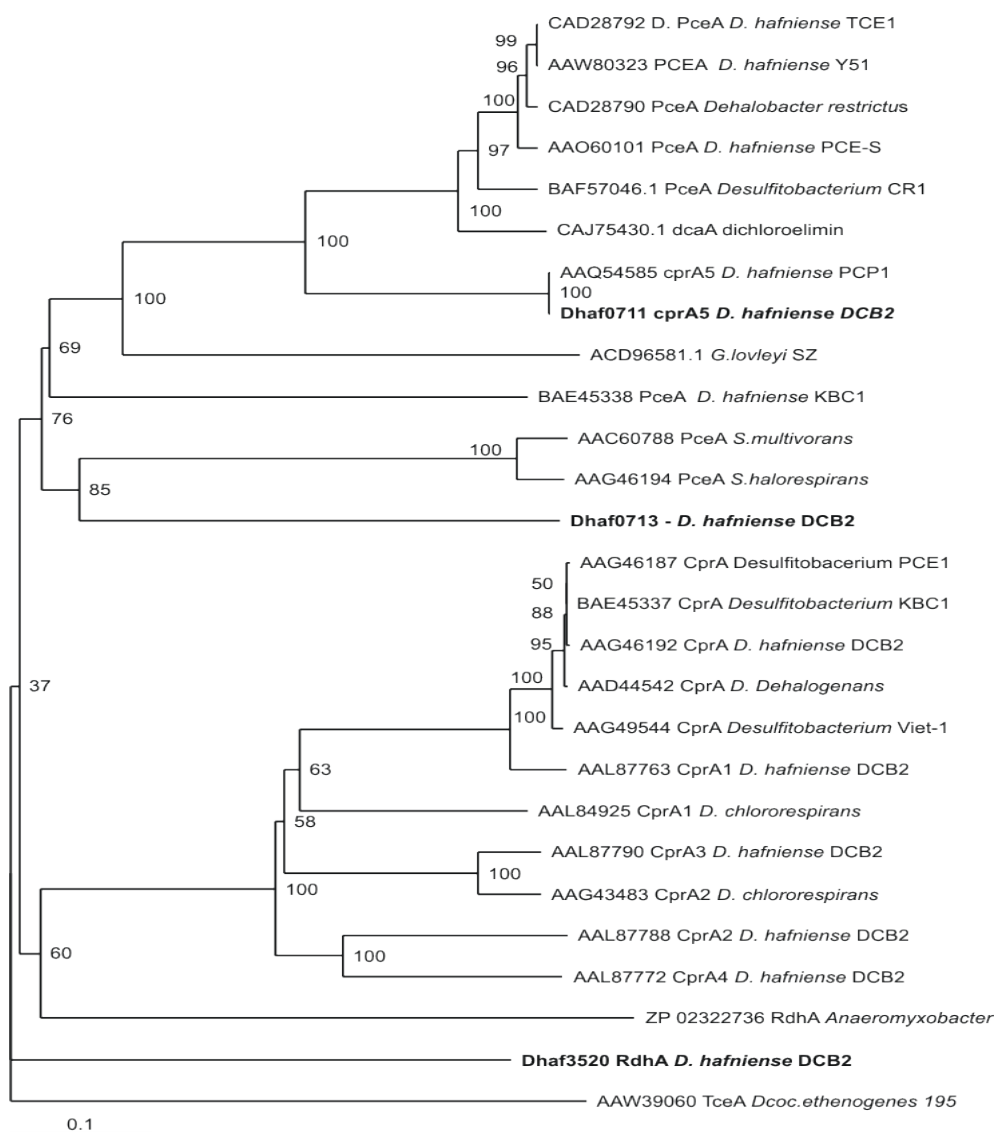
Analysis of the complete genome sequence of *D. hafniense* DCB-2 (NC\_011830) revealed seven chlorophenol reductive dehalogenase genes, five of which are predicted to be potentially functional. These five *cprA* genes (*cprA1-cprA5*) have an average G+C content of 52.4%, and average length of the encoded proteins is 481 amino acids with 18-91% amino acid sequence identity with the *ortho*-chlorophenol reductive dehalogenase CprA from *D. dehalogenans* (Table 2). *CprA1* to *cprA5* are located within a 48 kb chromosomal region of the *D. hafniense* DCB-2 genome, together with several clustering genes similar to those present in the *cprTKZEBACD* cluster of *D. dehalogenans*. The *cprBA* gene arrangement is the same for all the clusters except for *cprA5*, which has the hydrophobic membrane anchor-encoding gene *cprB* positioned downstream. Unlike *cprA* of *D. dehalogenans* and DCB-2 *cprA1-A4*, the *cprA5* sequence is on average 250 bp longer and encodes a protein that is 60 amino acids (or more) larger. CprA5 is only 18.3% identical to the CprA of *D. dehalogenans* while CprA1-A4 are over 50% similar to *D. dehalogenans* CprA. Phylogenetic analysis shows that CprA5 clusters with the CprA5 of strain PCP-1 and resembles tetrachloroethene reductive dehalogenases rather than the characterized chlorophenol reductive dehalogenases (Fig. 1).

### Induction of *cprA* transcription by chlorophenols

In order to test for specific gene induction by different substrates, we tested the chlorophenols 2,3-DCP (*meta*) and 2,4-DCP (*ortho*) for their *cprA*-inducing potential. Transcript levels of *cprA1 – cprA5* and the *macA* gene were examined during the exponential growth phase of *D. hafniense* DCB-2 cultures in the presence of either putative inducer, and normalized to 16S rRNA gene expression levels. Gene-specific primer sets were designed to quantify transcript levels for each of these genes and qPCR was performed on cDNA produced by reverse transcription of total RNA isolated from cells at 3 time points; 0, 4 and 44 h after addition of 2,3-DCP or 2,4-DCP. Clear differences in the expression patterns were observed for the 6 genes in the presence of each chlorophenol (Fig. 2). Normalized transcript levels of *cprA1* did not increase significantly ( $P>0.05$ ) in the presence of 2,3-DCP even after 44 h. As

**Table 2.** Gene and amino acid properties of the putative chlorophenol reductive dehalogenases of *D. hafniense* DCB-2.

	<i>cprA1</i>	<i>cprA3</i>	<i>cprA2</i>	<i>cprA5</i>	<i>cprA4</i>
Locus Tag / Accession	Dhaf_0737 / ZP_01372867	Dhaf_0713 / ZP_01372889	Dhaf_0711 / ZP_01372891	Dhaf_0696 / ZP_01372906	Dhaf_0689 / ZP_01372912
Gene size (bp)	1338	1377	1467	1653	1392
GC Content %	52	51	54	54	51
Amino acid sequence length (aa)	445	458	488	550	463
Signal Peptide	Yes	Yes	No	Yes	Yes
Identity to CprA of <i>D. dehalogenans</i> (% aa)	91.2	59.6	53.2	18.3	55.1

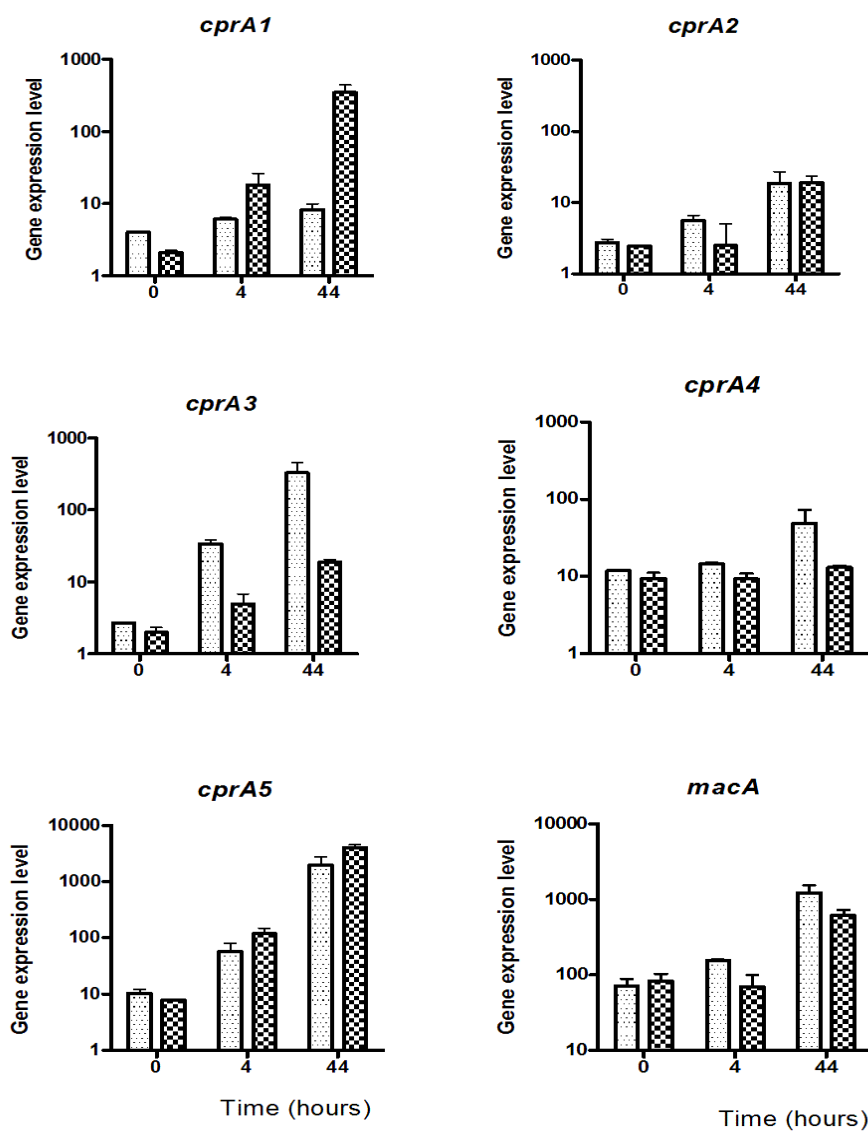


**Fig. 1:** Phylogenetic analysis of the predicted amino acid sequences of characterized and putative reductive dehalogenases of *Desulfotobacterium*. Also included are protein sequences from *Dehalobacter*, *Sulfurospirillum*, *Geobacter* and *Anaeromyxobacter*. The *Dehalococcoides ethenogenes* TceA amino acid sequence was used as an outgroup. Amino acid sequences were aligned by ClustalW. Genebank accession numbers are shown for each protein, the Rdh's of *D. hafniense* DCB2 are shown in bold. The scale bar represents 0.1 changes per amino acid position. 100 Bootstrap values are shown.

expected this gene was significantly induced by the *ortho*-chlorophenol 2,4-DCP and the normalized transcription levels increased from  $2.1 \pm 0.2$  at the time of induction to  $18 \pm 11.8$  and  $349 \pm 132.9$  transcript copies /ml at 4 and 44 h respectively. The *cprA2* gene did not show any significant increase in transcription in the presence of either 2,3-DCP or 2,4-DCP after 4 or 44 h. Induction of transcription of *cprA3* by 2,3-DCP was significant ( $P < 0.001$ ), increasing from 2.7 copies to  $33.9 \pm 6.4$  and  $330.0 \pm 181.0$  transcript copies / ml at 4 and 44 h respectively. No significant induction, however, was observed in the presence of the *ortho*-chlorophenol. *CprA4* was not induced by either compound and had a constant average expression level of  $9.6 \pm 1.2$ . Highest expression was observed for *cprA5* especially after 44 h in the presence of either 2,3-DCP or 2,4-DCP. For cells grown on 2,3-DCP the normalized expression levels were  $56.4 \pm 32.5$  and  $1975.0 \pm 1053.6$  after 4 and 44 h, respectively, and the increases compared to time of induction were significant at  $P < 0.001$ . The expression level was higher than for *cprA3*, which was also induced by 2,3-DCP. In the presence of 2,4-DCP *cprA5* was induced after 4 and 44 h to a normalized expression level of  $119.2 \pm 42.1$  and  $4050.0 \pm 650.5$ , respectively. Expression of *cprA5* was 7 and 12 fold higher than that of *cprA1* at 4 and 44 h, respectively.

The CprK regulators have been previously shown to interact *in vivo* and *in vitro* with *cprA* promoters and also with a promoter of *macA*, a putative methyl-accepting chemotaxis protein encoding gene, and therefore the induction of this gene was also examined in the presence of chlorophenols. The *macA* gene expression level was several fold higher than those observed for any of the *cprA* transcripts at the initial time point (Fig. 2), being 23 – 38 fold higher than *cprA1-cprA3* and 6-9 fold higher than *cprA4-cprA5*. There was a significant increase after 44 h in the presence of either compound, to an expression level of  $1223.5 \pm 461.7$  for 2,3-DCP and to  $615.0 \pm 158.4$  in the presence of 2,4-DCP.

Quantification of the overall gene expression, as a sum of the individual gene expression levels after 4 and 44 h in the presence of either compound were determined. Higher expression levels were observed in the presence of 2,4-DCP being 1.5 and 2 fold higher after 4 and 44 h respectively than in the presence of 2,3-DCP.



**Fig. 2:** Transcript levels for the *cprA1-cprA5* and *macA* genes of *D. hafniense* DCB-2. Gene expression level (gene copies /ml) is expressed normalized to 16S rRNA at point of induction, 4 and 44 h after addition of either 2,3-DCP (dotted bars) or 2,4-DCP (black square bars). Error bars represent standard deviations of duplicates.



## Discussion

*D. hafniense* DCB2 can dechlorinate several chlorophenols at the *ortho*- and *meta*-positions [107,111]. In our study, we observed higher cumulative reductive dehalogenase transcript levels in the presence of the *ortho*-chlorophenol 2,4-DCP compared to the *meta*-chlorophenol 2,3-DCP. This could explain the observed preference for *ortho*-chlorophenols in *D. hafniense* DCB-2 cell extract activity tests [105].

Transcription of *cprA* is regulated by the CRP-FNR-type regulator CprK in *D. dehalogenans*, and it has been shown that the corresponding gene in *D. hafniense* DCB2, *cprA1*, is activated by CprK1 in the presence of *ortho*-chlorinated compounds [40,98]. *In vitro* DNA-binding experiments have also demonstrated that CprK1 forms an active DNA-binding complex only with *ortho*-substituted compounds (2,4-DCP, CHPA) and not with *meta*-substituted compounds (2,3-DCP or 3,5-DCP) [36,112]. Our gene expression data support these observations: *cprA1* transcription was only induced by the *ortho*-chlorophenol (2,4-DCP) and not by the *meta*-chlorophenol (2,3-DCP). Other studies showed that *cprA1* is also induced by the *ortho*-chlorophenol CHPA [30,104], confirming that it encodes indeed an *ortho*-chlorophenol reductive dehalogenase. In comparison, *cprA* of *D. dehalogenans*, the gene most similar to the DCB2 *cprA1* gene, is also highly specific to *ortho*-chlorinated aromatic substrates [39,112,113].

Two genes, *cprA2* and *cprA4*, did not respond significantly to either 2,3-DCP or 2,4-DCP suggesting that they are induced by other (halogenated) effectors. The *cprA2* and *cprA4* genes are also phylogenetically strongly related (95% nucleotide sequence identity) and might have similar specificities for chlorophenols that were not tested here. These effectors remain to be elucidated in future experiments. Previously, RT-PCR experiments in strains PCP-1, TCP-A and DCB-2 had shown that *cprA2*, *cprA3* and *cprA4* were transcribed both in the absence and presence of 2,4,6-TCP (*ortho*-dechlorination position) and 3,5-DCP (*meta*-dechlorination position) [105]. In this experiment we did not find any significant induction of *cprA2* and *cprA4*, however, our results suggest that *cprA4* is constitutively expressed. Two genes, *cprA4* and *cprA5*, had 4- 7 times higher transcript levels at the point of induction compared to the other *cprA* paralogues in the presence of either 2,3-

DCP or 2,4-DCP suggesting transcripts are more abundant in the cell than transcripts of *cprA1- cprA3* even in the absence of the effectors. This is in agreement with observations made by in-vivo promoter probe assays which showed that a strong inducible promoter upstream of the *macA-cprKBA4* gene cluster appeared to be constitutively active [36]. The *macA* gene is located upstream of the *cprKBA4* gene cluster and had previously been shown to be up-regulated in the presence of different *ortho*- and *meta*-substituted chlorophenols. It was hypothesized that transcription of *cprK4* and *cprBA4* could be initiated from *macA* by readthrough [36]. Since *cprA5B* is located approximately 10kb downstream of the *macA* promoter, constitutive readthrough transcription of the *cprA5* cluster from this promoter can at this point not be excluded. We measured expression of *macA*, and unlike the *cprAs* it was highly expressed even at the initial time point, and transcripts increased further with time confirming its induction under organohalide respiring conditions. Transcript levels of *macA* were furthermore two fold higher in the presence of 2,3-DCP than with 2,4-DCP, suggesting a stronger response to the *meta*-substituted chlorophenol.

Only *cprA3* was significantly induced in the presence of the *meta*-chlorophenol 2,3-DCP but not by the *ortho*-chlorophenol 2,4-DCP. The *cprA3* gene clusters together with two putative regulator genes: *cprK2* and *cprK3*. CprK2 was found to be activated by an *ortho*-chlorophenol (CHPA), but did not activate transcription from the *cprA3* promoter [36]. The unsuccessful attempt to express *cprK3* heterologously prevented further transcriptional studies but it is possible that this regulator senses 2,3-DCP and induces expression of the *cprA3* gene in the presence of this compound.

Interestingly *cprA5* expression was induced both in the presence of the *ortho*- and the *meta*-chlorophenol suggesting that the encoded CprA5 enzyme has the ability to dechlorinate at both *meta*- and *ortho*- substituent positions as was observed for the closely related CprA5 of strain PCP-1 [105]. Transcript levels of *cprA5* were higher in the presence of *ortho*-chlorophenol 2,4-DCP than the *meta*-chlorophenol 2,3-DCP. The induction by both compounds may be due to the different gene arrangement observed for *cprA5* compared to the other *cprAs*. Firstly all known *cprA* encoding genes are linked to a gene encoding for a small hydrophobic protein (*cprB*)

of around 100 amino acids which has three membrane-spanning helices [35,114]. The predicted hydrophobic membrane anchor encoding gene *cprB5* is positioned downstream of *cprA5* similar to that of the 3,5-DCP reductive dehalogenase gene cluster of strain PCP-1, as also observed for some tetrachloroethene reductive dehalogenases [34]. Since the strong organohalide respiration-inducible promoter is located directly upstream of the reductive dehalogenase gene, the transcription of *cprA5* might be more efficient, than when transcription starts upstream of *cprB* and proceeds further to the *cprA* gene via read-through. Secondly it is likely that the CprK-type regulator that switches on expression of *cprA5* has a broad effector range and is activated both by *meta*- and *ortho*-chlorinated compounds. A candidate for this role is CprK4, a regulator that forms specific protein-DNA complexes in the presence of both 2,3-DCP and 2,4-DCP [36]. Thirdly we hypothesize that the promoter region of *cprA5* probably contains a less stringent CprK-binding site (dehalobox) allowing binding of either a 2,3-DCP activated CprK-homologue or a 2,4-DCP activated CprK-homologue. This is supported by data by Gábor et al. [36], who reported that the promoter of *cprA5* contains a putative dehalobox (ATACG-N4- AATAA), which is only distantly related to the dehalobox consensus (TTAGT-N4-ACTAA). Unfortunately, efforts to determine, which of the regulators activates transcription of *cprA5* were unsuccessful as all attempts to obtain stable expression constructs with the dehalobox of the *cprA5* promoter region failed.

The *cprA5* gene sequence is on average 250 bp (60 amino acids) longer than that of the other *cprAs* of *D. hafniense* DCB2 or *D. dehalogenans*. Possibly these additional amino acids confer structural properties that give the ability to dechlorinate at both *meta*- and *ortho*- positions. There is currently no protein structure available for any of the reductive dehalogenases from organohalide respiring bacteria, including CprA5, however, it would be interesting to see in the future whether its larger size plus the base difference located at nucleotide 818 (C for T), resulting in an amino acid substitution at position 273 (proline in strain DCB-2 for a lysine in strain PCP-1) [34] has any influence on the ability to dechlorinate chlorophenols with both *ortho*- and *meta*- substituents.

In conclusion, we investigated by quantitative reverse transcriptase PCR, which chlorinated phenols could induce the five *cprA* paralogues of *Desulfitobacterium hafniense* strain DCB-2, using as representative substrates 2,3-DCP for the *meta*-dechlorination position and 2,4-DCP for the *ortho*-dechlorination position. We observed clear differences in the induction patterns, with 2,4-DCP strongly inducing *cprA1* and *cprA5*, while 2,3-DCP strongly induces *cprA3* and *cprA5*. From previous studies we suggested that the regulators that are responsible for the transcriptional activation are CprK1 (for the *ortho*-chlorinated compounds) and CprK4 (for the *meta*-chlorinated compound). While the role of CprK1 in the transcriptional activation of *cprA1* has been already confirmed, [36] further studies need to be conducted to confirm the role of these two regulators on the activation of *cprA3* and *cprA5*. Most likely, additional regulatory proteins, including the yet uncharacterized CprK3, are involved in the transcriptional regulation and induction of the transcription of *cprAs* resulting in differential specificities and responses to *meta*- and *ortho*- chlorophenols.

### Acknowledgements

We would like to acknowledge support by the Netherlands Genomics Initiative (NGI) Innovative Cluster Ecogenomics (BSIK03011), as well as the European Commission through the BACSIN (FP7-KBBE-211684) collaborative project of the seventh framework program.

# CHAPTER 4

## **Microbial community- and metabolite dynamics of an anoxic dechlorinating bioreactor**

### Chapter 4

Farai Maphosa, Hauke Smidt, Willem M. de Vos, Wilfred F.M. Röling

*A modified version of this chapter is under review for Environmental Science and Technology*

**ABSTRACT**

Monitoring and quantification of organohalide respiring bacteria is essential for optimization of in-situ bioremediation of anoxic subsurface sites contaminated with chloroethenes. Molecular monitoring and model simulations were applied to determine degradation performance of an in-situ dechlorinating bioreactor and its influence on the contamination plume. *Dehalococcoides* was the dominant dechlorinating microorganism as revealed by qPCR targeting 16S rRNA- and chloroethene reductive dehalogenase-encoding genes (*tceA*, *vcrA*, *bvcA*). The presence of all three reductive dehalogenases genes indicated co-existence of several distinct *Dehalococcoides* populations in the bioreactor and groundwater. Mass balancing revealed that main dechlorinating activities were reduction of *cis*-dichloroethene and vinyl chloride. Saturation analysis showed that while performance of the bioreactor improved due to especially the addition of molasses, dechlorinating microorganisms were fast-growing. Once near-complete dehalogenation was achieved, *Dehalococcoides* only grew slowly and did not further increase in population density. The bioreactor influenced dechlorinating populations in the plume with subsequent decrease in chlorinated compound concentrations over time. In the present study, a combination of molecular diagnostics with mass-balancing and kinetic modeling allowed for detailed insight into the complex interplay of organohalide respiring bacteria and metabolite dynamics in the *in-situ* dechlorinating bioreactor, giving crucial information for optimizing bioremediation.

## Introduction

Bioremediation of sites contaminated with chlorinated ethenes, such as commonly used industrial solvents tetrachloroethene (PCE) and trichloroethene (TCE) [115,116], makes use of organohalide respiring bacteria (OHRB) that can anaerobically reduce these toxic compounds to harmless end-products like ethene. Key OHRB belong to the genus “*Dehalococcoides*” [117-121], which reductively dechlorinate PCE, TCE, *cis*-dichloroethene (cDCE), and vinyl chloride (VC) to ethene. Also important are microorganisms from other genera including *Desulfotobacterium*, *Dehalobacter*, *Geobacter* and *Sulfurospirillum*, which are able to dechlorinate PCE to cDCE. Reductive dechlorination of PCE past cDCE, however, has so far been linked exclusively to *Dehalococcoides* spp. [15,27,33,57,122]. In general, reduction of chloroethenes by OHRB is mediated by reductive dehalogenase (Rdh) enzymes. Rdh-encoding genes have emerged as appropriate diagnostic targets in studying and monitoring bioremediation of contaminated sites [16,46,49,57,122-124].

Quantitative PCR protocols have been developed for detection and monitoring of *Dehalococcoides* populations and their Rdh-encoding genes [124]. Specific targets are *tceA*, encoding the Rdh catalyzing stepwise reductive dehalogenation of TCE to VC in strains 195 [125] and FL2 [119,126], *vcrA*, coding for the DCE- and VC- to ethene reducing dehalogenase of strains VS [33] and GT [127], and *bvcA* that encodes the VC Rdh of strain BAV1 [32].

Continuing efforts are made in evaluating dechlorination potential of sites and monitoring their *in-situ* bioremediation based on the presence of *Dehalococcoides* populations, and the abundance and expression of genes that encode Rdh's associated with the dechlorination of cDCE to VC and ethene [46,49,50]. Information on growth kinetics and dynamics of *Dehalococcoides* spp., however, has so far only come from laboratory studies [128,129]. Therefore, the purpose of this study was to (i) describe and quantify key dechlorinating populations using genes that code for 16S rRNA and complementary Rdh's (*tceA*, *bvcA* and *vcrA*) for an on-site dechlorinating anaerobic bioreactor installed for the *in-situ* bioremediation of a chloroethene-contaminated site, (ii) use mass-balancing and kinetic modeling to quantify the dechlorinating activities of the bioreactor, and (iii) integrate molecular data with mathematical analysis to determine bioreactor performance and degradation dynamics of *Dehalococcoides*

populations.

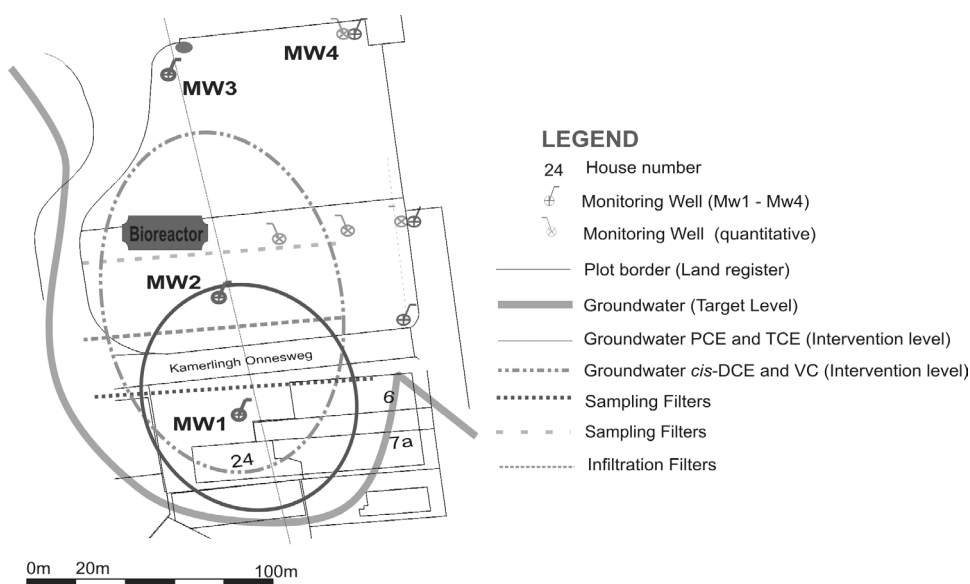
## Methods

### Site description

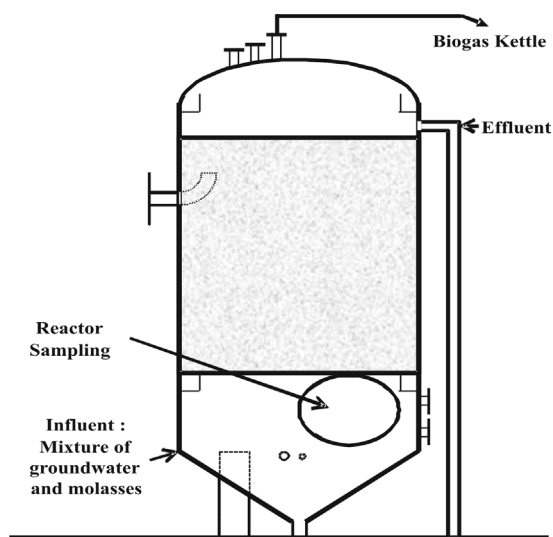
The study site, located in Ede, The Netherlands, is undergoing *in-situ* bioremediation of a contaminated plume of chlorinated ethenes using an anaerobic bioreactor (Fig. 1-2). The site is largely covered by an asphalt layer for a car parking lot. The underlying soil consists of alternate 2 m layers of moderate to extremely coarse sand and peat up to 9 m depth and then fine sandy and clay soil to approximately 35 m below ground level. The contamination plume has south westerly direction along with the regional groundwater flow, was about 300 m long and 75 m wide and was estimated to migrate at about 2m/year. Distribution of the chlorinated solvents (PCE and TCE) at this site has been studied since 1996 and the source of pollution was attributed to dry cleaning and automobile operations that once took place at the site. PCE, TCE, cDCE and VC were present in ground water at concentrations of 400, 8800, 27000 and 150 µg/L respectively. Mineral oils were also observed in the plume at concentrations above 5000 µg/L.

An anaerobic dechlorinating bioreactor has been installed on site for in-situ bioremediation in December 2005. Groundwater is pumped into the bioreactor, and effluent containing a dechlorinating community is discharged back into the plume (Fig. 1). Inflow rate to the 7 m<sup>3</sup> bioreactor is 400 l/h and residence time is about 17.5 hours. Molasses (138 g/L) was dosed into the influent stream to the bioreactor in 1:1500 dilution. To enhance bioremediation in the plume 15 m<sup>3</sup> of molasses were added as carbon and energy source at 150 infiltration wells, covering 3600 m<sup>2</sup> of the pollution plume, in July 2006 and March 2007. Furthermore, to improve degradation of chlorinated compounds beyond cDCE, the bioreactor was bioaugmented in August 2007 by addition of 1.8 m<sup>3</sup> IJlst groundwater, which was rich in VC-degrading *Dehalococcoides* [130] (Fig. 2).





**Fig 1:** Overview of the chloroethene polluted research location in Ede, The Netherlands, Ground plan of the study site and concentration of chloroethenes in the plume and the position of the infiltration wells used to add electron donor and effluent discharge to the plume. Intervention levels for PCE: 40  $\mu\text{g/l}$ , TCE: 500  $\mu\text{g/l}$ , DCE: 20 $\mu\text{g/l}$  and VC: 5 $\mu\text{g/l}$ . Arrow shows the direction of groundwater flow.

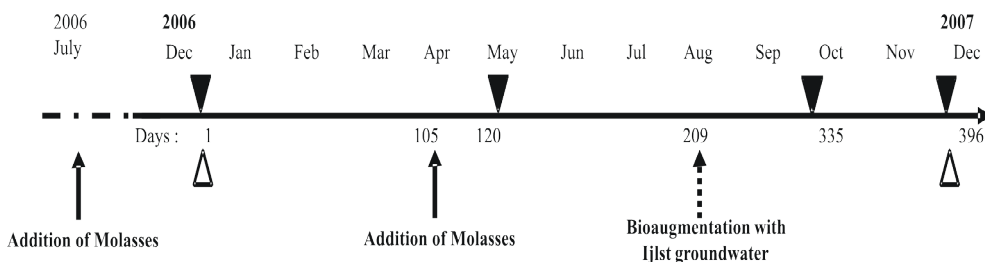


**Fig. 2:** Bioreactor drawing indicating the sampling points for influent, effluent and the reactor vessel

### Sample collection

Samples for physico-chemical analysis were obtained monthly from the influent, effluent and reactor content (Fig. 1 and 2) and occasionally from plume monitoring wells. On-site analysis was done to determine the influent and effluent temperature, pH, oxygen, and redox potential (mV), and chemical analyses were performed in a certified laboratory (Analytico Milieu B.V., Barneveld, Netherlands) using ISO and EPA protocols ([www.analytico.com](http://www.analytico.com)). These analyses included chlorine hydrocarbons (PCE, TCE, *cis*- and *trans*- isomers of DCE, and VC, di-, tri- and tetrachloromethane, di- and tri-chloroethane), inorganic chemicals (nitrate and nitrite, sulfate, sulfite and sulfide, iron and manganese), and volatile fatty acids (acetate, propionate and butyrate).

Samples for biomolecular analyses were obtained during a period of 12 months in December 2006 and in May, October and December 2007 (Fig. 3). Groundwater samples were taken from monitoring wells (MW1 and MW2; Fig. 1) in December 2006 and December 2007 to assess microbial communities present in the plume. Samples were obtained from monitoring wells at different depths (2-3m, 5-6m, 8-9m, and 11-12m), after three volumes of the sampled well were pumped out and discarded. The bioreactor was sampled from the influent, reactor content and effluent (Fig. 3). The reactor samples had a high content of particulate sludge matter resembling as close as possible the inside of the bioreactor without disturbing the operating conditions. All samples were transported anaerobically to the laboratory at +4°C.



**Fig. 3:** Sampling and monitoring scheme for the Ede site study. Samples for physicochemical analysis were collected monthly. Samples for molecular analysis obtained from bioreactor (solid triangles) and groundwater samples (open triangles). Points of intervention with addition of molasses to the plume and IJlst groundwater to bioreactor are shown by arrows.

### Chemical Analysis of Samples

The water temperature, redox potential and concentrations of chlorinated compounds were measured on a monthly basis. Water temperature ranged between 10 °C in winter and 20 °C in summer while pH was between 6.3 and 6.5. DOC (dissolved organic carbon) was on average  $19.7 \pm 2.5$  mg/L and  $25.0 \pm 4.6$  mg/L for the influent and effluent, respectively. Molasses was continuously dosed into the bioreactor and could have contributed to the higher DOC measurements observed for the effluent.

### DNA extraction

Samples (100-150 ml) were filtered through 0.2 µm polycarbonate filters (Millipore BV, Amsterdam, Netherlands) using a Millipore filtering system to collect microbial biomass. Filters were either stored at -20 °C or immediately processed for DNA extraction, whereby filters were cut into small strips and placed in bead beating tubes supplied with the FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH, USA). Genomic DNA extraction was done according to the manufacturer's instructions.

### DGGE analysis of PCR-amplified 16S rRNA gene fragments

PCRs were performed using the GoTaq DNA Polymerase Kit (Promega, Leiden, The Netherlands). PCR for V6-V8 regions of the bacterial 16S rRNA gene was performed using primers 968-GC-F and 1401-R [131]. *Dehalococcoides*-specific PCR was performed with a nested PCR approach, using primers DeF and DeR [132] for the first PCR, followed by a nested PCR using primers 968-GC-F [131] and DHC1350R [133]. DGGE analysis was done as described previously [53]. The similarity between DGGE profiles was determined by calculating pairwise similarity indices of densitometric curves of the profiles using Pearson's correlation. UPGMA dendrograms were constructed from this data.

### *Dehalococcoides* 16S rRNA gene clone library construction and sequencing

*Dehalococcoides*-specific PCR was performed on the May 2007 samples for cloning and sequencing using primers DeF and DeR [132] as described for *Dehalococcoides*-specific DGGE. PCR products were purified using the Zymo DNA Clean & Concentrator Kit (Zymo Research, San Francisco, CA, USA), ligated into cloning

vector pGEMT-Easy (Promega) and cloned into *E. coli* XL1-Blue cells (Stratagene, La Jolla, CA, USA), according to the manufacturer's instructions. White colonies were screened for correct-sized inserts by colony PCR on lysed cells with primers T7 and SP6, and checked by agarose gel electrophoresis. An extra check for *Dehalococcoides*-specific 16S rRNA gene amplicons was done on the correct-sized inserts using primers DHC730F and DHC1350R [133]. PCR parameters were: 2 min at 95°C, 35 cycles of 30 sec at 95°C, 40 sec at 55°C, 90 sec at 72°C, followed by 5 min at 72°C. Sequence diversity of cloned inserts was assessed by restriction fragment length polymorphism (RFLP) of amplicons using a mix of restriction enzymes (MspI, CfoI and AluI). Restriction fragments were then separated on 12% Poly(NAT) Wide-Mini S-4x25 PreCast Gels and electrophoresis performed with a SEA 2000 Apparatus (Elchrom Scientific, Cham, Switzerland). Ethidium bromide-stained gels were photographed using a UV illuminator. RFLP patterns were clustered to identify clones representing the same ribotype and inserts of selected clones for each RFLP-pattern were re-amplified by T7-SP6 PCR. T7-SP6 PCR products were purified and sequenced (Baseclear, Leiden, The Netherlands), and DNA sequences were manually checked for chimeric structures using Bellerophon [134]. Phylogenetic analyses were performed using the ARB software package [4], and a maximum likelihood tree was constructed using algorithms as implemented in ARB.

### Quantitative PCR

Quantitative PCR (qPCR) was performed using the iQ SYBR Green Supermix kit and the iQ5 iCycler (BioRad) for 16S rRNA genes of total and dehalogenating bacteria (*Dehalococcoides*, *Desulfitobacterium*, *Dehalobacter*) and *Dehalococcoides* Rdh genes (*tceA*, *bvcA*, *vcrA*) [4,48,135]. PCR protocols were followed as previously described [53], and respective gene copy numbers were calculated as copies/ml of water sample.

### Modeling Approach

Percentage of chlorinated compound consumed and consumption rates of PCE, TCE, cDCE and VC were calculated by mass-balancing based on the concentration of PCE, TCE, cDCE and VC measured in the influent and effluent of the reactor. To

fully dechlorinate PCE, TCE, cDCE and VC, respectively 8, 6, 4 and 2 electrons are required. Electron-accepting capacity (EAC) of influent and effluent can thus be calculated as:

$$EAC_x = 8*[PCE]_x + 6*[TCE]_x + 4*[cDCE]_x + 2*[VC]_x \quad [1]$$

In which  $[-]_x$  is either the concentration in the influent or effluent ( $\mu\text{M}$ ).

Percentage consumption of EAC in the influent was calculated as:

$$\% \text{ EAC consumption} = 100 \times \frac{EAC_{\text{influent}} - EAC_{\text{effluent}}}{EAC_{\text{influent}}} \quad [2]$$

As TCE, cDCE and VC are intermediates in organohalide respiration and are being produced in the reactor from their parent molecules, the measured concentrations of these compounds do not directly inform about their consumption (in  $\mu\text{M}$ ) in the reactor. The consumption was obtained as follows:

$$\Delta\text{PCE} = [\text{PCE}]_{\text{influent}} - [\text{PCE}]_{\text{effluent}} \quad [3]$$

$$\Delta\text{TCE} = \Delta\text{PCE} + [\text{TCE}]_{\text{influent}} - [\text{TCE}]_{\text{effluent}} \quad [4]$$

$$\Delta\text{cDCE} = \Delta\text{TCE} + [\text{cDCE}]_{\text{influent}} - [\text{cDCE}]_{\text{effluent}} \quad [5]$$

$$\Delta\text{VC} = \Delta\text{cDCE} + [\text{VC}]_{\text{influent}} - [\text{VC}]_{\text{effluent}} \quad [6]$$

Rates of consumption were obtained by multiplication of the consumption with the influent inflow rate  $f$  (in  $\text{L/h}$ ). The dilution rate  $D$  ( $\text{h}^{-1}$ ) of the reactor was calculated as:

$$D = f/V \quad [7]$$

where  $V$  is volume of the reactor in liters.

An estimate of *Dehalococcoides* numbers expected per ml groundwater was obtained via:

$$X = \Delta\text{EAC} * (Y/B) \quad [8]$$

In which  $Y$  is yield (in  $\text{g biomass/mole electrons}$ ),  $B$  is weight ( $\text{g biomass}$ ) per bacterium and  $\Delta\text{EAC} (= \text{EAC}_{\text{influent}} - \text{EAC}_{\text{effluent}})$  expressed in  $\mu\text{M}$ .

Growth rate of *Dehalococcoides* in the reactor was modeled according to the equation given by Becker [128] :

$$\mu_i = \mu_{\max,i} \frac{\frac{PCE}{K_{s,PCE}} + \frac{TCE}{K_{s,TCE}} + \frac{cDCE}{K_{s,cDCE}} + \frac{VC}{K_{s,VC}}}{1 + \frac{PCE}{K_{s,PCE}} + \frac{TCE}{K_{s,TCE}} + \frac{cDCE}{K_{s,cDCE}} + \frac{VC}{K_{s,VC}}} \quad [9]$$

In which  $\mu_{\max,i}$  is the maximal growth rate and  $K_{s,[PCE]}$ ,  $K_{s,[TCE]}$ ,  $K_{s,[cDCE]}$  and  $K_{s,[VC]}$  are the affinity constants for PCE, TCE, cDCE and VC, respectively. Equation was parameterized with published affinity constants and concentrations measured in the effluent (Table 1). Saturation (the degree to which maximum growth rate was obtained) was calculated by dividing equation 9 by  $\mu_{\max,i}$ .

### Statistical Analysis

Significance of the observed differences in gene copies number over time for influent, reactor and effluent was analyzed using Student's Unpaired *t*-test.

**Table 1:** Constant values used to calculate saturation of dehalogenating capacity.

Microorganism Culture	Substrate	Ks (μM)	Reference
<i>Dehalococcoides ethenogenes</i>	PCE	0.54	[128]
	TCE	0.54	
	cDCE	0.54	
	VC	290	
<i>Dehalobacter restrictus</i>	PCE	0.54	[128]
	TCE	0.54	
<i>Desulfuromonas michiganensis</i>	PCE	0.54	[128]
	TCE	0.54	

**Table 1:** Constant values used to calculate saturation of dehalogenating capacity continued

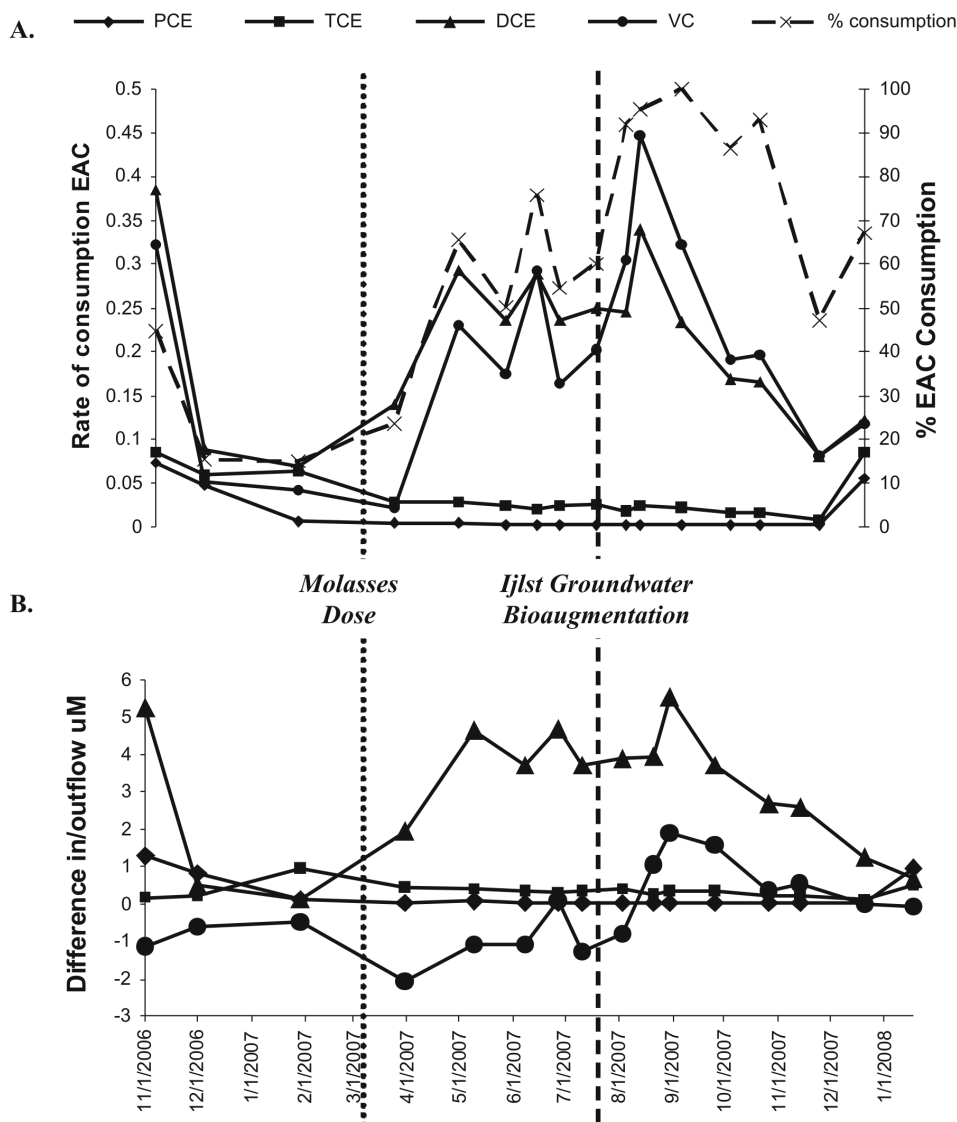
Microorganism / Culture	Substrate	Ks (μM)	Reference
<i>Desulfitobacterium</i> sp. PCE1	PCE	0.54	[128]
<i>Dehalococcoides</i> BAV-1 containing consortium	cDCE	8.9	[27]
	tDCE	8.5	
	VC	5.8	
<i>Dehalococcoides</i> consortia (VS)	TCE	9	[136]
	cDCE	3.3	
	VC	2.6	
<i>Dehalococcoides</i> consortia (KB-1/VC)	TCE	10	[136]
<i>Dehalococcoides</i> consortia (Pinellas)	TCE	10.5	[136]
Point Magu enrichment culture	PCE	3.9	[137]
	TCE	2.8	
	cDCE	1.9	
	VC	602	
Evanite enrichment culture	PCE	1.6	[137]
	TCE	1.8	
	cDCE	1.8	
	VC	62.6	
Enrichment Culture (toluene)	PCE --> ethene	0.86 + 0.71	[138]
Enrichment Culture ( lactate)	PCE --> ethene	0.23 + 1.45	[138]
Methanogenic dechlorinating consortium	PCE	0.1 +/- 0.05	[139]
Methanogenic dechlorinating consortium	PCE	2.8	[140]
	TCE	1.5	
	cDCE	3	
	VC	360	
Dechlorinating Consortium	PCE	0.11	[136,141]
	TCE	1.4	
	cDCE	3.3	
	VC	2.6	
Dechlorinating Consortium	PCE	70.7	[142]
	TCE	17.4	
	cDCE	11.9	
	VC	383	

## Results

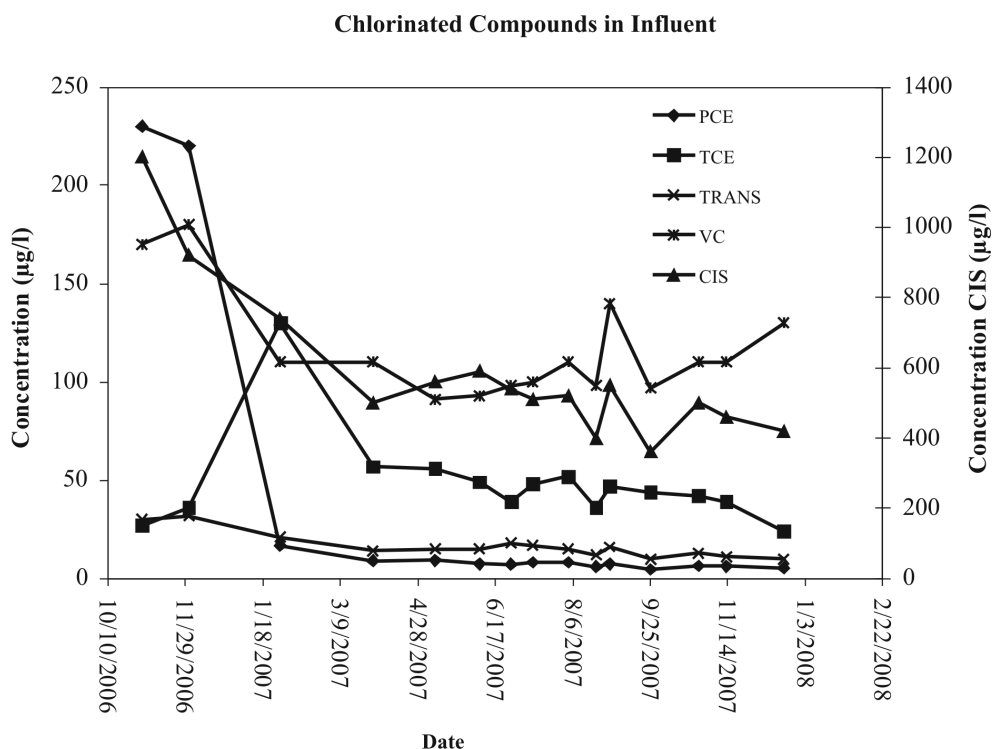
### Bioreactor performance

Dechlorination performance of the reactor improved over time (Fig. 4). Calculation of the electron accepting capacity (EAC) of the chlorinated compounds in the influent and effluent of the bioreactor, followed by subsequent determination of the percentage consumption of EAC in the reactor revealed that during the first 5 months only 20% of the EAC present in the influent was consumed (Fig. 4A). However, a steady increase of electron acceptor consumption was then observed till after 10 months almost 100% of the inflowing chlorinated compounds were consumed. This high EAC was maintained for 4 months, after which it dropped to 50%. Secondly, consumption rates of the different chlorinated compounds revealed, by correcting for the internal production of chlorinated intermediates, that consumption rates of VC were very high ( $> 0.15 \mu\text{mol/L/h}$ ) in the period April 2007 – November 2007 (Fig. 4A), despite the fact that often a net increase of VC in the reactor was observed (Fig. 4B). The EAC for VC consumption was comparable to that of cDCE and much higher than those of PCE and TCE. This relates to the fact that in the influent cDCE in general was present at the highest concentration. Rate of VC consumption increased from May 2007 and reached a peak in September 2007, one month after the bioreactor was bioaugmented with IJlst groundwater that contained actively VC-reducing populations (Fig. 4A). The plume chloroethene concentrations decreased over time as shown by decrease in influent chloroethene concentrations (Fig 5).





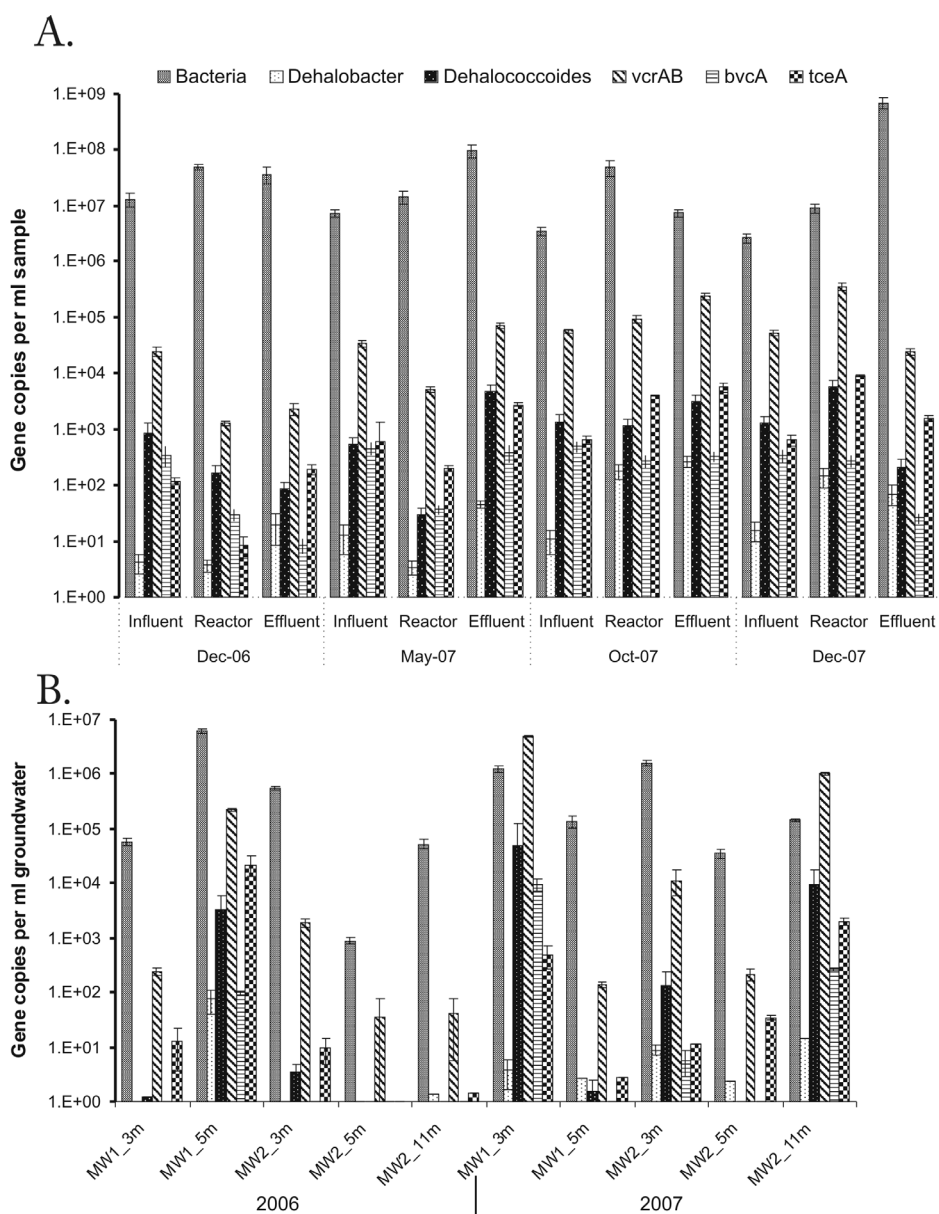
**Fig. 4.** A) Rate of consumption and % consumption of electron accepting capacity of individual chloroethenes and B) Fluctuations in the concentrations of chloroethenes measured in the influent and effluent. Negative values represent net production of a given chloroethene in the bioreactor. Points of molasses and IJlst groundwater addition are shown by the dotted and dashed lines, respectively.



**Fig. 5:** Concentration of chlorinated compounds in the influent stream to the bioreactor.

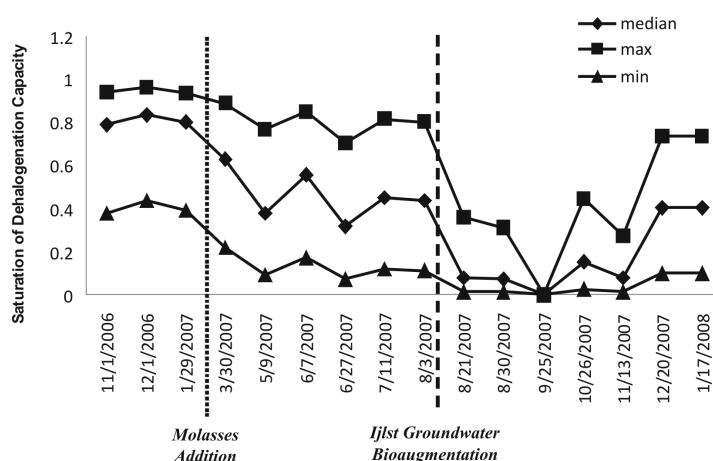
### Abundance and spatial distribution of 16S rRNA- and dehalogenase-encoding genes in bioreactor and groundwater samples

The presence of *Dehalobacter* and *Dehalococcoides* species 16S rRNA genes and *Dehalococcoides* Rdh genes *tceA*, *vcrA* and *bvcA* was confirmed in all samples (Fig. 6A). Total bacterial 16S rRNA gene copies were generally higher in the effluent ( $2.0 \times 10^8 \pm 4.8 \times 10^7$  copies/ml) than in the influent and reactor, which averaged  $6.6 \times 10^6 \pm 1.4 \times 10^6$  copies/ml and  $3.7 \times 10^7 \pm 1.0 \times 10^7$  copies/ml, respectively. *Dehalobacter* species were detected at very low abundance throughout the system at all time points and ranged from 10 to 100 16S rRNA gene copies/ml in the reactor and effluent. Numbers observed in the influent were negligible. *Desulfitobacterium* species were not detected in any of the reactor samples, including the field samples.



**Fig. 6:** Quantitative PCR analysis of bioreactor samples. The graph shows the results for 16S rRNA genes of total bacteria, *Dehalococcoides* and *Dehalobacter*, and Rdh genes *vcrA*, *bvcA* and *tceA*. Each bar represents the average of the results of triplicate qPCRs for **A)** Bioreactor samples –influent, reactor and effluent and **B)** groundwater samples from monitoring MW1 -3m and 5 m depths and MW2- 3m, 5m and 11m depths for December 2006 and 2007.

*Dehalococcoides* spp. 16S rRNA gene copies in the reactor were around  $10^3$  copies/ml, which was lower than expected since the bioreactor operating conditions were designed so as to optimize for their growth and proliferation. In agreement with the improvement in reactor performance over time (Fig. 5) numbers of *Dehalococcoides* increased from  $0.09 \times 10^3 \pm 0.03 \times 10^3$  in December 2006 to  $5.76 \times 10^3 \pm 1.72 \times 10^3$  copies/ml in December 2007. *Dehalococcoides* 16S rRNA gene copy numbers and growth varied significantly over time in the reactor compared to the influent and effluent ( $P = 0.0066$ ) and were higher in the effluent. Bioaugmentation with IJlst groundwater in August 2007 did not influence significantly the numbers of *Dehalococcoides*. However, an improvement was observed for the consumption of chlorinated compounds (Fig. 4B). Since *Dehalococcoides* was the dominant dechlorinating bacteria observed, saturation of its dehalogenation capacity was calculated based on reported affinity constants [128,132,143] and concentrations of chlorinated compounds in the reactor. This analysis revealed that in the period that consumption of chlorinated compounds was only 20% and started to increase to nearly 100% (December 2006 – August 2007), the saturation was nearly complete (i.e. near 1) (Fig. 7). However, when near 100% consumption of chloroethenes was achieved, the saturation dropped strongly, to about 10% (September 2007- December 2007).



**Fig. 7:** Saturation of *Dehalococcoides* dehalogenating capacity. Calculated from published affinity constants and measured concentrations of chlorinated compounds in the bioreactor.

If *Dehalococcoides* would grow at its maximum growth rate, then at least one of the chlorinated compounds PCE, TCE, cDCE and VC needs to be present at a concentration exceeding its affinity constant, so that the dehalogenation capacity of the dechlorinating microorganisms is close to full saturation. Thus while the performance of the reactor improved, dechlorinating microorganisms appear to be fast growing and, once near complete dehalogenation was achieved, *Dehalococcoides* only grew slowly and did not increase in number due to electron-acceptor limitation.

The *vcrA* gene was numerically dominant ( $1.0 \times 10^4$  copies/ml), compared to *bvcA* and *tceA* (Fig. 6A). Total numbers of VC Rdh-encoding genes, i.e. the sum of *vcrA* and *bvcA*, were higher than those observed for *Dehalococcoides* 16S rRNA genes. There was a significant and gradual increase in *vcrA* in the bioreactor over time suggesting proliferation of populations that harbor this gene, again in line the increase in cDCE and VC consumption. Total bacterial and Rdh gene numbers transferred to the plume in the effluent were either equal or higher than those in the suspended content of the reactor.

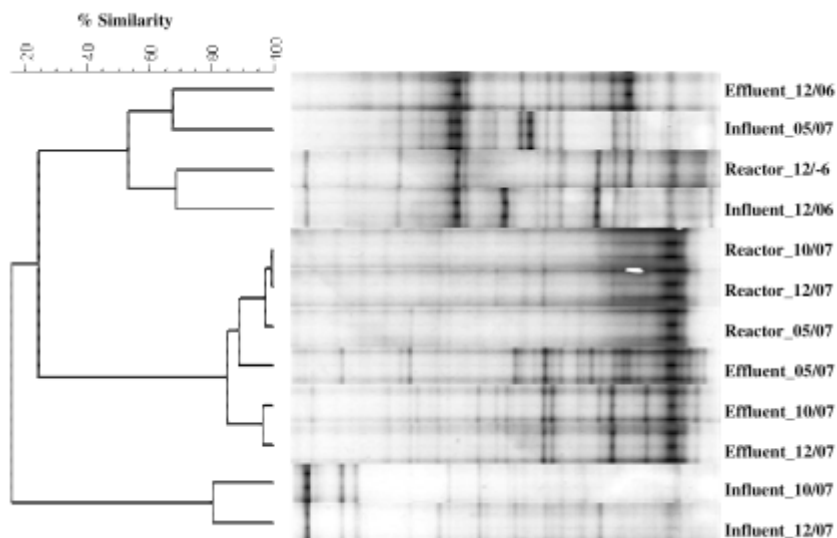
In field samples, *Dehalococcoides* and its catabolic genes were most dominant, while *Dehalobacter* was insignificant ( $<78$  16S rRNA gene copies/ml) and *Desulfitobacterium* was not detected (Fig. 6B). *Dehalococcoides* 16S rRNA and Rdh gene copy numbers were lower in December 2006 compared to those observed in the reactor and its effluent. However, they increased significantly over the year at both monitoring wells showing influence of the bioreactor effluent that was being pumped back into the plume around MW1 and MW2. For example in MW1, 2-3 m depth, *Dehalococcoides* was  $1.1 \times 10^5 \pm 7.7 \times 10^4$  copies/ml groundwater in December 2007, which was higher than observed in the bioreactor. In line with observations made for the bioreactor, *vcrA* copy concentrations were 1-2 orders of magnitude higher than *Dehalococcoides* spp. 16S rRNA gene copies. The *bvcA* gene was present in December 2007 field samples, but not in December 2006, except for one sample from MW1 at 5-6 m depth. The *bvcA* gene was observed at all time points in the bioreactor, suggesting that there was transfer to and stimulation of dechlorinating populations harboring *bvcAB* genes in the reactor, whereas they could have been outcompeted in situ.

### Community Profiling of Bioreactor samples

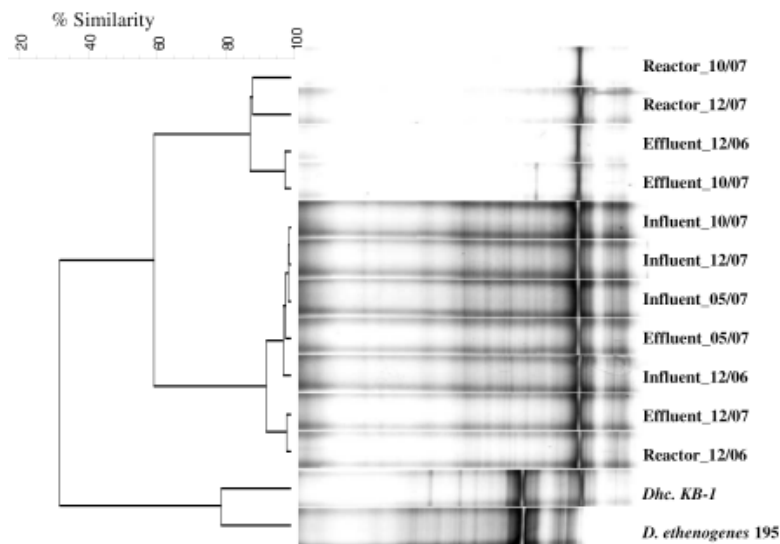
DGGE profiles of bacterial communities in different sections of the bioreactor (influent, reactor and effluent) and field samples (MW1 and MW2) were complex; the number of bands ranged from 10 to 36. Cluster analysis showed that 2006 samples of effluent and reactor were more similar to the corresponding influent sample (Fig. 8) in line with the low bioreactor performance at that time. Reactor and effluent 2007 samples clustered together sharing over 80% similarity. Influent communities showed more changes over time (31.6 % average similarity) possibly due to fluctuations in the field situation, and were only 20-40% similar to the reactor and effluent communities. Reactor and effluent communities were on average over 85% similar to each other over the monitoring period. Nevertheless, microbial communities of the reactor in December 2006 shared only 46 % similarity with the highly similar profiles of May, October and December 2007 samples suggesting that a stable community developed and was maintained throughout the rest of the year, despite changes in bioreactor operation, addition of substrate and bioaugmentation with IJlst groundwater. This stable community of the reactor did not seem to be affected by changes in influent communities. Only one dominant band could be observed for the influent, reactor and effluent *Dehalococcoides*-specific DGGE profiles, regardless of the sampling point and time (Fig. 9).

### Cloning and Sequencing of *Dehalococcoides*

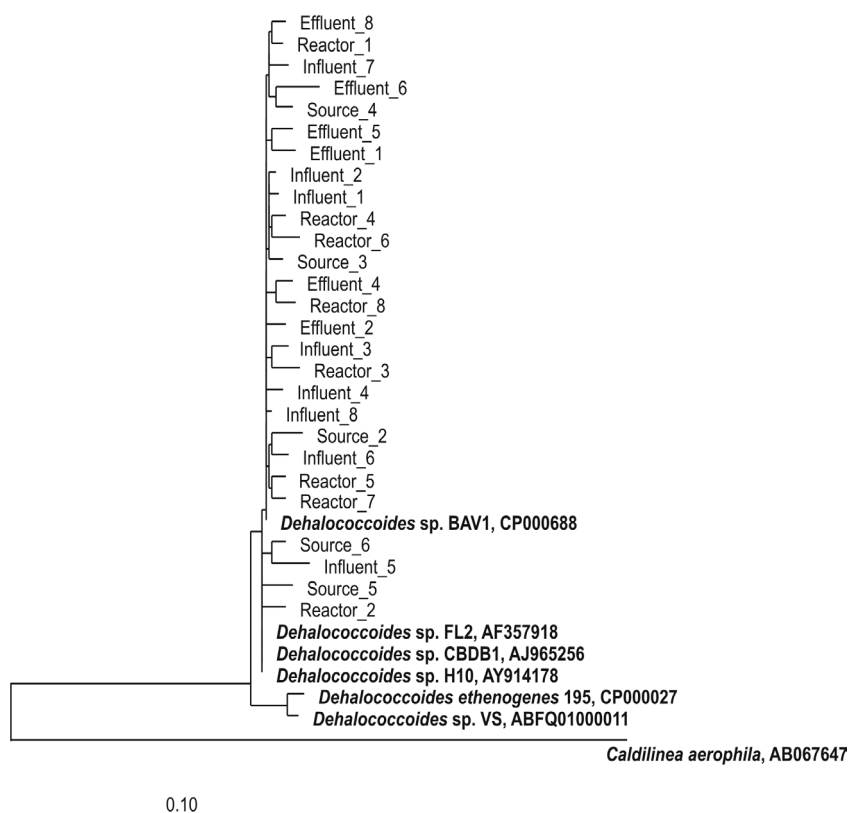
*Dehalococcoides* 16S rRNA gene clone libraries were constructed from the bioreactor samples and monitoring well MW1 obtained in May 2007 before bioaugmentation with IJlst groundwater (Fig. 10). Consistent with the qPCR data, which showed presence of Rdh's from *Dehalococcoides* sp. BAV1 from all sampling points of the bioreactor, sequences retrieved from both the bioreactor and field samples were most closely related to *Dehalococcoides* BAV1. However, no 16S rRNA sequences relating to *Dehalococcoides* sp. VS were retrieved in spite of dominant presence of *vcrAB* detected by qPCR in all samples.



**Fig. 8:** Clustering of 16S rRNA gene-based DGGE profiles of Bacteria after Pearson's correlation. For each sample the section of bioreactor and time of sampling month/year is shown.



**Fig. 9:** Clustering of *Dehalococcoides* specific DGGE profiles from bioreactor samples after Pearson's correlation. For each sample the section of bioreactor and time of sampling month/year is shown.



**Fig. 10:** Phylogenetic tree of representative *Dehalococcoides* sequences obtained from the influent, reactor, effluent in May 2007 and monitoring well MW1 (source) in December 2006 and 2007. Maximum likelihood algorithm as implemented in ARB was used to generate a phylogenetic tree from an ARB alignment. Reference sequences are given with their respective GenBank accession numbers. The reference bar indicates sequence dissimilarity of 10%. MW1 was selected because it was also located at the source point of the plume. The sequences retrieved from both the bioreactor and field samples were most closely related to *Dehalococcoides* BAV1.



## Discussion

The Ede site has been undergoing bioremediation since 2006; an indigenous dechlorinating population had been present in the groundwater but dechlorinated PCE to mostly cDCE. Therefore a dechlorinating bioreactor had been installed aiming at complete in-situ bioremediation. Improved performance of the bioreactor as well as molasses addition to the field influenced conditions towards complete dechlorination, as indicated by congruent changes in concentrations of chlorinated compounds and microbial community structure in the reactor and contaminant plume. Concentrations of chlorinated compounds in the influent stream to the bioreactor decreased over time resulting in PCE and TCE concentrations below the intervention levels, 40 µg/L and 500 µg/L respectively, by 2007 with continuing decrease of DCE and VC concentrations. There was an increase in *Dehalococcoides* and its Rdh genes in the plume between 2006 and 2007 in line with increased performance of the bioreactor which showed increasing numbers of dechlorinating microorganisms in time and especially an increase in *vcrA* and *bvcA* encoding enzymes responsible for complete dechlorination. This causal relationship between dechlorinating activity and removal of the target contaminants from the aquifer further strengthens the use of such bioremediation techniques as biocurtains for preventing spread of pollution plumes [144]. Mixed *Dehalococcoides* communities existed in both the field and the bioreactor. Appearance of *Dehalococcoides* communities with *bvcA* and increase of those with *vcrA* show the direct influence of the bioreactor since these communities were enriched in the reactor and transferred to the plume via the effluent stream. Sequence analysis of *Dehalococcoides* 16S rRNA gene amplicons showed that similar populations were present in the bioreactor system and in the field. Competition models [128] have shown that when other organohalide respiring populations are present, they may be able to out-compete *Dehalococcoides* populations for limiting amounts of electron donor at sites where daughter products accumulate. Our data showed that the indigenous dechlorinating populations were not able to completely dechlorinate PCE to ethene, however, once electron donor was added, dechlorination past DCE could be observed with concomitant increase in cell numbers of *Dehalococcoides* spp. harboring *vcrA* and *bvcA*.

High cDCE and VC consumption rates are in line with dominant presence of *vcrA*. The qPCR results indicate that the high consumption rate between April and November 2007 was likely due to the increase in *Dehalococcoides* populations in the bioreactor and concomitant increase of *vcrA*, *bvcA* and *tceA*. Rates of VC consumption were comparable to those of cDCE and much higher than those of PCE and TCE, indicating that main activities of the dechlorinating microbial communities were cDCE and VC reduction. *Dehalococcoides* spp. that harbor the *vcrA* and *bvcA* genes have been shown to dominate in DCE and VC enriched environments [46,124], and the presence and expression of *vcrA* and *bvcA* was found consistent with degradation of chlorinated compounds to ethene [49,145,146].

Co-existence of several populations of *Dehalococcoides* in the bioreactor and field samples as evidenced by presence of several *Rdh* genes, known to be present in different *Dehalococcoides* strains. An important observation from this study is that although *Dehalococcoides* are reported to have one 16S rRNA gene per genome and one *tceA*, *vcrA* or *bvcA* gene/genome [3,124,147], the sum of functional gene copies observed was greater than 16S rRNA gene copies. Copy numbers of *tceA* were in most cases equal to *Dehalococcoides* 16S rRNA gene copies. A similar trend was observed for *bvcA*, which was present in numbers either equal or lower than that of 16S rRNA gene copies. Most interestingly, copy numbers of *vcrA* were in all cases 1-2 orders of magnitude higher than the 16s rRNA gene. One possible explanation for this observation would be a massive expansion of *vcrA* copy numbers to 10-100 copies per genome. Given the fact that currently available genomes of *Dehalococcoides* spp. isolates harbor up to 36 non-identical single-copy *Rdh* genes [3,19,109], a more likely explanation would be the presence of cDCE/VC dechlorinating bacteria that are not picked up by the current *Dehalococcoides* qPCR primer set used in these analyses. Novel groups of OHRB, in consortia with *Dehalococcoides* populations that reductively dechlorinate PCE to cDCE have recently been identified in river- and marine sediments by stable isotope probing [55,148]. Hence, it is not unlikely that yet undiscovered groups harbor cDCE and VC dechlorinating ability. Secondly, the higher numbers of *vcrA* could be a result of horizontal gene transfer events. Studies of *Dehalococcoides* genomes have shown *Rdh* genes to be associated with mobile genetic elements [149], and secondly a

different evolutionary history of *vcrA* genes [150]. Our results indicate the need for further analysis to determine extent of mobility of VC Rdh-encoding genes from environmental samples.

Since *Dehalococcoides* was the main OHRB observed in the bioreactor and field samples the expected number of cells/ml was estimated to be  $3.7 \times 10^5$ . The estimated number of cells was in agreement with observed *vcrA* copy numbers, which were two orders of magnitude higher than for *bvcA* and *tceA* and two orders of magnitude higher than *Dehalococcoides* 16S rRNA gene numbers. Although conversions from 16S rRNA gene copies to biomass, gene copy numbers per cell and cell volume are required and errors in these estimates can lead to significant differences in biomass yields [129], our results indicate biomass yield calculations are in line with pollutant consumption. Saturation analysis showed that while the performance of the bioreactor improved, dechlorinating microorganisms were fast growing, whereas, once near complete dehalogenation was achieved, *Dehalococcoides* only grew slowly and did not increase in number.

Additionally to the above-discussed possible failure to detect *vcrA*-harboring populations by *Dehalococcoides*-specific 16S rRNA gene-targeted assays, the observation that measured cell numbers (based on 16S rRNA gene copies) in the suspended biomass were smaller than estimated could be an indication that a considerable part of biomass is growing in biofilms in the reactor. The reactor can be likened to a chemostat, a continuous culturing system from which suspended biomass is continuously removed with the effluent. Steady state growth rates in a chemostat equal the dilution rate, which is inflow rate (400 L/h) divided by volume (7 m<sup>3</sup>). This results in a dilution rate of 0.06 h<sup>-1</sup>, a rate much higher than published maximum growth rates of *Dehalococcoides* (0.02 h<sup>-1</sup>; [151]). Thus, if *Dehalococcoides* would solely grow in suspension it would not be able to maintain itself in the reactor, even if growing at maximum rates (which requires full saturation of dechlorination capacity (Fig. 3)). Though limited information is available for dechlorinating biofilms, these are known to be present in most natural and engineered systems, such as contaminated soils and aquifers or engineered permeable barriers [147,152]. Biofilms of dechlorinating bacteria have been observed to develop predominantly inside pores of support materials, with no or little biofilm formation on external

surfaces [153]. Considering the design and operation parameters of the bioreactor in this study, it is likely that the microorganisms form these biofilms in the porous granular material allowing them to form the stable microbial communities observed. In a membrane-biofilm reactor study, the biofilm community was shown to adapt to TCE and increased in numbers of *Dehalococcoides* spp. and *tceA* and *bvcA* [147].

Overall, there is a complex interplay within dechlorinating microbial communities, including the co-existence of several populations of *Dehalococcoides* spp. strains. Modeling the microbial and metabolic dynamics of dechlorinating populations as described in the present study is important for optimizing performance of cleanup strategies for contaminated plumes beyond information provided by molecular diagnostic endpoints.

### Acknowledgements

This work was supported by the Ecogenomics Project of the Netherlands Genomics Initiative. We acknowledge Floris van der Waal (P&J Milieuservices B.V.) for providing site data and samples used in this study and the valuable contribution of Iemke Bisschops and Miriam van Eekert (Lettinga Associates Foundation) on bioreactor design and operation.

# CHAPTER 5

## **Chasing Organohalide Respirers: Quantitative Multiplex Detection of Dechlorinating Bacteria using a Ligation Probe-based Detection System Coupled to OpenArrays™**

Farai Maphosa, Ronald van Doorn, Willem M. de Vos, Cor Schoen, Hauke Smidt

*Manuscript in preparation*

**ABSTRACT**

Bioremediation management strategies for sites contaminated with chlorinated compounds require monitoring technologies that enable simultaneous detection and quantification of a wide range of organohalide respiring bacteria. Many multiplex and quantitative detection methods suffer from compromises between the level of multiplexing, throughput and accuracy of quantification. In this paper we report the development and application of a high-throughput, ligation-based circularization assay for simultaneous quantitative detection of multiple organohalide respiring bacteria and their key reductive dehalogenases. The ligation probes are long oligonucleotides with target complementary regions at their 5' and 3' ends. Upon perfect target hybridization, the ligation probes are circularized via enzymatic ligation, subsequently serving as template for individual and standardized amplification via unique probe-specific primers. Adaptation to OpenArrays™, which can accommodate up to 3072, 33 nl PCR amplifications, allowed high-throughput real-time quantification. The ligation probes were designed to target 10 reductive dehalogenases and 16S rRNA genes from 9 phylogenetic groups, involved in reductive dechlorination. These were used to demonstrate the application of the detection system and all probes specifically detected their corresponding targets and provided discrimination against non-target sequences with very similar ligation target sites. The multiplex assay was used to profile four different field samples and it was possible to detect and quantify organohalide respiring bacteria with comparable results but larger reproducibility, sensitivity and dynamic range than conventional qPCR regimes. This system gives new possibilities for the simultaneous and high-throughput analysis of samples and potential biomarker targets.

## Introduction

High-throughput technologies are needed for monitoring the biodiversity and functional capabilities of microorganisms involved in breakdown of halogenated pollutants in the environment. Since the first microorganism, *Dehalococcoides ethenogenes* strain 195, capable of reducing the common groundwater contaminants tetrachloroethene (PCE) and trichloroethene (TCE) completely to ethene, was isolated a decade ago [122], many more microorganisms have been isolated and identified to play crucial roles in the organohalide respiration of a wide and diverse range halogenated pollutants. Key organohalide respiring bacteria (OHRB) belong to the bacterial genus *Dehalococcoides* [117,118,120], which can reductively dechlorinate PCE, TCE, cis-dichloroethene (cDCE), and vinyl chloride (VC) to ethene in addition to other recalcitrant chloroorganic pollutants. Also important are bacteria belonging to the genera *Desulfitobacterium*, *Dehalobacter*, *Sulfurospirillum*, *Anaeromyxobacter*, *Desulfuromonas* and *Geobacter*, which are able to degrade PCE to cDCE, and also dehalogenate several other halo-organic pollutants [15,18,25]. In addition, there are strong indications for the existence of additional, yet unexplored, microbial taxa that have the capacity to dehalogenate, and now several of them have been isolated or identified [54,55,148].

OHRB possess reductive dehalogenase (Rdh's) enzymes which are responsible for the breakdown of the halogenated compounds. These key enzymes of the halo-respiration have been biochemically and genetically characterized, and found to constitute a novel class of vitamin B12-depending enzymes coupling the oxidation of an electron donor to the reductive elimination of chlorides or other halogen ions [30,154]. As various OHRB taxa include species that do not grow by organohalide respiration and possibly do not contain these Rdh's, it is necessary to couple phylogenetic detection to functional information. For the latter, the well-characterized Rdh's have emerged as the most appropriate detection targets in studying and monitoring bioremediation of contaminated sites [46,53]. For example the similarity in 16S rRNA gene sequence between *Dehalococcoides* spp. does not necessarily translate into dehalogenation ability - this was observed for the isolates BAV1 and FL2, which share a highly similar (greater than 99.9% identity) 16S rRNA gene sequence, but only strain BAV1 can use VC as an electron acceptor

for growth [119,127]. Thus, it is important to include the Rdh genes as targets for characterization of dehalogenating ability.

Quantitative PCR protocols have been developed [124] for 16S rRNA targets and a variety of Rdh genes, including the TCE to VC Rdh gene (*tceA*) of *Dehalococcoides* strains 195 [125] and FL2 [119,126], the VC Rdh genes from *Dehalococcoides* strains VS [33] and GT [57] (denoted *vcrA*), and that from *Dehalococcoides* strain BAV1 (termed *bvcA*) [32]. Many of the routinely used assays are focused on *Dehalococcoides* spp. and their genes, and are generally overlooking other dehalogenating microorganisms. Yet, contaminated sites often contain a mixture of pollutants and environmental conditions that support the growth of a variety of OHRB.

To achieve the most reliable characterization of a soil's dehalogenating capacity, a wide range of OHRB should be monitored quantitatively. To date, quantitative PCR (qPCR) provides the most reliable means of target quantification [155-157]. Although several successful multiplex qPCR-based detection methodologies for microorganisms have been realized, the attainable level of multiplexing is relatively low, typically restricted to detection of only a few target microorganisms per assay [156,158-161]. Moreover, the accuracy of quantification in samples with highly unbalanced (disparate) target ratios is often limited [162-164]. Reliable detection and accurate quantification of several microorganisms in a single sample, therefore, requires separate reactions making large-scale screening of samples more laborious, time consuming and expensive. To increase efficiency, reliability and reduce expenses, it is desirable to develop simple multiplex assays that can specifically detect and quantify a high number of microorganisms simultaneously.

Recently, a novel ligation-based probe assay was developed that bridged the gap between the demands for highly specific target recognition and the high-throughput, multiplex quantification of microorganisms [165]. The high specificity was achieved via the use of ligation-based circularization probes which allow discrimination of microorganisms based on a single point mutation in the targeted DNA sequence [165-167]. This point mutation specificity is of prime importance for the reliable characterization of a soil's dehalogenating capacity, since microorganisms that are closely related to target OHRB and containing only one or a few mismatches



in the probe target sites, may have highly disparate ecologies from the target organism itself. The real-time PCR methodology combined with OpenArray™ technology additionally provided sensitive and highly accurate quantification in a high-throughput fashion. The OpenArray has 48 sub-arrays, allowing parallel testing of up to 48 samples, and each sub-array contains 64 microscopic holes of 33nl volumes. These hydrophilic holes and hydrophobic array surface coatings ensure the sample remains isolated through surface tension. The holes are automatically loaded with master mix (Fig 1) [60,165].

In the implemented strategy mixtures of multiple circularization probes are ligated on fragmented DNA samples [165]. Target recognition is achieved by specific hybridization of the specific sequences flanking the target of these circularization probes, and efficient ligation occurs only if the end nucleotides perfectly match the DNA target. Next, the circularized probes are individually amplified by using PRI-lock-unique primer pairs via universal real-time PCR in OpenArrays™, enabling accurate target quantification in a highly multiplex format. (Fig 1)

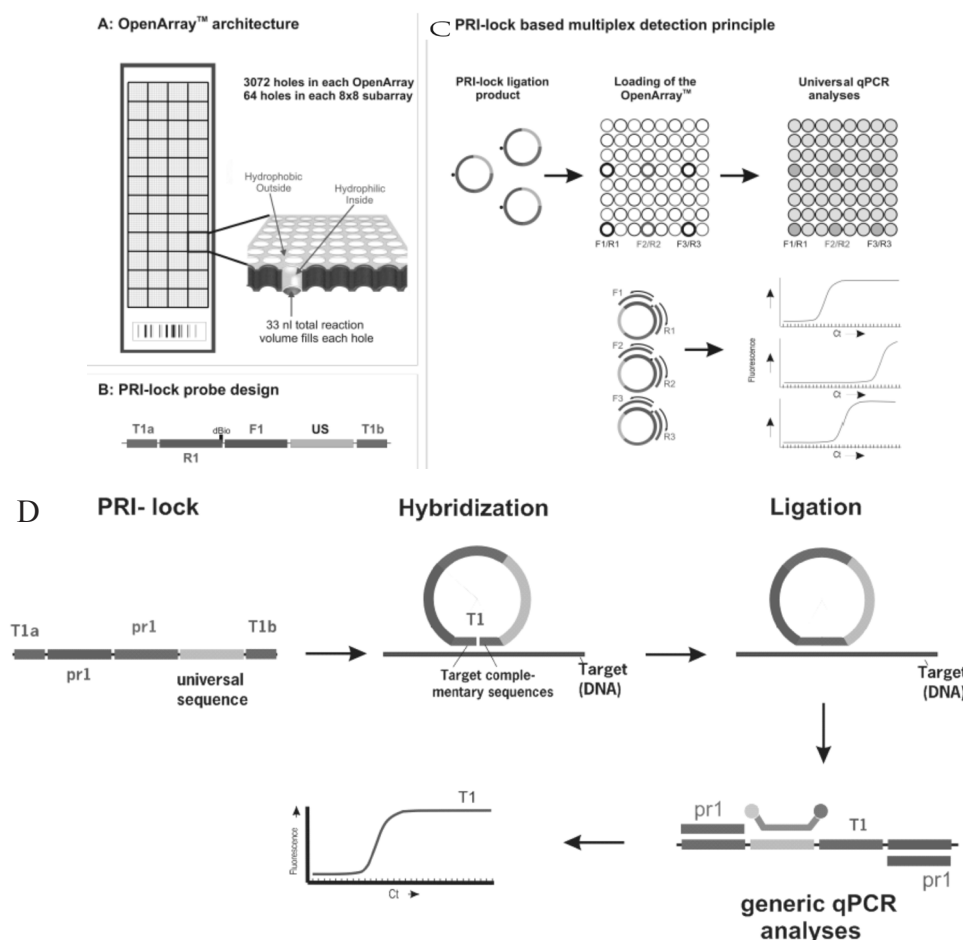
In this paper we report the development of a ligation-based, quantitative detection assay for the high-throughput detection of OHRB and the subsequent characterization of the organohalide respiration capacity of soils in contaminated environments. The specificity, sensitivity, and dynamic range of detection of the developed assay were determined by using 19 target-specific circularization probes that were specific to their target sequences over a linear quantification range of 5-6 orders of magnitude and could be used to multiplex detection on the OpenArray system™.

## Materials and Methods

### Site Samples

Four samples with different composition of OHRB as characterized by routine conventional qPCR were used in application of the PRI-lock probe system and for testing the multiplex assay on the Biotrove OpenArray system™.

Sample 1 was an enrichment sample made from river sediments contaminated with PCE and hexachlorobenzene. Original site and enrichment were



**Fig. 1:** Schematic overview of the proposed assay. (A) OpenArray™ architecture. The OpenArray™ has 48 subarrays and each subarray contains 64 microscopic through-holes of 33 nl volume. The primers are pre-loaded into the holes. The sample combined with the reaction mix is auto-loaded due to the surface tension, provided by the hydrophilic coating of the holes and the hydrophobic surface of the array. (B) Ligation (PRI-lock) probe design. T1a and T1b indicate target complementary regions. Unique primer sites ensure specific amplification (forward: F1 and reverse: R1) and each Ligation Probe (PRI-lock) contains a universal sequence (US) and a desthiobiotin moiety (dBio). (C) Multiple target specific Ligation probes are ligated on fragmented DNA samples. T1a and T1b bind to adjacent sequences of the target and in case of a perfect match, the probe is circularized by a ligase. Unreacted Ligation probes are removed by exonuclease treatment. Circularized probes are loaded and independently amplified on the Biotrove OpenArray™ platform using Ligation probe specific primers. The amplification is monitored using SYBR-Green and the ligated probes are quantified based on the threshold cycle number (CT). (Figure adapted from van Doorn et. al. 2007 [165]).

previously described, and were found to contain populations of *Dehalococcoides* and *Sulfurospirillum* spp. [168]. Secondly, a sample from an *in-situ* dechlorinating bioreactor was also used. This bioreactor is being used to bioremediate a site contaminated with chloroethenes and mineral oils (Chapter 4 of this thesis). Thirdly, two groundwater samples were included that were obtained from a contaminated site located in the province of Groningen, the Netherlands, that was contaminated by spillage of the dry-cleaning solvents PCE and TCE. These two field samples were taken at different time points during routine monitoring of bioremediation activities that are currently taking place at the site (Bioclear b.v., Groningen, the Netherlands). At this site a *Dehalococcoides* culture and carbon source had been injected into the soil to bioaugment and biostimulate reductive dechlorination of the chlorinated compounds. Sample “Field-1” was taken at the beginning of the monitoring round when bioaugmentation had been completed and “Field-120” was taken 120 days later. The two field samples were used to evaluate the applicability of the PRI-lock system for the monitoring of environmental samples.

### DNA extraction

Field samples (100-150 ml) were filtered through 0.2 µm polycarbonate filters (Millipore BV, Amsterdam, Netherlands) using a Millipore filtering system to collect microbial biomass. Filters were either stored at -20 °C or immediately processed for DNA extraction, whereby filters were cut into small strips and placed in bead beating tubes supplied with the FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH, USA). Genomic DNA extraction was done according to the manufacturer’s instructions.

### Ligation circularization probe design

The PRI-lock probe target complementary regions were engineered according to previously described design criteria [166] and were connected by a 60 bp compound linker sequence. The linker sequence contained a 20-bp generic TaqMan probe-binding sequence and two unique primer binding sites for specific PCR amplification (Table 1). All primer pairs have equal melting temperatures to allow universal TaqMan and SYBR-Green based detection in real-time PCR. The

primer pairs were chosen from the GeneFlex™ TagArray set (Affymetrix, Santa Clara, CA, USA) so as to minimize probe secondary structures and to optimize primer  $T_M$  and primer specificity. Potential for secondary structures, primer  $T_M$  and primer specificity were predicted using Visual OMP 6.0 software (DNA Software Inc., Ann Arbor, MI, USA). The prediction parameters were set to match ligation ([monovalent] = 0.025 M; [Mg<sup>2+</sup>] = 0.01 M; T = 65°C, [probe] = 250 pM) and PCR conditions ([monovalent] = 0.075 M; [Mg<sup>2+</sup>] = 0.005 M, T = 60°C). When necessary, PRI-lock probe arm sequences were adjusted to avoid strong secondary structures that might interfere with efficient ligation as described previously [166].

To monitor the ligation efficiency and provide a reference for normalization, an Internal Ligation Control (ILC) PRI-lock probe was constructed [165]. The target complementary regions of the ILC PRI-lock detect an artificial DNA sequence, of which a fixed amount was added to each ligation reaction, resulting in a standard ILC CT value of 23 for the iQ5 Biorad and the Biotrove OpenArray™. Variation in the ligation reaction was monitored by comparing the observed ILC CT values with the fixed ILC CT value. The  $C_T$  values were standardized as described previously [165].

The PRI-lock probes listed in Table 1 and all the other oligonucleotides used in this study, were synthesized by Eurogentec SA (Seraing, Belgium)

## Ligation

PCR fragments and genomic DNAs were used as templates for ligation. Prior to ligation, the DNA targets were fragmented by digestion using EcoRI, HindIII and BamHI (New England Biolabs Inc., Ipswich, MA, USA) for 15 min at 37°C. Cycled ligation was performed in 10 µl reaction mixture containing 20 mM Tris-HCl, pH 7.5, 20 mM KCl, 10 mM MgCl<sub>2</sub>, 0.1% Igepal, 0.01 mM rATP, 1 mM DTT, 20 ng sonicated salmon sperm DNA and 4 U Pfu DNA ligase (Stratagene, La Jolla, CA, USA). For multiplex detection, the optimized concentrations of the individual probes ranged from 100 pM to 10 nM. Reaction mixtures were prepared on ice and rapidly transferred into a thermal cycler. Before ligation, samples were denatured at 95°C for 5 min. The samples were subsequently subjected to 20 cycles of 30 s at 95°C and 5 min at 64°C. After the final cycle, the reaction was immediately cooled to 4°C.

### Exonuclease treatment

After the ligation reaction, 10 µl of exonuclease mixture [10 mM Tris-HCl, pH 9.0, 4.4 mM MgCl<sub>2</sub>, 0.1 mg/ml BSA, 0.5 U Exonuclease I (USB Europe GmbH, Staufien, Germany) and 0.5 U Exonuclease III (USB Europe GmbH)] was added to each reaction, and the samples were incubated at 37°C for 30 min, followed by enzyme inactivation at 95°C for 2.5 h.

### Real-time PCR

Amplification of ligated probes was monitored in real-time using the iQ5 iCycler real-Time PCR system (Biorad, Veenendaal, The Netherlands). Reactions were performed in 1× TaqMan® universal PCR Master Mix (Applied Biosystems, Carlsbad, CA, USA) containing, 0.12 µl UNG (Applied Biosystems), 1 µl ligation-exonuclease product, 300 nM of the PRI-lock probe-specific forward and reverse primers and 100 nM TaqMan probe. The reaction mixture was initially incubated at 50°C for 2 min, followed by 10 min denaturation at 95°C, and 40 cycles of 15 s at 95°C and 1 min at 60°C. Fluorescence was recorded in the 60°C step of each cycle,

### Biotrove OpenArray™ real-time PCR

Amplification of ligated circularization probes was followed in real-time using an OpenArray™ NT Cyclor (BioTrove Inc., Woburn, MA, USA). Samples were loaded into OpenArray™ plates using the OpenArray™ NT Autoloader according to the manufacturer's protocols. Each subarray was loaded with 5.0 µl mastermix containing 2.5 µl ligated PRI-lock mixture and reagents in a final concentration of 1× LightCycler® FastStart DNA Master SYBR Green I mix (Roche Diagnostics GmbH, Mannheim, Germany), 0.2% Pluronic F-68 (Gibco, Carlsbad, CA, USA), 1 mg/ml BSA (Sigma-Aldrich, St Louis, USA), 1:4000 SYBR Green I (Sigma-Aldrich), 0.5% (v/v) Glycerol (Sigma-Aldrich), 8% (v/v) deionized formamide (Sigma-Aldrich). The PCR OpenArray™ thermal cycling protocol consisted of 90°C for 10 min, followed by cycles of 28 s at 95°C, 1 min at 60°C and 70 s at 72°C (imaging step). Simulation of Biotrove PCR conditions in Visual OMP 5.0 software (DNA Software Inc.) estimated that the 60°C as annealing temperature in the Biotrove OpenArray™. The Biotrove OpenArray™ NT Cyclor System Software uses a proprietary calling

algorithm that estimates the quality of each individual CT value by calculating a CT confidence value for the amplification reaction. PCR amplification with a CT confidence lower than 100 were regarded as background signals. The remaining positive amplification reactions were analyzed for amplicon specificity by studying the individual melting curves. A Chi square analysis to determine number of replicates that should be positive in an unknown sample measurement was used to rule out any false positives when melting curve analysis of negative controls showed amplifications.

## Results

### Probe design

The ligation-based circularization probes were designed for 10 phylogenetic groups of OHRB and 10 Rdh's that are important in characterization of microbial reductive dehalogenation potential at contaminated sites. Using the 16S rRNA sequences for taxonomic and the Rdh-encoding genes for functional information, target and non-target sequences were selected and probes designed (Table 1). Target complementary regions were selected based on alignments and phylogenetic trees with the target and closest non-target bacteria 16S rRNA sequences or Rdh gene sequence (Table 2). Probes were also designed to distinguish between very closely related sequences like those of *Dehalococcoides ethenogenes* 195 and *Dehalococcoides* sp strain VS (Cornell subgroup) from *Dehalococcoides* strains CBDB1, FL2, BAV1 (Pinellas subgroup) which share 99 % sequence identity. Mismatches between the target sequence and the closest non-target sequence in the 3' arm of each probe provided for probe specificity (Table 2).

### Evaluation of assay performance

An Internal Ligation Control (ILC) PRI-lock probe was used to correct for the variability arising from the ligation reaction, since the target detection and quantification are dependent on the ligation and real-time PCR amplification steps, as described previously [165].

**Table 1:** Sequences of the target complementary regions and unique primer sequences for the ligation probes. All sequences are depicted as 5' to 3'.

organism	5' Ligation Probe target sequence	3' Ligation Probe target sequence	Primer 1 (Reverse)	Primer2 (Forward)
<i>Dehalobacter</i>	GCTCAACTCCGTAAAGCATTTGG	TGAAAGTGGGA	ATAACTGTGAGCGTGTGATG	CATCCAGCTCAACGTATCCA
<i>Dehalococcoides</i>	GGGAAAGCCAACTGATGAAGTAC	GAATTGTCAGTAAA	AGAAACAGTCAGAGGTCGGT	TACGAAACGTCTTAGCATCC
<i>Dehalococcoides</i>	TCATCACITTCACITTTCCCTTGCTGGA	AAACTGTACGTGA	CATAGGACTGACACGAGCTG	AGAACTGTACACGCTGCCAT
<i>Dehalobacterium</i>	CGAGTTGCAGATAACAATCCGAAC	GACTTCATGTTCT	ATCAGATCGACACCGTAGCT	CTTGTGCATGTACTCGACTG
<i>Sulfurospirillum</i>	CAAGGTCATCTAGCGGAAGC	GTCGCAAGTTATTA	ATAGTCTCTGAGCCGGATCT	GAGTCTGCTGTTGACCATC
<i>Anaeromyxobacter</i>	GGCGGATGTCGGGAAATGC	GCTGAACGATCC	TGATTACGGACTGTGCTTTG	CACCACACTGGAGCGGTATA
<i>Desulfuromonas</i>	CGTACTTTGTGGACACAGGGGGTCAGTTCGGCT	TTACATCCCGAT	ACTCGGTCACTCAGCATATA	CAATACCTGTGACGAGCCTG
<i>Sedimentibacter</i>	GGTCACITATGTAAAGGGCCACAG	GCTTAGCTTTGCA	ATTCTAATCAATCGTCGGGG	CTTTGGCGTGTAGTGGTTT
<i>Geobacter</i>	GAACTTATGCCCTCTCCCATC	CTTTGACCGTCT	ATATTGGAAAGTCGCCCTGG	CTTCGGTGTCAAGTATCTCT
<i>LAHN CLUSTER</i>	TCACGACTATGGGTACACCA	TGTGAAGCTTTCT	GATGAATGTCTGCTGCTGCTC	CGGCGTGATGATCACTACTTT
<b>Rdase gene</b>				
<i>Sul_pceA</i>	TTCACTAATTGCTTTTGAAGGACACTC	AAGTTCTTGGAC	ATAGTCTCTGAGCCGGATCT	GAGTCTGCTGTTGACCATC
<i>Dsb_pceA</i>	TCTGGGAGCCGATCTAGTCG	AGCAACCCGTC	ATATTGGAAAGTCGCCCTGG	CTTCGGTGTCAAGTATCTCT
<i>Det_pceA</i>	GTCTTTGAATCATATCCAGTCTATAGG	CAAGTGAAGGC	GATGAATGTCTGCTGCTGCTC	CGGCGTGATGATCACTACTTT
<i>verA</i>	AGCACTGCTATACCAACCG	ATACAGTATATGATCC	TGATTACGGAGTGTGCTTTG	CACCACTGTGAGCGGTATA
<i>bvcA</i>	ATTACACCTATACTGGAGTAACCAAAA	CTGGTGATTGGC	ATTCTAATCAATCGTCGGGG	CTTTGGCGTGTAGTGGTTT
<i>tceA</i>	ATTAACAGGTTCTCTTCAGGCG	CGGTGGCGATG	AGAAACAGTCAGAGGTCGGT	CCTCACGATTTCTCAAAGTAT
<i>cprA1</i>	ACTTTACAGAACTTCTGCGG	CTGCAGCCCA	ATAACTGTGAGCGTGTGATG	CATCCAGCTCAACGTATCCA
<i>cprA5</i>	CCCTGATCTTCGGCTGCG	CGTGATGAGCTG	ATCAGATCGACACCGTAGCT	CTTGTGCATGTACTCGAACC
<i>dca</i>	ATGCTGACCAAAAGAAAGAACGC	CATGACACCGT	ACTCGGTCACTCAGCATATA	CAATACCTGTGAGCGCTG
<i>cbrA</i>	CCGGGTTGACCCCAAAAAGTGT	GACACGTATCCAAC	CATACGACTGACACGAGCTG	AGAACTGTACACGCTGCCAT

**Table 2:** OHRBs 16S rRNA and reductive dehalogenase genes targeted by ligation probe and their closest non-target organism or reductive dehalogenase gene. The 3' arm of the ligation probe is shown with the main discriminating nucleotides underlined.

Target organism	3' Ligation probe sequence (5'-3')	Closest Non-Target Organism
<i>Dehalobacter</i>	TGAAAGTC <u>CGGA</u>	<i>Desulfitobacterium</i>
<i>Dehalococcoides</i> 195 + rel	GAATTGTCC <u>AGTAAA</u>	<i>Dehalococcoides</i> CBDB1 + rel
<i>Dehalococcoides</i> CBDB1 + rel	AAACTGTACGT <u>GCA</u>	<i>Dehalococcoides</i> 195 + rel
<i>Desulfitobacterium</i>	GACTTCATGT <u>TCT</u>	<i>Dehalobacter</i>
<i>Sulfurospirillum</i>	GTGCGASTT <u>ATTA</u>	<i>Campylobacter</i>
<i>Anaeromyxobacter</i>	GCTGAACGAT <u>TCC</u>	<i>Desulfuromonas</i>
<i>Desulfuromonas</i>	TTGACATCC <u>CGAT</u>	<i>Geobacter</i>
<i>Sedimentibacter</i>	GCTTAGTCTT <u>GCA</u>	<i>Dehalobacter</i>
<i>Geobacter</i>	CTTTGACCGT <u>TCT</u>	<i>Desulfuromonas</i>
LAHN CLUSTER	TGTGAAGCTT <u>TCT</u>	<i>Dehalococcoides</i> 195 + rel
<b>Target rdh gene</b>		<b>Non-Target rdh gene</b>
<i>Sul_pceA</i>	AAGTTCTTG <u>ACC</u>	<i>Dsb_pceA</i>
<i>Dsb_pceA</i>	AGCAACCGT <u>TC</u>	<i>cprA5</i>
<i>Det_pceA</i>	CAAGTGTAAC <u>GGC</u>	<i>tceA</i>
<i>vcrA</i>	ATACAGTATATG <u>ATCC</u>	<i>bvcA</i>
<i>bvcA</i>	CTGGTGA <u>TGGC</u>	<i>vcrA</i>
<i>tceA</i>	CGGTGCGC <u>ATG</u>	<i>bvcA</i>
<i>cprA1</i>	CTGCAGC <u>CCA</u>	<i>cprA3</i>
<i>cprA5</i>	CGTGATGAGC <u>CTG</u>	<i>Dsb_pceA</i>
<i>dcA</i>	CATGACAC <u>CCGT</u>	<i>Dsb_pceA</i>
<i>cbrA</i>	GACACGTATCCAAC	<i>cbdbA88</i>

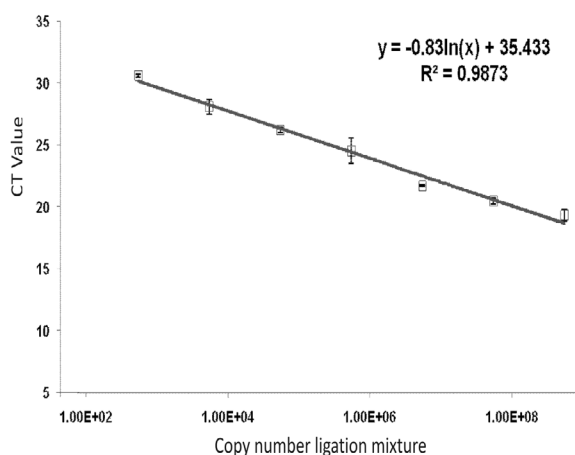
**Table 3:** Specificity of the PRI-lock probes for the target and non-target DNA. CT values are shown, 100pM of probe was ligated to 1 pg of DNA.

PRI-lock Probe	Target CT	Non-target CT	No-template control CT
<i>Dehalobacter</i>	19.3	31.0	31.4
<i>Dehalococcoides</i> 195 + rel	19.1	30.3	30.9
<i>Dehalococcoides</i> DCDB1+ rel	22.3	32.6	35.4
<i>Desulfitobacterium</i>	23.4	31.5	31.4
<i>Sulfurospirillum</i>	16.9	32.3	33.0
<i>Anaeromyxobacter</i>	13.5	28.1	28.2
<i>Desulfuromonas</i>	21.8	30.1	31.1
<i>Sedimentibacter</i>	18.2	28.8	29.3
<i>Geobacter lovleyi</i>	14.3	29.5	27.1
<i>Sul_pceA</i>	17.9	32.6	32.2
<i>Dhb_pceA</i>	16.7	30.6	30.2
<i>Det_pceA</i>	16.8	33.9	34.2
<i>vcrAB</i>	23.9	32.6	31.6
<i>bvcAB</i>	19.7	31.4	37.6
<i>tceA</i>	18.9	35.5	36.0
<i>cprA1</i>	15.1	28.2	28.3
<i>cprA5</i>	19.8	32.9	33.5
<i>dcA</i>	23.0	31.8	31.7
<i>cbrA</i>	23.2	31.6	32.9



The specificity of each PRI-lock probe to their individual targets was evaluated using two probe mixtures (one mixture of 10 phylogenetic probes and one mixture of 10 functional probes) in separate ligation reactions followed by real-time PCR. With the exception of the Lahn cluster probe, all the probes specifically detected their corresponding targets (Table 3). To ensure specificity towards the targeted microorganisms, all non-target organisms containing sequences similar to the target ligation regions for the PRI-lock probe mixtures, were tested. In all cases, the detected amplification signals were not significantly different from the non-target controls, indicating that no a-specific ligation occurred on the non-target DNAs (Table 3).

In order to establish detection sensitivity and linear quantification range, calibration curves were constructed based on CT values measured in a multiplex setting on a 10-fold dilution series of all the target DNA sequences. This approach was possible, as it had been shown previously that specific ligation of each probe is independent of other probes [165]. Detection was observed with the developed PRI-lock probes over a range from  $10^3$  to  $10^8$  copies / ligation mixture for both the taxonomic and Rdh gene probes (Table 4). Additionally, the DNA targets could be quantified in a linear fashion over 5 - 6 orders of magnitude (Fig. 2, Table 4). Linear correlation of the observed CT values and the logarithmic target concentrations in the ligation mixtures was higher than  $R^2 = 0.98$  in all cases (Table 4), demonstrating that the PRI-lock system can be used with confidence for reliable quantification over a wide concentration range for both the taxonomic and functional targets.



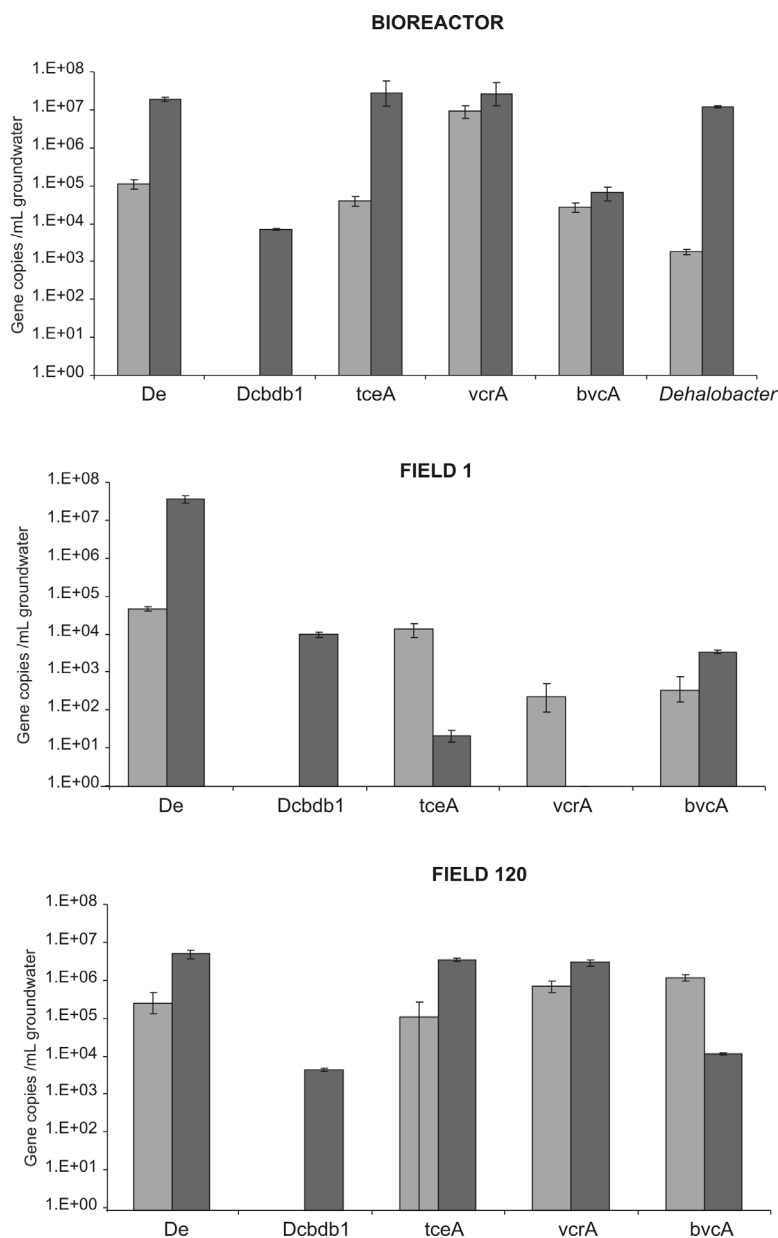
**Fig. 2:** Typical calibration curve to assess the sensitivity and linear range of quantification of the *Dehalococcoides* Ligation probe in conventional real-time PCR. CT values were normalized using the Internal Ligation Control PRI-lock probe.

**Table 4:** Summary of the parameters of the calibration formulas obtained with 10-fold dilution series of the ligation mixture of each probe - the relevant information on r-squared, sensitivity and linear detection range is presented.

PRI-lock probe	Calibration curve formula	R <sup>2</sup> Value	Detection limit	Linear quantification range
<i>Dehalobacter</i>	$y = -0.84\ln(x) + 30.8$	0.989	10 <sup>3</sup>	10 <sup>8</sup> – 10 <sup>3</sup>
<i>Dehalococcoides ethenogenes</i>	$y = -0.79\ln(x) + 29.8$	0.984	10 <sup>3</sup>	10 <sup>8</sup> – 10 <sup>3</sup>
<i>Dehalococcoides CBDBI</i>	$y = -0.96\ln(x) + 35.3$	0.996	10 <sup>3</sup>	10 <sup>8</sup> – 10 <sup>3</sup>
<i>Desulfotobacterium</i>	$y = -0.87\ln(x) + 29.8$	0.984	10 <sup>3</sup>	10 <sup>8</sup> – 10 <sup>3</sup>
<i>Sulfurospirillum</i>	$y = -0.85\ln(x) + 28.6$	0.994	10 <sup>2</sup>	10 <sup>8</sup> – 10 <sup>2</sup>
<i>Anaeromyxobacter</i>	$y = -1.02\ln(x) + 33.4$	0.976	10 <sup>4</sup>	10 <sup>8</sup> – 10 <sup>4</sup>
<i>Desulfuromonas</i>	$y = -0.75\ln(x) + 31.7$	0.983	10 <sup>2</sup>	10 <sup>8</sup> – 10 <sup>2</sup>
<i>Sedimentibacter</i>	$y = -1.06\ln(x) + 32.3$	0.977	10 <sup>3</sup>	10 <sup>8</sup> – 10 <sup>3</sup>
<i>Geobacter</i>	$y = -0.75\ln(x) + 31.7$	0.970	10 <sup>4</sup>	10 <sup>8</sup> – 10 <sup>4</sup>
<i>Sul_pceA</i>	$y = -1.11\ln(x) + 36.1$	0.999	10 <sup>4</sup>	10 <sup>8</sup> – 10 <sup>4</sup>
<i>Dsb_pceA</i>	$y = -1.12\ln(x) + 34.8$	0.994	10 <sup>3</sup>	10 <sup>8</sup> – 10 <sup>3</sup>
<i>Det_pceA</i>	$y = -1.16\ln(x) + 36.1$	0.988	10 <sup>3</sup>	10 <sup>8</sup> – 10 <sup>3</sup>
<i>vcrA</i>	$y = -1.03\ln(x) + 33.5$	0.975	10 <sup>3</sup>	10 <sup>8</sup> – 10 <sup>3</sup>
<i>bvcA</i>	$y = -1.36\ln(x) + 35.3$	0.992	10 <sup>3</sup>	10 <sup>8</sup> – 10 <sup>3</sup>
<i>tceA</i>	$y = -1.11\ln(x) + 34.6$	0.996	10 <sup>3</sup>	10 <sup>8</sup> – 10 <sup>3</sup>
<i>cprA1</i>	$y = -1.22\ln(x) + 37.7$	0.993	10 <sup>4</sup>	10 <sup>8</sup> – 10 <sup>4</sup>
<i>cprA5</i>	$y = -1.10\ln(x) + 32.1$	0.997	10 <sup>2</sup>	10 <sup>8</sup> – 10 <sup>2</sup>
<i>dcaA</i>	$y = -1.09\ln(x) + 34.3$	0.995	10 <sup>3</sup>	10 <sup>8</sup> – 10 <sup>3</sup>
<i>cbrA</i>	$y = -1.03\ln(x) + 36.2$	0.998	10	10 <sup>8</sup> – 10 <sup>3</sup>

## Application to environmental samples

Using the developed probe sets, we tested the application of the assay on genomic DNAs from four biological samples – a laboratory enrichment culture, an *in-situ* bioreactor and two different time measurements of a contaminated site undergoing bioremediation. These samples had been previously analyzed using regular qPCR assays for *Dehalococcoides* spp. and their Rdh genes *tceA*, *bvcA* and *vcrA*. Additionally, the presence of *Dehalobacter*, *Desulfotobacterium* and *Sulfurospirillum* had been detected in some of the samples using regular qPCR assays. The PRI-Lock probes and primers detected their target sequences in the environmental samples as had been expected based on the results from regular qPCR assays (Fig. 3) indicating



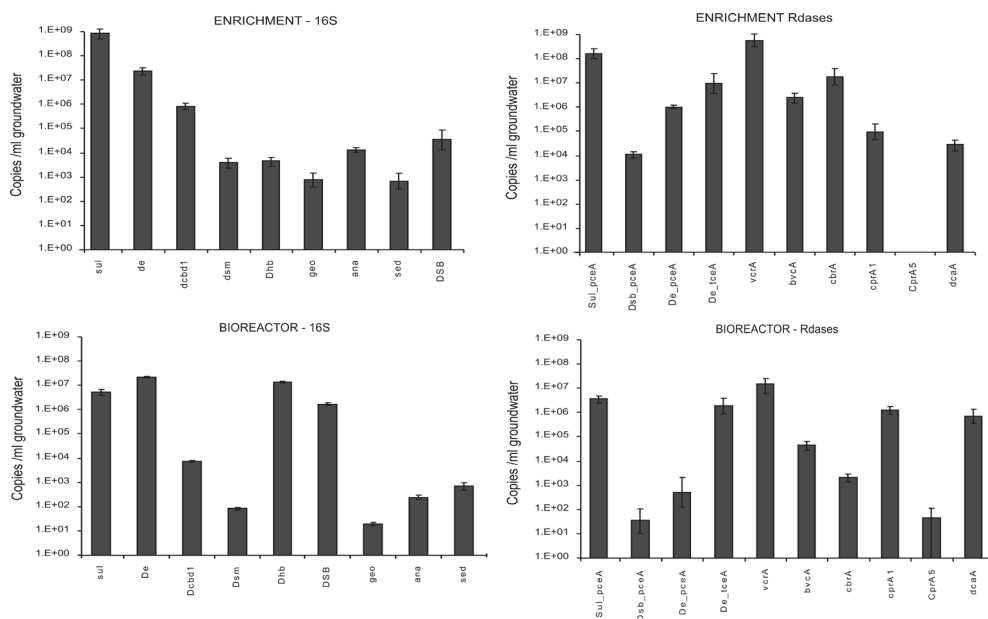
**Fig. 3:** Comparison of conventional qPCR (grey bars) and the ligation-based system (black bars) for quantitative analysis of *Dehalococcoides* spp. 16S rRNA gene and reductive dehalogenase genes (*tceA*, *vcrA* and *bvcA*). -De: *Dehalococcoides* Cornell subgroup, Dcbdb1 *Dehalococcoides* Pinellas subgroup.

that the detection system developed here is suitable for the reliable monitoring of OHRB in environmental samples.

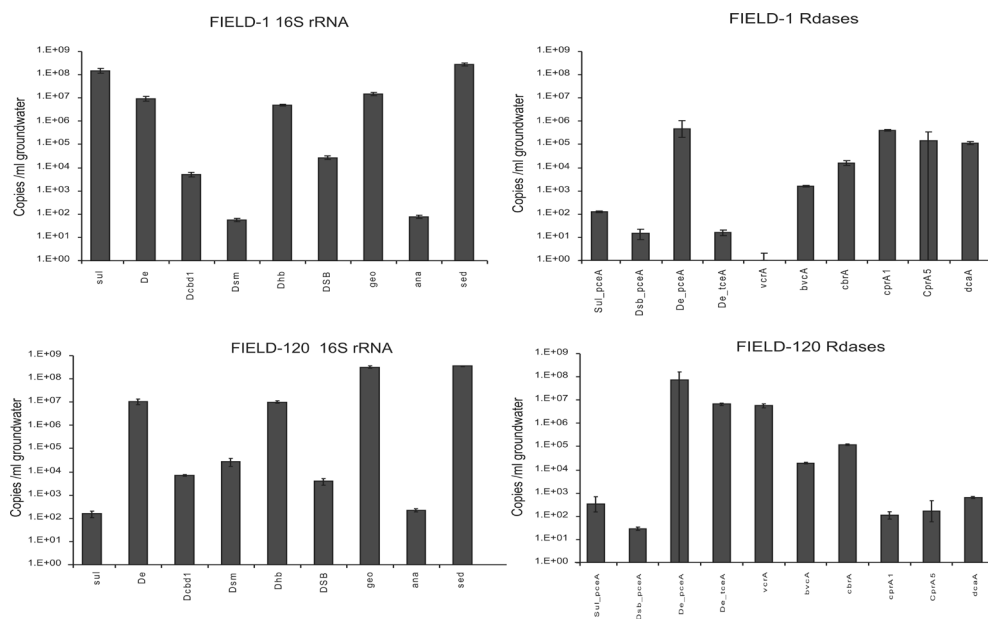
The results obtained with the multiplex ligation-based detection system were comparable to those obtained in regular qPCR analysis, with quantitative differences in only a few cases (Fig. 3). As an example, we detected *vcrA* and *bvcA* in the bioreactor at concentrations of  $10^7$  and  $10^4$  gene copies/ml sample, respectively, as had been found previously using the regular quantification assay (Chapter 4). For some of the targets, copy numbers/ml were found to be generally higher for the ligation-based detection system than obtained by regular qPCR assays. For example the concentration of the *tceA* gene as observed using the ligation-based detection was in general 2 orders of magnitude higher than with conventional qPCR estimations. *Dehalococcoides* spp. 16S rRNA gene copy numbers, like the *tceA*, were also higher by more than 2 orders of magnitude with the ligation-based detection system compared to those obtained by conventional PCR.

### Presence of other OHRB

Using the ligation-based circularization probes developed in this study, we were able to detect other phylogenetic groups of OHRB, including *Sulfurospirillum* and *Geobacter* that had not previously been quantified in the environmental samples. For example in the enrichment samples  $10^9$  16S rRNA gene copies/ml groundwater were detected for *Sulfurospirillum* (Fig. 4). Moreover, Rdh genes like *pceA* from *Sulfurospirillum* or *cbrA* from *Dehalococcoides* could also be detected in the biological samples. In the bioreactor sample, *Dehalobacter* was detected in agreement with data from regular qPCR. *Desulfitobacterium* was also detected at  $10^6$  gene copies/ml groundwater although it had not been detected by conventional qPCR.



**Fig. 4A:** Quantification of OHRB 16S rRNA and reductive dehalogenases genes using the ligation-based system in the Enrichment and Bioreactor samples.



**Fig. 4B:** Quantification of OHRB 16S rRNA and reductive dehalogenases genes using the ligation-based system on the field samples.

### Application of the ligation-based multiplex quantitative detection on the Biotrove OpenArray™ platform

Analysis of a single sample for 19 targets on a conventional qPCR machine is laborious, time consuming and expensive. Hence, it is desirable to have a multiplex assay for the large scale screening of samples at sufficiently high spatiotemporal resolution for robust monitoring of prospects and progress of bioremediation. We applied the multiplex assay developed in this study on the Biotrove OpenArray™ quantitative PCR system, which provides the possibility for the parallel analysis of up to 48 samples against all the 19 probes. Analysis on the Biotrove was done with SYBR Green qPCR mix and did not include specific calibration curves – hence, only relative comparisons are reported here (Table 5).

In general, we observed the same high specificity of probes also on the Biotrove OpenArray™. There was very low background signal in comparison to the amplification signal using conventional qPCR, and the detection of OHRB was specific to targets previously observed, demonstrating that multiplex assays are feasible on the OpenArray™ system.

The four field samples were also analyzed on the OpenArray™ to determine the potential for multiplexing in a single assay. Amplification signals were obtained for all the probes. Based on the melting curve analysis probes were specific to their target. One probe, Dhb\_pceA targeting the *Dehalococcoides pceA* gene, had some unspecific amplification in 4 of the replications of the negative control. Using Chi square analysis it could be determined, which of the measurements in unknown samples can be considered positive.

To get an indication of the assay performance, the CT values obtained with each probe on the OpenArray™ were compared to CT values obtained on the normal qPCR machine (Table 5). There are differences between the CT values and in some cases there is no correlation to what was expected. For example in the Enrichment culture, we expected to detect *Sulfurospirillum* and both *Dehalococcoides* Pinellas and Cornel subgroups but there was no signal in any of the replicates for *Sulfurospirillum* and *Dehalococcoides* Pinellas. Another difference is that the CT value on the OpenArray™ is often much lower (corresponding to higher copy numbers) than

**Table 5:** Comparison of CT Values obtained for the four samples on the OpenArray against those on the standard qPCR machine using SYBR supermix. N CT: number of replicates used to calculate the mean CT value. On the OpenArray™ each sample was analyzed for a total of 32 replicates.

Probe	Enrichment						Bioreactor						Field 1						Field 120					
	OpenArray			Standard qPCR			OpenArray			Standard qPCR			OpenArray			Standard qPCR			OpenArray			Standard qPCR		
	Mean CT	N	CT	Mean CT	N	CT	Mean CT	N	CT	Mean CT	N	CT	Mean CT	N	CT	Mean CT	N	CT	Mean CT	N	CT	Mean CT	N	CT
sulf	-	-	-	14.5	3	3	27.6	15	3	15.1	3	3	12.0	30	3	15.7	3	3	22.5	14	3	27.7	3	3
De	-	-	-	19.5	3	3	14.5	32	3	19.7	3	3	-	-	-	20.0	3	3	18.9	15	3	20.3	3	3
DCBDBI	21.1	9	30.3	30.3	3	3	19.4	32	3	30.6	3	3	23.8	29	30.4	3	3	9.6	30	30.6	3	3	3	3
Dsm	-	-	-	31.8	3	3	-	-	-	31.4	3	3	21.6	32	20.8	3	3	10.7	32	18.1	3	3	3	3
Dsb	-	-	-	20.2	3	3	-	-	-	16.7	3	3	25.8	27	19.1	3	3	13.4	32	21.9	3	3	3	3
Dhb	20.0	27	30.8	30.8	3	3	-	-	-	20.5	3	3	10.3	32	20.9	3	3	22.2	32	20.7	3	3	3	3
Ana	27.1	9	33.0	33.0	3	3	9.8	32	3	32.4	3	3	-	-	-	18.6	3	3	24.3	25	20.1	3	3	3
Geo	26.8	13	32.3	32.3	3	3	26.8	19	3	31.9	3	3	-	-	-	32.4	3	3	18.8	32	32.0	3	3	3
Sed	24.4	24	34.4	34.4	3	3	-	-	-	34.1	3	3	-	-	-	20.9	3	3	15.3	32	20.3	3	3	3
Sul_pecA	18.4	32	23.5	23.5	3	3	-	-	-	23.3	3	3	27.5	3	34.3	3	3	-	-	34.1	3	3	3	3
Dhb_pecA	27.1	22	30.7	30.7	3	3	22.7	31	3	25.2	3	3	23.8	18	31.9	3	3	10.1	29	26.9	3	3	3	3
Det_pecA	-	-	-	33.2	3	3	26.7	27	3	35.5	3	3	20.7	31	35.5	3	3	22.6	26	35.2	3	3	3	3
De_teeAB	-	-	-	20.2	3	3	20.5	30	3	20.1	3	3	25.3	18	35.2	3	3	13.7	32	21.3	3	3	3	3
vcrA	14.7	15	20.5	20.5	3	3	12.9	31	3	20.2	3	3	9.6	30	37.0	3	3	21.6	32	21.3	3	3	3	3
bvcAB	27.7	4	25.7	25.7	3	3	9.8	10	3	25.7	3	3	-	-	-	29.7	3	-	-	27.0	3	3	3	3
cprA1	26.5	6	33.3	33.3	3	3	-	-	-	25.0	3	3	13.4	32	25.9	3	3	21.0	31	36.5	3	3	3	3
cprA5	23.2	16	-	-	-	-	-	-	-	21.8	3	3	-	-	-	16.6	3	3	15.1	31	9.8	3	3	3
cbrA	-	-	-	26.9	3	3	18.0	6	3	32.2	3	3	-	-	-	29.6	3	3	14.3	32	28.0	3	3	3
dcA	18.2	32	31.9	31.9	3	3	15.6	30	3	24.2	3	3	-	-	-	-	-	14.5	32	-	-	-	-	-

observed on the normal qPCR machine. This can be seen for the DCBDB1 and Dhb probes, which gave CT values of 21 and 20 on the OpenArray™, while they were 30 on the normal qPCR machine. Analysis of the Bioreactor and Field samples show CT values that more in line with those from the normal qPCR machines.

To assess the potential of the assay for use in routine monitoring we compared the CT values obtained for the two time points of the Field Samples to see if the trends observed from qPCR on the normal machine were the similar. Basically most of the probes showed a similar trend. Some exceptions are for example the probes for *Anaeromyxobacter* (Ana) and *Geobacter* (Geo), which were not detected in the Field at day 1 but have a low CT after 120 days. On the normal qPCR machine the CT values for these probes had CT values of 18 and 20 respectively at day 1. Similar variations were observed for the Rdh genes.

## Discussion

The detection and quantification by real-time PCR of different populations and functional genes in a microbial community is important for good characterization of organohalide respiration potential. In this study we have developed a multiplex assay targeting taxonomic groups of OHRB and Rdh's, the key enzymes in reductive dehalogenation. The ability for a highly multiplexed assay allows for rapid screening of a large number of samples for multiple targets.

Using the ligation based detection system, high specificity could be achieved for the desired targets, whereas significant amplification signals were observed in the absence of target DNA. Specificity was furthermore demonstrated by the ability to distinguish between closely related strains like those of *Dehalococcoides*. It is important to distinguish between these strains because they exhibit different reductive dehalogenating capacities [32,33,124]. Taxon-specific detection of distinct groups of microorganisms in environmental samples is crucial whenever a key process is linked to these taxa [169]. To further aid in distinguishing between different *Dehalococcoides* spp. subgroups, it has previously been recognized that in addition to 16S rRNA gene-targeted detection, analysis of Rdh genes is also important [53,124]. In the test samples analyzed in the present proof-of-principle study, we were able to confidently detect the presence of either *bvcA*, *vcrA* or *cbrA* target



sequences. A presence of these genes would confirm that at least *Dehalococcoides* spp. strains BAV1, VS or CBDB1 are present, as confirmed by a positive signal from the respective taxonomic probe. The same can be done for *Desulfotobacterium* spp. strains confirming the presence of those strains, which harbor e.g. the *dcaA* gene or *cprA5* gene. The presence of multiple targets in the same sample did not influence the ligation efficiency or detection as previously demonstrated [165], and independent detection of targets could be achieved.

The detection limits of  $10^2$ - $10^3$  copies/ml groundwater obtained here are in line with what is commonly reported for conventional qPCR. Some examples of detection limits reported in qPCR detection of *Dehalococcoides* are  $10^2$  / g from aquifer material [144],  $10^2$  copies / g soil [170], and  $10^2$ - $10^3$  copies/L of groundwater [124,171]. In our assay, for example, the *Dehalococcoides* 195 + rel (Cornell subgroup) probe sensitivity was  $10^3$  and its associated Rdh gene probes *Det\_pceA* and *tceA* had a sensitivity of  $10^3$  DNA copies/ml groundwater. In contrast *Sulfurospirillum* probe sensitivity was  $10^2$  copies/ml groundwater while its associated *Sul\_pceA* probe detected  $10^4$  target DNA copies/ml groundwater.

Screening of environmental samples showed that the ligation detection system can indeed be used to quantify OHRB with confidence. The results obtained in the present study were relatively comparable to those observed previously by conventional qPCR. However, often gene copy numbers were 2 orders of magnitude higher for the ligation-based circularization system. This could be due to the higher sensitivity of the ligation-based circularization system in detecting specific gene targets, and also due to the Taqman qPCR as compared to the SYBR green qPCRs that were used in the conventional quantification analysis [172,173]. The Taqman PCR is specific to the circularized probe whereas the Sybr green assay depends on ability of primers to access the target DNA directly. In the ligation-based circularization assay there is an exonuclease step that is performed to remove any unligated DNA fragments that may interfere in the qPCR. Another reason for these differences maybe due to the DNA extraction, as a freshly obtained DNA extract from original sample was used for the analysis on the OpenArray. Although a standard protocol was used for all extractions, differences resulting from DNA extraction cannot be ruled out. For example it has been demonstrated that differences in *Dehalococcoides* BAV1

16S rRNA gene copy numbers were observed based on the DNA extraction method used [27,57]. PCR inhibition maybe another cause for the observed differences, as it has been reported especially in soil samples [48,174]. Furthermore, it should be noted that in the set-up used in this study, the internal ligation control can correct for potential inhibitions that can cause ligation variability [165].

#### *Analysis of environmental samples*

The enrichment samples contained *Dehalococcoides*, *Desulfitobacterium* and *Dehalobacter* spp., and all these genera were detected as present within this sample and in significant quantities. Interestingly *Sulfurospirillum* was also detected by this system, in line with clone library results from the enrichments which had *Sulfurospirillum*-like sequences [168]. The respective Rdh genes were also found to be present. Of interest is the detection of the chlorobenzene reductive dehalogenase *cbrA* gene from *Dehalococcoides* CBDB1, since this is the gene for the enzyme involved in degrading chlorobenzenes [31,135]. The culture was enriched on hexachlorobenzene and PCE.

The bioreactor profile showed higher numbers of *Dehalococcoides* of the Cornell group ( $10^7$  copies/ml) compared to the Pinellas *Dehalococcoides* ( $10^4$  copies/ml). The bioreactor had been bioaugmented with groundwater that had been reported to contain *Dehalococcoides* sp. strain VS [130], which harbors the *vcrA* gene in its genome. *Dehalococcoides* VS is within the Cornell group. In agreement with results from regular SYBR green qPCR  $10^7$  *vcrA* genes/ml groundwater were detected in the present study. There were higher numbers of *Dehalobacter* and *Desulfitobacterium* and the respective *pceA* and *cprA* genes detected in comparison to regular qPCR, suggesting that the assay was more sensitive in detecting these targets. Although the bioreactor was being optimized for growth and degradation activity in presence of *Dehalococcoides*, the presence of these genera is not surprising as they play a crucial role in degrading PCE to cDCE and are often found in dechlorinating consortia.

The potential for assessing temporal dynamics of OHRB could be demonstrated in the field samples. The contaminated site had been bioaugmented with a *Dehalococcoides* culture, and carbon source was added to further stimulate their growth. Although *Dehalococcoides* populations did not increase significantly,

we found a marked increase in the corresponding Rdh genes *tceA*, *vcrA* and *bvcA*. This suggests a change in the *Dehalococcoides* community over a period 120 days, although it should be noted that in the present proof-of-principle study, only a limited number of environmental samples has been analyzed. It is important to see such changes in the desired OHRB, in this case *Dehalococcoides*, in evaluating bioremediation strategies. Equipped with such information and kinetic modeling as performed in Chapter 4, management strategies for further optimization can be made at sites.

The multiplex assay enabled the detection of other OHRB's like *Geobacter*, *Desulfuromonas* and *Anaeromyxobacter*. These microorganisms have, in contrast to members of the genera *Dehalococcoides* and *Dehalobacter*, a versatile metabolism and are often found in diverse environments. *Geobacter* was detected in the Field samples. *Geobacter lovleyi* has been isolated for its ability to reductively dechlorinate PCE [127], which was the main contaminant at this site. *Desulfuromonas* spp. copy numbers increased during the 120 days monitoring period. Though it cannot be concluded that *Desulfuromonas* is actively dechlorinating at the site studied here, our data indicate that they accumulated during the biostimulation for dechlorinating species. This is important to know because competition for carbon source among dechlorinating bacteria can influence progress at bioremediation sites [128]. *Anaeromyxobacter* was not detected in significant quantities at any of the sites investigated here.

We could show that the ligation-based circularization system can be coupled to the OpenArray, however, more optimization will be needed to allow for accurate quantification of targets. The ligation-based circularization system has been coupled to the OpenArray™ system for the detection of plant pathogens [165] and thus it should be possible to fully adopt this assay for OHRBs. It is difficult to explain the variation observed between the two qPCR systems. Many qPCR parameters such as standard linear range, amplification efficiency and CT determination methods affect accuracy levels of an assay [46,48]. The CT values were generated on different qPCR machines and also different qPCR mastermix (SYBR green was used on the Biotrove). It has not been possible to use Taqman qPCR on the Biotrove until recently. Optimization of the Taqman assay will now be done and standard curves

for quantification of targets will then be possible on the Biotrove. Nonetheless using a relative quantification estimation we could establish expected trends for the Field samples where we observed a change in CT values according to expectations from data obtained on normal qPCR machines.

In conclusion, we have developed a multiplex assay that can be applied on the Biotrove Openarray. Using PRI-lock probes, multiplex detection of target OHRBs or Rdh-encoding genes can be done in a single reaction, and in combination with specific primers for each probe, qPCR can be done on the OpenArray platform. With further development and optimization of the Taqman qPCR protocol on the OpenArray it will be possible to develop calibration curves that can be used for absolute quantification in a highly multiplexed fashion. This will enable specific, high throughput, quantitative detection of multiple OHRB's over a wide range of target concentrations.

### **Acknowledgements**

This work was supported by the Ecogenomics Project of the Netherlands Genomics Initiative. We acknowledge Shakti Leiten from Bioclear for providing site data and samples used in this study.

# CHAPTER 6

## General Discussion and Future Perspectives

*A modified version of this chapter, with parts of chapter 1, is in press in Trends in Biotechnology*

The aim of the research described in this thesis was to unveil and explore the genetic potential related to the self-attenuating capabilities of soils with organohalide contamination using an integrated ecogenomics strategy. Ecogenomics involves multidisciplinary approaches capitalizing on genomics and related high throughput technologies applied to environmental systems (see **Chapter 1**). Below the obtained results that included functional metagenomics, molecular diagnostics and modeling studies applied to model and field systems, are summarized and placed in a general context. Moreover, an outlook is provided on the developments in the area of ecogenomics and how the present work has contributed to this and our understanding of bioremediation processes of soils contaminated with organohalides.

### **Polyhalogenated organic pollutants**

There is wide diversity of highly toxic and persistent chlorinated organics in the environment that require implementation of remediation techniques to clean them up (Table 1). Most of these contaminants accumulate in the anoxic subsurface of soil and groundwater systems and are recalcitrant to aerobic degradation [175]. Halogenated hydrocarbons with a high degree of substitution are generally more readily degraded under anoxic conditions. This and the fact that most of the contaminated sites such as aquatic sediments, submerged soils and ground water are oxygen depleted, makes anaerobic bacteria that are capable of organohalide respiration, good candidates for bioremediation [176]. Organohalide respiration (previously also termed (de) halo-respiration or respiratory reductive dehalogenation), the use of halogenated compounds as terminal electron acceptors in anaerobic respiration, is a key process for their degradation in the anoxic subsurface [15].

### **Options for the remediation of organohalide-polluted environments**

Organohalides like tetrachloroethene (PCE) and trichloroethene (TCE) are Dense Non Aqueous Phase Liquid (DNAPLs). Their low solubility and high density makes them challenging to remediate, and a common method that has been used to treat subsurface water contaminated with DNAPLs is pump and treat [177]. Pump and treat remediation consists of pumping out the contaminated water from the aquifer and treating it *ex situ*. A major drawback to this method is that the source

**Table 1:** Properties of organohalides. Only the chlorinated compounds are shown. Chlorophenols were used in laboratory culture expression studies of the *Dehalobacter-Sedimentibacter* coculture (Chapter 2) and *Desulfotobacterium hafniense* DCB-2 (Chapter 3). Tetrachloroethene (PCE) and its derivatives were the main contaminants of the field study (Chapter 4), and were tested as potential substrates for reductive dehalogenases of *Dehalobacter* sp. strain E1 (Chapter 2).

	Ortho-, meta-, para-CP <sup>a,b</sup>	2,4-DCP <sup>b,c</sup>	2,4,6-TCP <sup>b,c</sup>	PCP <sup>a,b</sup>	PCE <sup>d</sup>	HCB <sup>a</sup>	PCBs <sup>a</sup>
<b>Occurrence</b>	Mostly bound to soil particles, mineral soils.	adsorption is strong in acidic,	organic soils,	but low in basic,	Mostly in vapour phase	Mostly in vapour phase	Lipophilic semi-volatile, mostly absorbed to particles in air, soil.
<b>Abiotic degradation</b>	Photodegradable ( $t_{1/2}$ varies)	Degradation in minutes by UV irradiation in aqueous solution	Degradation by UV irradiation on silica gel ( $t_{1/2}$ 17 h).	Photodegradable ( $t_{1/2}$ 14 days in soil).	Persistent (no photodegradation in water).	Persistent ( $t_{1/2}$ is 1.5-2 years in the atmosphere).	Persistence with halogen substitution ( $t_{1/2}$ between 10 days to 1.5 year).
<b>Natural production<sup>e</sup></b>	Metabolites of microbes, fungi (2,4-DCP is produced by <i>Penicillium</i> sp.), ticks (2,6-DCP secreted as sex pheromone), and grasshoppers (2,5-DCP as repellent); can occur in proteins (chlorinated tyrosine in proteins from marine sponges).				Volcanic activity, barley	Halogen containing minerals	Indirectly by volcanic activity.
<b>Anthropogenic Production</b>	Wood preservation, inhibitors of microbial growth in oils, textiles, and pharmaceutical products.	Intermediate in 2,4 D synthesis, used as repellent in mothballs.	Germicide, preservation of wood, leather, glue, and textiles.	Wood preservation by-product of water treatment.	Solvent in metal and dry cleaning, used in grain fumigation.	By-product in pesticide synthesis, intermediate. In dye production.	Dielectric and heat-exchange fluids, also used in microscope oil.
<b>Toxicokinetics</b>	Readily absorbed across the skin and the gastrointestinal tract, accumulation in liver and kidney. Detoxification by conjugation to glucuronates or sulphates in the liver.				Rapid absorption by skin or lung, accumulation in fatty tissues.	High absorption by the digestive tract, low dermal intake. Mainly catabolized in the liver.	Rapidly absorbed in fatty tissues.
<b>Systemic effect</b>	Generally, toxic effect of chlorophenols involves the uncoupling of mitochondrial oxidative phosphorylation.	Increased spleen and liver weights and haematological or immunological effects, one death incidence after dermal exposure.	Leukaemia and liver cancer in rats.	Alterations in hepatic enzyme activities, increased liver weight, histopathological and carcinogenic effects.	Dermal irritation, liver and kidney damage, alterations in metabolism, long term neurotoxic effect, risk of cancer.	Effect on immune function, liver activity, cancer, death (Turkey 1955-1959).	Introduction of cytochrome P450 (ligand for AhR), endocrine disturbances ligand for TR $\beta$ respiratory tract toxicity, cancer.

<sup>a</sup> (WHO, 2003); <sup>b</sup> (WHO, 1989); <sup>c</sup>(ATSDR, 1999); <sup>d</sup> (WHO, 1984); <sup>e</sup> (Gribble, 1994); <sup>f</sup> (Schriks et al., 2006). Abbreviations: CP, chlorophenol; DCP, dichlorophenol; TCP, trichlorophenol; PCP, pentachlorophenol; PCE, tetrachloroethene; HCB, hexachlorobenzene; PCBs, polychlorinated biphenyls; 2-4D, 2,4-dichlorophenoxyacetic acid; AhR, aryl hydrocarbon receptor; TR, thyroid hormone receptor.

of contamination and the chlorinated ethenes adsorbed to the soil are not treated. Although large masses of contaminants may be removed, the bulk pure phase DNAPL remains absorbed to the soil *in situ* and serves as a continuous polluting source. Additional physicochemical methods for remediation include surfactant flushing, soil vapor extraction and reduction by metallic iron [178].

In contrast to these physicochemical approaches, bioremediation is the use of living organisms as the catalysts for detoxification reactions that result in pollutants clean up. Some of the different processes by which bioremediation can be implemented are phytoremediation, natural attenuation, biostimulation, and bioaugmentation.

Phytoremediation, the use of plants in bioremediation enables the mitigation of an environmental problem without the need to excavate the contaminant material and dispose of it elsewhere. Phytoremediation may be applied wherever the soil or static water environment has become polluted or is suffering ongoing chronic pollution and has been successfully employed in the restoration of abandoned metal-mine workings, reducing the impact of sites where polychlorinated biphenyls have been dumped during manufacture and mitigation of on-going coal mine discharges. Contaminants such as metals, pesticides, solvents, explosives, crude oil and its derivatives, have been mitigated in phytoremediation projects worldwide [179,180]. Plants such as mustard plants, alpine pennycress and pigweed have proven to be successful at hyperaccumulating contaminants at toxic waste sites [179]. As such phytoremediation is considered a clean, cost-effective and non-environmentally disruptive technology, as opposed to mechanical cleanup such as soil excavation or pumping polluted groundwater. Over the past 20 years, this technology has become increasingly popular and has been employed for soils contaminated with lead, uranium, and arsenic [181]. However, one major drawback of phytoremediation is a requirement for long-term commitment as it is dependent on plant growth, tolerance to toxicity and bioaccumulation capacity.



A passive form of bioremediation is natural attenuation, which is basically described by the Environmental Protection Agency EPA as “Naturally occurring processes in soil and groundwater environments that act without human intervention to reduce the mass, toxicity, mobility, volume, or concentration of contaminants in those media. These *in situ* processes include biodegradation, dispersion, dilution, adsorption, volatilization, and chemical or biological stabilization or destruction of contaminants.” This requires that the relevant microbial reactions and microorganisms are present as an important component in natural attenuation [182]. However natural attenuation is a slow process and intervention is often required for enhanced and effective cleanup of contaminated soils.

Often at contaminated sites the microorganisms of interest are present, in very low numbers or may not be metabolically active due to unfavourable aquifer's conditions, such as elevated redox potential or lack of nutrients. In such cases biostimulation may be applied by the addition of nutrients, electron acceptors and electron donors, to stimulate the growth of these organisms. Biostimulation has been successfully applied in cleaning sites with organohalide contaminants [144]. Biostimulation can e.g. be used to create bioreactive barriers stopping the spread of pollution plumes [144].

At times microorganisms of interest are not detected at a site that has been considered for bioremediation. The most appropriate remediation strategy then obviously involves the introduction into the site of microorganisms that are able to perform the detoxification reactions of interest. Firstly microorganisms are grown in reactors *ex-situ*, and once a required density and metabolic activity is reached they are then introduced into the soil. For example several organohalide respiring cultures are now being used for bioaugmentation purposes. These include Bioclear *Dehalococcoides* Consortia (Bioclear BV, The Netherlands), KB-1 Culture [183,184], and the Bachman Road Culture [133,144]. In general, the benefits of bioremediation have awakened research focused on populations that are able to detoxify widespread compounds, ranging from chlorinated ethenes to polyhalogenated compounds like polychlorinated dibenzodioxins (PCDDs).

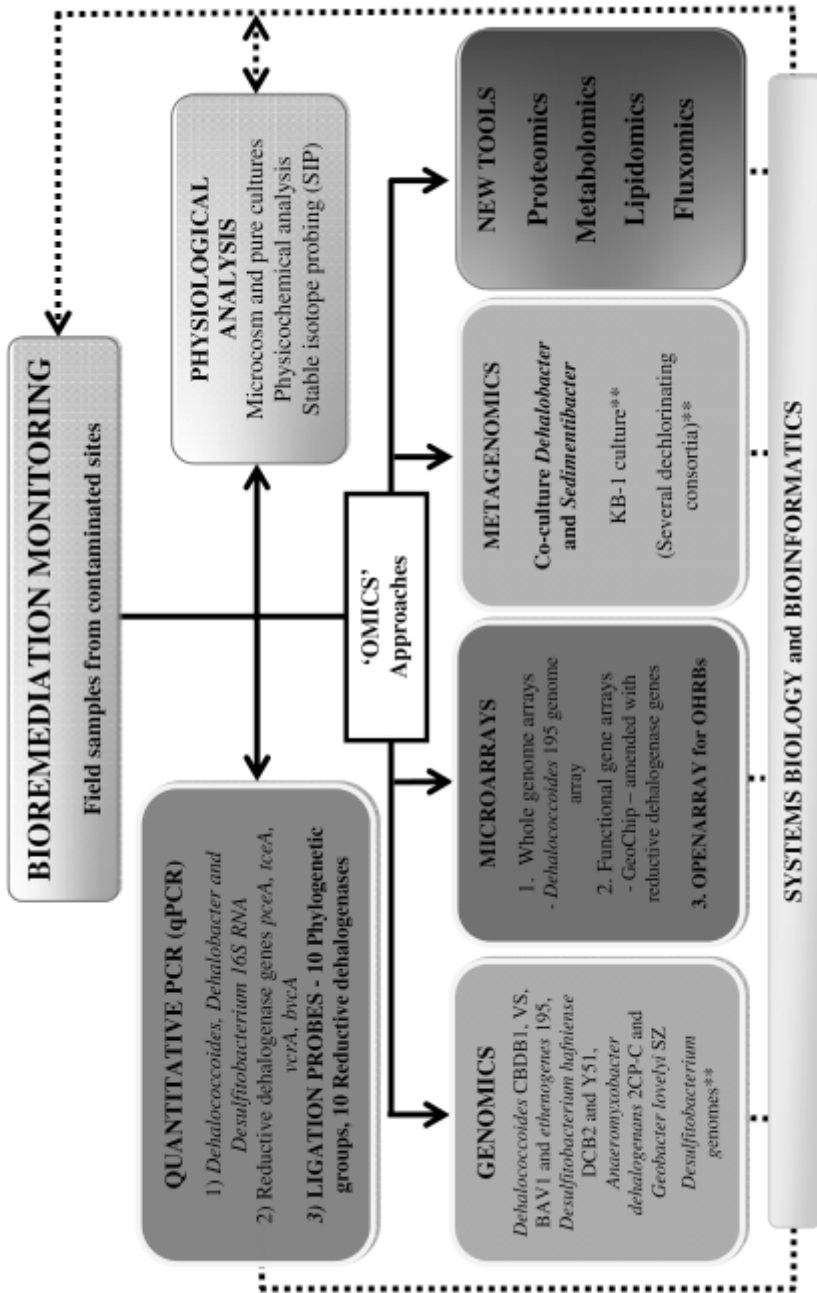
*Bioremediation technologies in this study*

This study was conducted at sites where biostimulation and bioaugmentation were employed as bioremediation strategies. Biostimulation was achieved by injection of carbon source such as molasses directly into the plume. *In-situ* dechlorinating bioreactors were installed at the sites to bioaugment them with *Dehalococcoides* spp. that are capable of degrading dichloroethene and vinyl chloride (VC) to ethene. It had been observed that PCE was degraded with concomitant accumulation of the more toxic VC, resulting in the need to intervene via biostimulation and bioaugmentation. Effective design and realization of intervention strategies, however, require knowledge of catabolic potential and *in-situ* dynamics of organohalide-respiring and co-metabolizing microorganisms to be stimulated, or added, to the polluted site.

**A need for an ecogenomics toolbox towards more effective bioremediation**

The inception, of ‘omic’-based techniques has revolutionized the field of microbial ecology and is shifting the focus towards the functional characterization of complex microbial ecosystems. To this end, **Chapter 1** provided an overview of the current state of the art as to how ecogenomics approaches have aided development of high-throughput-techniques for detection and monitoring organohalide-respiring bacteria (OHRB), and for providing improvements to selection, specificity and sensitivity of target biomarkers, including their application to evaluate bioremediation strategies (Fig. 1). The use of these molecular and biochemical approaches has provided a more effective and direct assessment of the microbial communities involved in bioremediation beyond classical microbiological techniques.

Bioremediation of sites requires an understanding of the microbial community and its response to a given treatment, in addition to information obtained from the assessment of geochemical parameters. Obtaining information can be a costly exercise due to site-specific environmental conditions, operative mechanisms and rates of bioremediation. The incorporation of ecogenomic tools now allows us to open the “black box” between the environment and bioremediation processes, like organohalide respiration, and enabling us to understand the mechanisms behind such adaptations to toxic organohalide compounds [6]. Ecogenomics defines phylogenetic and functional biodiversity at the DNA-, RNA- and protein levels [7-10], in order

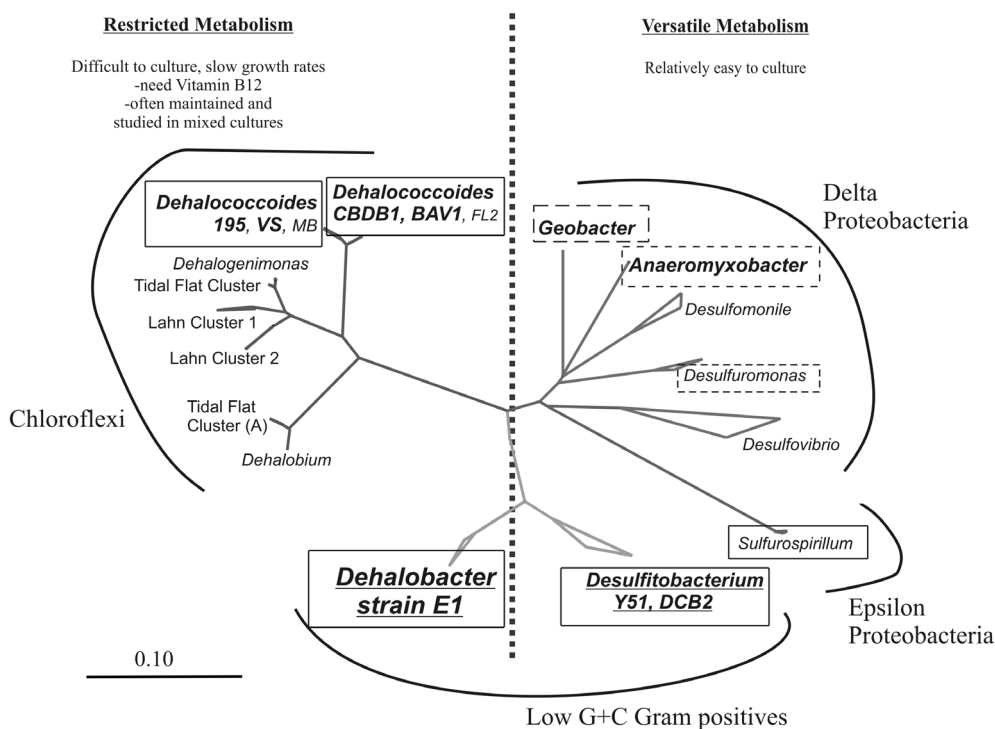


**Fig. 1.** The 'ecogenomics toolbox': current set of techniques in use (i.e. genomics, microarrays, metagenomics) and those under development for use ('new tools') in monitoring bioremediation. \*\*Denotes genomes and metagenomes expected in 2010.

to elucidate the functions and interactions of organisms at an ecosystem level in relation to ecological and evolutionary processes [11,185]. It answers environmental questions providing crucial data that can help manage and optimize remediation strategies.

### **Functional genomics of microbial consortia – The way ahead**

The use of next generation technology (NGT) sequencing is further revolutionizing biology. As sequence throughput has now reached the tera base (1000 Gbase) level in single papers and is still expanding, the analysis of such complex systems as the soil microbial ecosystem comes within reach. For example a microbial gene catalogue for the human gut has recently been described through metagenome sequencing [186]. The living soil is the most biodiverse environment on earth: it is estimated to contain approximately 1,000 Gbp of microbial genome sequences per gram of soil and thus metagenomic sequencing of soil remains rudimentary and constitutes a new and ambitious challenge, now met by the recently established TerraGenome initiative [187]. Pyrosequencing using the Roche 454 instrument, which has further improved with the titanium technology, is currently considered one of the most robust and reliable NGT sequencing approaches and hence, as is described in **Chapter 2**, we decided to determine the genome of a representative of the genus *Dehalobacter*, an important member of the OHRB that had not previously been characterized at the genome level (Fig 2). This *Dehalobacter* sp. strain E1 grows in coculture with *Sedimentibacter* sp. strain B4. This coculture was the first known culture capable of metabolic dechlorination of  $\beta$ -HCH under anaerobic conditions [22]. The coculture grows better than available pure cultures of *Dehalobacter* and can serve as a better model for gaining insight into the *in situ* situation than isolates. This is important in view of the applicability of the ecogenomics toolbox, including transcriptomics and proteomics, to elucidate physiology and to provide the necessary (kinetic) data for metagenome-scale metabolic modeling of this intriguing syntrophic partnership. Such models can be expected to be instrumental in providing novel leads towards optimized bioremediation strategies, e.g. by identification of more selective substrates that can be added *in situ*.



**Fig. 2:** Phylogenetic tree of OHRB based on bacterial 16S rRNA sequences. Alignment and phylogenetic analysis were performed with the ARB software and tree was constructed using the neighbor-joining (NJ) method [26]. The reference bar at the bottom center indicates the branch length that represents 10% sequence divergence. Bold face shows strains with full genome sequence available including *Dehalobacter* sp. strain E1. Strain names shown only for organisms of specific interest to this study. Rectangles show genera for which primers and probes for ligation-probe based analysis have been developed for both 16S rRNA and reductive dehalogenase genes. Dashed boxes –only ligation probes for 16S rRNA have been developed. Underlined genera include the main study organisms in this research.

Building on the consortium genome elucidated in the framework of this thesis, we expanded our knowledge on the reductive dehalogenation potential of *Dehalobacter* sp. E1. Genome based prediction showed that strain E1 had genes predicted to encode reductive dehalogenases (Rdh's) most similar to characterized PCE and chlorophenol dehalogenating enzymes. It had been shown previously that the coculture could not degrade PCE and DCA [22]. We therefore assessed the potential

for degradation of TCE and DCE and several chlorophenols. The coculture was able to degrade TCE but not any of the other compounds tested. Further biochemical and degradation studies for this TCE degradation can now be conducted.

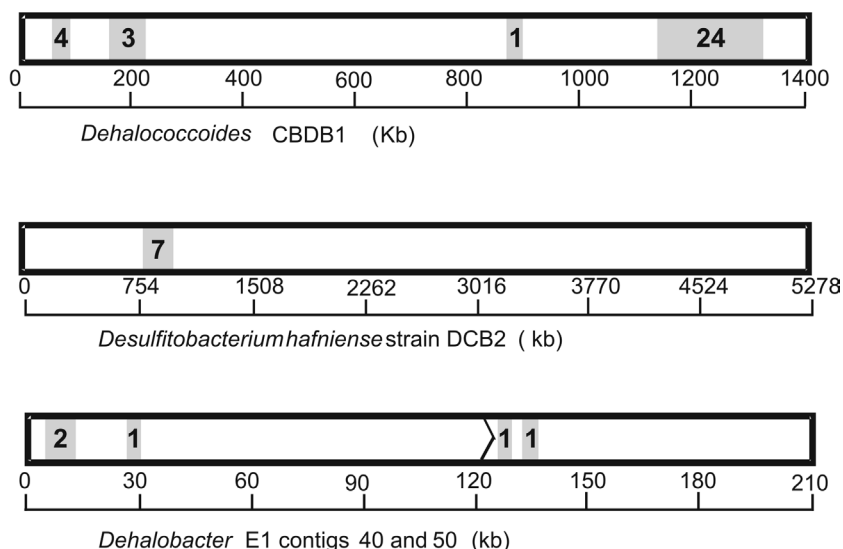
This first genome insight into the *Dehalobacter* genus will now enable for further comparisons with genomes from other sequenced OHRB. Currently eight genomes have been sequenced for *Anaeromyxobacter dehalogenans* [23], *Desulfitobacterium hafniense* strains Y51 and DCB-2 [24] and *Dehalococcoides ethenogenes* strain 195 [10], *Dehalococcoides* sp. strain CBDB1 [7] and *Dehalococcoides* sp. strains BAV1 and VS [3] (Fig. 2). Additional genomes of other isolates are currently being elucidated (<http://www.genomesonline.org>). Furthermore, several metagenomes of consortia with *Dehalococcoides* spp. as dominant organohalide respirers are being sequenced (Fig. 1). Whereas much progress has been made over the past years with respect to genome sequencing of *Dehalococcoides* spp. (four complete genomes, two draft genomes) and subsequent exploration by transcriptomics and proteomics [123,188], no efforts have been undertaken previously towards a systematic exploration of genomes of organohalide respirers belonging to the genus *Dehalobacter*, closely related to the genus *Desulfitobacterium* within the *Firmicutes*. Whereas *Desulfitobacterium* spp. are versatile in their spectrum of electron donors and acceptors, *Dehalobacter* spp. are strongly restricted in their energy metabolism, being dependent on respiratory reductive dehalogenation with hydrogen as electron donor in syntrophic interaction with *Sedimentibacter* spp.

The availability of the *Dehalobacter* and *Sedimentibacter* metagenome as reported in here can be expected to stimulate systematic studies to understand the respiratory landscape of these organisms. Secondly the functionality of organohalide respiring *Firmicutes* at increasing levels of complexity, ranging from the intricate molecular interactions inside individual cells to the interaction networks of syntrophic partners in their natural environment can now also be studied. Bioremediation of environmental pollutants depends on these intricate multispecies metabolic networks that are present in a polluted scenario. The inherent complexity of such microbial systems, as well as their interaction with their biological and physical-chemical environment is growingly amenable to the conceptual frame and the

growing tool box of systems biology [11,185]. To this end, the availability of the metagenome sequence of the defined *Dehalobacter-Sedimentibacter* partnership, used as a model for syntrophic microbial consortia that are pivotal to the clean-up of subsurface anaerobic environments polluted with halogenated organic compounds, can guide interventions for stimulating the performance of desired biodegradation processes. This can include identification of potential alternative substrates, leading to improved strategies for efficient in situ stimulation and maintenance of these dedicated degraders.

The repertoire of reductive dehalogenase genes in *Dehalobacter* is different from those found in *Dehalococcoides*. The identified gene products, key enzymes in organohalide respiration, cluster with those from various *Desulfitobacterium hafniense* strains. It has been suggested that multiple events of homologous recombination, as well as horizontal gene transfer have been, and still are, involved in the generation of strain-specific sets of multiple reductive dehalogenase gene loci in individual strains that have significant differences in their substrate spectra for dehalogenation [3,7,10,24]. Comparative genomics between *Dehalobacter* and *Dehalococcoides* genomes will help give insight into the evolutionary mechanisms underlying this dedication to organohalide respiration by these two intriguing yet phylogenetically distant bacteria. For example identification of either common or lacking biosynthetic pathways between these two genera might give us clues to this dedication. Genome dynamics can also be studied to find indications of evolutionary mechanisms of adaptations towards organohalide respiration. For example comparison of the four *Dehalococcoides* genomes showed that the *rdhAB* genes are compartmentalized within specialized high plasticity regions close to the origin of replication and may enhance opportunistic adaptations to respiratory niches in the anoxic subsurface [3]. Similarly, in *Desulfitobacterium hafniense* DCB2 reductive dehalogenases are also concentrated in the vicinity of each other. A similar situation can probably be expected, as two of the large *Dehalobacter* contigs (contig 00040 and contig00050) both have two *Rdh* gene clusters within a 30 kb region (Fig. 3). It will be interesting to confirm whether the complete genome of *Dehalobacter* also has similar *Rdh* gene arrangements. To this end, generation of a complete genome sequence of *Dehalobacter* sp. E1 seems feasible to achieve, even from a

metagenome, as was demonstrated by the closure of genomes from the acid mine drainage microbiome [189], also considering recent developments towards routine



**Fig. 3:** Localisation of reductive dehalogenase genes in representative genomes of OHRB. Cartoon maps of the *Dehalococcoides* CBDB1 and *Desulfotobacterium hafniense* DCB2 genomes showing the position of reductive dehalogenase clusters, and 2 contigs of *Dehalobacter*. Numbers represent number reductive dehalogenases in close vicinity. (NB *Dehalobacter* contigs are shown on a different scale for visualization)

single cell genomics [190].

Lastly the presence and role played by *Sedimentibacter* has to be further elucidated. A systems biology approach to studying the coculture may open up this “blackbox” showing which pathways or factors the *Sedimentibacter* produces, and may have implications for optimising bioremediation. *Dehalobacter* sp. E1 grows strictly in the presence of the *Sedimentibacter* sp., however, the mechanism of this obligate syntrophy is still to be alluded. Comparative genomics and proteomic analysis can help us answer these questions. The *Sedimentibacter* sp. strain B4 genome harbours genes for pathways involved in carbohydrate fermentations and vitamin biosynthesis. Proteomic studies of the coculture can now be conducted in the presence of different carbon sources. Furthermore, cocultures as the one studied in this thesis are routinely maintained in the presence of a high concentration of Vitamin



B12 in the media, as cobalamin is an essential cofactor for functional reductive dehalogenases [22,102]. Further analysis of the pathways in these two organisms can help show the reasons why this is so and give indications as to how this can be overcome even at bioremediation sites. Strain-resolved comparative proteomics has been shown for an acidophilic microbial community [191] and such experiments can help in identification of proteins that are produced by each of these two syntrophic partners during growth.

### Functionality of reductive dehalogenases

OHRB dechlorinate several different chlorinated compounds, however, knowledge as to which reductive dehalogenase is responsible for which specific conversion is still limited. **Chapter 3** describes the differences in induction and transcriptional patterns for *cprA1* to *cprA5* reductive dehalogenase genes in *Desulfitobacterium hafniense* DCB-2 cells cultured in the presence of 2,3-dichlorophenol and 2,4-dichlorophenol as model substrates for the *meta*- and *ortho*- dechlorination positions. Clear differences in the induction patterns were observed. In the presence of 2,4-DCP, *cprA1* and *cprA5* were strongly induced, while 2,3-DCP induced *cprA3* and *cprA5*. Our results suggest that the gene product of *cprA5* has a broad range substrate specificity whereas those for *cprA1* and *cprA3* have a more restricted activity towards *ortho*- and *meta*-chlorophenols, respectively. The genes *cprA2* and *cprA4* for which transcription had been reported to be induced by both *ortho*- and *meta*-chlorophenols were not induced by these substrates, however, *cprA4* appears to be constitutively expressed. Similarly in *Dehalobacter* (**Chapter 2**) we observed differential expression patterns for the reductive dehalogenase genes during growth of the coculture on  $\beta$ -HCH. This expression of multiple reductive dehalogenase-encoding genes has also been reported in *Dehalococcoides* spp. [38,124,126]. Expression of multiple reductive dehalogenases probably allows for the adaptation of these OHRB to the different compounds and breakdown products that they dechlorinate. As more studies are carried out with transcriptomics and proteomics at increasing sensitivity, it shall soon become clear why there has been such an accumulation of genes in these bacteria. Also the regulatory networks involved in transcription of these genes in the presence of different chlorinated compounds as well as potential alternative electron acceptors

will be important at least for optimizing biostimulation at bioremediation sites.

### **Environmental diagnostics of OHRB**

In studying bioremediation potential *in situ* testing for OHRB and respiration is important. A challenge is the very low abundance of organohalide respirers in the environment. They are known to constitute less than 1% of the microorganisms present in samples [53]. Organisms do not grow to high density even under laboratory conditions, posing a detection challenge when monitoring them in field samples from bioremediation sites. There is still a knowledge gap about the diversity, distribution and functioning of *Dehalococcoides* and other groups of OHRB in different environments despite their presence at several contaminated locations. Highly sensitive and specific detection methods are therefore needed to identify and monitor OHRBs. To this end we used quantitative PCR (qPCR) methods to detect and quantify OHRB in environmental samples (**Chapters 4 and 5**). Molecular profiling of *Dehalococcoides* in a dechlorinating bioreactor was done and coupled to kinetic model simulations (**Chapter 4**). Overall, a complex interplay within dechlorinating microbial communities, including the co-existence of several populations of *Dehalococcoides* spp. strains was observed [72,129,146]. A key conclusion from this study is that modeling of the microbial and metabolic dynamics of dechlorinating populations is important for optimizing performance of cleanup strategies for contaminated plumes beyond information provided by molecular diagnostic endpoints. The estimation of *Dehalococcoides* growth yields in the bioreactor suggested the presence of biofilm formation, which need to be considered in design and monitoring of bioremediation strategies. Biofilms of dechlorinating bacteria have been observed to develop predominantly inside pores of support materials, with no or little biofilm formation on external surfaces [153]. Considering the design and operation parameters of the bioreactor in this study, it is likely that the microorganisms form these biofilms in the porous granular material allowing them to form the stable microbial communities observed. In a membrane-biofilm reactor study, the biofilm community was shown to adapt to TCE and increased in numbers of *Dehalococcoides* spp. and *tceA* and *bvcA* [147]. Taking biofilm formation into account is important because efficient delivery and distribution of

organisms and optimum concentrations of electron donors to the subsurface remains a big challenge, especially in heterogeneous settings [6]. Additionally, the complex interactions of multiple contaminants on *Dehalococcoides* and other OHRB needs to be further studied in greater detail to give confidence to implementation of strategies at mixed waste sites.

Bioremediation management strategies require monitoring technologies that enable simultaneous detection and quantification of a wider range of microorganisms involved in reductive dechlorination. Many multiplex, quantitative detection methods available suffer from compromises between the level of multiplexing, throughput and accuracy of quantification. In **Chapter 5** we demonstrate the development and application of a high-throughput, ligation based assay for simultaneous quantitative detection of multiple reductive dechlorinating populations and their key reductive dehalogenases. The ligation probes (LCR) are long oligonucleotides with target complementary regions at their 5' and 3' ends. Upon perfect target hybridization, probes are circularized via enzymatic ligation, subsequently serving as template for individual, standardized amplification via unique probe-specific primers. Specificity is achieved in this ligation step at the 3' arm of the probe is shorter and has the specific match for the desired target sequence. The accurate identification and detection of OHRB and their key genes is important for evaluating management studies or querying field sites for reductive dehalogenation potential.

To increase efficiency of *in situ* detection and screening of many samples to achieve the required spatiotemporal resolution for robust field characterization, it is desirable to have a high throughput multiplex assay. Microarrays may enable highly parallel detection of diverse organisms. Presently microarrays are mainly used in microbial ecology to infer the composition of microbial communities or of defined guilds, and are limited by the expense in the technological aspects and setup from becoming a routine tool to be used in high throughput monitoring. Examples of microarrays used in soil studies include the Phylochip and GeoChip [62,192-195]. As alternative new multiplex microfluidics platforms capitalizing on the ease and relatively low cost of qPCR compared to microarrays are being developed as a favorable technique for fast analysis of microbial numbers and indicators of bioremediation potential from in-situ samples. To this end, we developed the Ligation

probe based detection system for use on the Biotrove OpenArrays™, which can accommodate up to 3072, 33 nl PCR amplifications, allowing high-throughput real-time quantification. This assay combines the multiplex capabilities and specificity of ligation reactions with high-throughput real-time PCR in the OpenArray™, resulting in a flexible, quantitative multiplex detection system for routine environmental monitoring. Ligation probes targeting 16S rRNA genes from nine phylogenetic groups involved in reductive dechlorination and ten reductive dehalogenase genes were used to demonstrate the application of the detection system. The probes specifically detected their corresponding targets and provided discrimination against non-target organisms with very similar ligation target sites. Comparable results were achieved for detecting organohalide respirers from environmental samples. Observed target gene copy numbers were, however, one to two fold higher with the Ligation probes compared to conventional qPCR. This might be due to the fact that the fundamental detection mechanisms and the associated limitations on quantitative analysis differ between these two methods. The primer binding sites in conventional qPCR were different from those of the Ligation probes. Secondly the conventional qPCR's were done using a SYBR Green protocol whereas the Ligation probe system uses Taqman qPCR chemistry. It has to be realized that qPCR as a highly sensitive technique is prone to errors and validity of resulting data sets has to be considered with regard to specificity of primers, efficiency in DNA extraction methods and errors arising from PCR methodology and instrumentation [46]. The Ligation probe based system, which was shown to be highly reproducible, should be easily adaptable for versatile detection and monitoring purposes based on other functional genes involved in bioremediation and probe sets can be easily expanded giving greater multiplexing possibilities. Cost efficiency is achieved by using nanoliter quantities on the OpenArray™ compared to microliter volumes in conventional PCR. Universal primer pairs and arrays may also be used further reducing costs.

### **New tools for the ecogenomics toolbox – Future perspectives**

In addition to genomics, transcriptomics and proteomics, several emerging approaches, such as lipidomics, metabolomics, fluxomics, glycomics, are increasingly

being applied in environmental studies (Fig. 1). These methods require advanced equipment for mass spectrometry and high-resolution chromatography, and have been developed for applications in related fields [185,196].

Proteomic approaches, though still in their infancy in studying OHRB, are increasingly being applied with the aim to elucidate the metabolism of organisms that cannot be grown in amounts that are sufficient to perform standard biochemical analyses, such as enzyme purification and activity measurements [197]. Proteomics has started to play a crucial role in identifying Rdhs that are produced during dehalogenation of different substrates and enzymes involved in the utilization of additional substrates by *Dehalococcoides* [42,197]. Non-gel based proteomics has been used to demonstrate that peptides from functional enzymes, including Rdhs and other putative respiratory enzymes, might be used as quantitative biomarkers with high specificity and sensitivity to differentiate closely-related strains in cultures and environmental samples [188,198]. Proteomics will further aid in functional characterization of the large diversity of Rdhs encoded in OHRB genomes, and is emerging as a powerful tool in bioremediation, allowing studies of protein–protein interactions and cell surface proteomics in discovery of new biomarkers (genes and proteins) as well as giving insight into metabolic pathways of dechlorination not clearly understood [185,199].

Lipidomics holds promise for a systems-based study of a wide range of lipids [196] that may eventually be used for routine profiling of dechlorinating microbial communities at contaminated sites. Integrated studies combining proteomics and metabolomics can be predicted to soon take a central position in the chase for OHRB in the postgenomic systems biology era. Ability to decipher and predict important processes and parameters will aid site remediation by answering the “when” and “how” questions of OHRB activity. Continuing improvements in instrumentation will likely result in increased integration of high throughput tools applied in studying these intriguing microorganisms. Nevertheless, various sources of variance will continue to influence the data, including methodological, factorial and error sources [8]. Furthermore, the large data sets generated by the above-mentioned approaches require advanced bioinformatics systems. A key challenge of bioinformatics will be the generation of databases relevant to processes and microorganisms involved in

organohalide respiration. With new genome and metagenome sequences being made available at ever increasing speed [200], such databases are foreseeable sooner rather than later, providing the necessary basis for ecosystems biology efforts towards descriptive and predictive models of OHRB functioning in polluted environments [11,12].

The chase for OHRB for use in bioremediation continues to advance. Additional OHRB isolates are and will be identified and characterized based on the insights on lifestyles of OHRB gained from their genomes and metagenomes. As our understanding of the biology of organohalide respiration increases, we are now getting answers from ecogenomics approaches to improve and expand the biomarkers we use in monitoring bioremediation strategies at sites with multiple contamination. With developments in proteomics and metabolomics in combination with genomic information correlations between in-situ detection and quantification of these biomarkers to determine activity will become more robust and reliable. Increased in-depth understanding and exploitation of organohalide respiration, of dedicated degraders and of corresponding microbial consortia, requires collaborations between microbial physiologists, molecular ecologists and bioinformaticians, modelers as well as input from geneticists and evolutionary biologists.

# References

## References

1. GeoSyntec, C. (2005) Bioaugmentation for remediation of chlorinated solvents: Technology development, status, and research needs. *Environmental Security Technology Certification Program (ESTCP)*. , 122
2. Weber, R., *et al.* (2008) Dioxin- and POP-contaminated sites-contemporary and future relevance and challenges. *Environ. Sci. Pollut. Res. Int.* 15, 363-393
3. McMurdie, P.J., *et al.* (2009) Localized plasticity in the streamlined genomes of vinyl chloride respiring *Dehalococcoides*. *PLoS Genet.* 5, e1000714
4. Futamata, H., *et al.* (2007) Reductive dechlorination of chloroethenes by *Dehalococcoides*-containing cultures enriched from a polychlorinated-dioxin-contaminated microcosm. *ISME J.* 1, 471-479
5. Hiraishi, A. (2008) Biodiversity of dehalorespiring bacteria with special emphasis on polychlorinated biphenyl/dioxin dechlorinators. *Microbes and Environments* 23, 1-12
6. Löffler, F.E., and Edwards, E.A. (2006) Harnessing microbial activities for environmental cleanup. *Curr. Opin. Biotechnol.* 17, 274-284
7. Kube, M., *et al.* (2005) Genome sequence of the chlorinated compound-respiring bacterium *Dehalococcoides* species strain CBDB1. *Nat. Biotechnol.* 23, 1269-1273
8. Ouborg, N.J., and Vriezen, W.H. (2007) An ecologist's guide to ecogenomics. *J. Ecol.* 95, 8-16
9. Wagner, M., *et al.* (2007) Unravelling microbial communities with DNA-microarrays: Challenges and future directions. *Microb. Ecol.* 53, 498-506
10. Seshadri, R., *et al.* (2005) Genome sequence of the PCE-dechlorinating bacterium *Dehalococcoides ethenogenes*. *Science* 307, 105-108
11. de Lorenzo, V. (2008) Systems biology approaches to bioremediation. *Curr. Opin. Biotechnol.* 19, 579-589
12. Singh, O.V., and Nagaraj, N.S. (2006) Transcriptomics, proteomics and interactomics: unique approaches to track the insights of bioremediation. *Briefings in functional genomics & proteomics* 4, 355-362
13. Thomas, M.A., and Klaper, R. (2004) Genomics for the ecological toolbox. *TRENDS Ecol. Evol.* 19, 439-445
14. Deweerdt, K.A., *et al.* (1990) *Desulfomonile tiedjei* gen. nov. and sp. nov., a novel anaerobic, dehalogenating, sulfate-reducing bacterium. *Arch. Microbiol.* 154, 23
15. Smidt, H., and de Vos, W.M. (2004) Anaerobic microbial dehalogenation. *Annu. Rev. Microbiol.* 58, 43-73
16. Holmes, V.F., *et al.* (2006) Discrimination of multiple *Dehalococcoides* strains in a trichloroethene enrichment by quantification of their reductive dehalogenase genes. *Appl. Environ. Microbiol.* 72, 5877-5883



17. Cheng, D., and He, J. (2009) Isolation and characterization of *Dehalococcoides* sp. strain MB, which dechlorinates tetrachloroethene to trans-1,2-dichloroethene. *Appl. Environ. Microbiol.* 75, 5910-5918
18. Futagami, T., *et al.* (2008) Biochemical and genetic bases of dehalorespiration. *Chemical Record* 8, 1-12
19. Tas, N., *et al.* (2009) The little bacteria that can; diversity, genomics and ecophysiology of *Dehalococcoides* spp. in contaminated environments. *Microbiol. Biotechnol.* 9999
20. Yoshida, N., *et al.* (2009) A novel *Dehalobacter* species is involved in extensive 4,5,6,7-tetrachlorophthalide dechlorination. *Appl. Environ. Microbiol.* 75, 2400-2405
21. Grostern, A., and Edwards, E.A. (2009) Characterization of a *Dehalobacter* coculture that dechlorinates 1,2-dichloroethane to ethene and identification of the putative reductive dehalogenase gene. *Appl. Environ. Microbiol.* 75, 2684-2693
22. van Doesburg, W., *et al.* (2005) Reductive dechlorination of beta-hexachlorocyclohexane ( $\beta$ -HCH) by a *Dehalobacter* species in coculture with a *Sedimentibacter* sp. *FEMS Microbiol. Ecol.* 54, 87-95
23. Thomas, S.H., *et al.* (2008) The mosaic genome of *Anaeromyxobacter dehalogenans* Strain 2CP-C suggests an aerobic common ancestor to the Delta-Proteobacteria. *PLoS ONE* 3, e2103
24. Nonaka, H., *et al.* (2006) Complete genome sequence of the dehalorespiring bacterium *Desulfitobacterium hafniense* Y51 and comparison with *Dehalococcoides ethenogenes* 195. *J. Bacteriol.* 188, 2262
25. Villemur, R., *et al.* (2006) The *Desulfitobacterium* genus. *FEMS Microbiol. Rev.* 30, 706-733
26. Ludwig, W., *et al.* (2004) ARB: a software environment for sequence data. *Nucleic Acids Res.* 32, 1363-1371
27. He, J., *et al.* (2003) Detoxification of vinyl chloride to ethene coupled to growth of an anaerobic bacterium. *Nature* 424, 62
28. Rahm, B.G., *et al.* (2006) Temporal expression of respiratory genes in an enrichment culture containing *Dehalococcoides ethenogenes*. *Appl. Environ. Microbiol.* 72, 5486
29. Neumann, A., *et al.* (1996) Purification and characterization of tetrachloroethene reductive dehalogenase from *Dehalospirillum multivorans*. *J. Biol. Chem.* 271, 16515
30. Van De Pas, B.A., *et al.* (1999) Purification and molecular characterization of ortho-chlorophenol reductive dehalogenase, a key enzyme of halo-respiration in *Desulfitobacterium dehalogenans*. *J. Biol. Chem.* 274, 20287-20292
31. Adrian, L., *et al.* (2007) Identification of a chlorobenzene reductive dehalogenase in *Dehalococcoides* sp. Strain CBDB1. *Appl. Environ. Microbiol.* 73,

7717-7724

32. Krajmalnik-Brown, R., *et al.* (2004) Genetic identification of a putative vinyl chloride reductase in *Dehalococcoides* sp. strain BAV1. *Appl. Environ. Microbiol.* 70, 6347
33. Müller, J.A., *et al.* (2004) Molecular identification of the catabolic vinyl chloride reductase from *Dehalococcoides* sp. strain VS and its environmental distribution. *Appl. Environ. Microbiol.* 70, 4880 - 4888
34. Thibodeau, J., *et al.* (2004) Purification, cloning, and sequencing of a 3,5-dichlorophenol reductive dehalogenase from *Desulfitobacterium frappieri* PCP-1. *Appl. Environ. Microbiol.* 70, 4532-4537
35. Villemur, R., *et al.* (2002) Occurrence of several genes encoding putative reductive dehalogenases in *Desulfitobacterium hafniense/frappieri* and *Dehalococcoides ethenogenes*. *Can J. Microbiol.* 48, 697 - 706
36. Gabor, K., *et al.* (2008) Divergent roles of CprK paralogues from *Desulfitobacterium hafniense* in activating gene expression. *Microbiol.* 154, 3686-3696
37. Letunic, I., and Bork, P. (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127-128
38. Wagner, A., *et al.* (2009) Transcription analysis of genes encoding homologues of reductive dehalogenases in *Dehalococcoides* sp. strain CBDB1 by using terminal restriction fragment length polymorphism and quantitative PCR. *Appl. Environ. Microbiol.* 75, 1876-1884
39. Pop, S.M., *et al.* (2004) Regulation of anaerobic dehalorespiration by the transcriptional activator CprK. *J. Biol. Chem.* 279, 49910-49918
40. Gabor, K., *et al.* (2006) Characterization of CprK1, a CRP/FNR-type transcriptional regulator of halo-respiration from *Desulfitobacterium hafniense*. *J. Bacteriol.* 188, 2604-2613
41. Levy, C., *et al.* (2008) Molecular basis of halo-respiration control by CprK, a CRP-FNR type transcriptional regulator. *Mol. Microbiol.* 70, 151-167
42. Fung, J.M., *et al.* (2007) Expression of reductive dehalogenase genes in *Dehalococcoides ethenogenes* strain 195 growing on tetrachloroethene, trichloroethene, or 2,3-dichlorophenol. *Appl. Environ. Microbiol.* 73, 4439-4445
43. Waller, A.S., *et al.* (2005) Multiple reductive-dehalogenase-homologous genes are simultaneously transcribed during dechlorination by *Dehalococcoides*-containing cultures. *Appl. Environ. Microbiol.* 71, 8257-8264
44. Rahm, B.G., and Richardson, R.E. (2008) *Dehalococcoides*; gene transcripts as quantitative bioindicators of tetrachloroethene, trichloroethene, and cis-1,2-dichloroethene dehalorespiration rates. *Environ. Sci. Technol.* 42, 5099-5105
45. Rowe, A.R., *et al.* (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
46. Cupples, A.M. (2008) Real-time PCR quantification of *Dehalococcoides* populations: Methods and applications. *J. Microbiol. Meth.* 72, 1-11
47. Smith, C.J., and Osborn, A.M. (2009) Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol. Ecol.* 67, 6-20
48. Smits, T.H., *et al.* (2004) Development of a real-time PCR method for quantification of the three genera *Dehalobacter*, *Dehalococcoides*, and *Desulfitobacterium* in microbial communities. *J. Microbiol. Methods* 57, 369-378
49. Lee, P.K.H., *et al.* (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
50. Nishimura, M., *et al.* (2008) Detection and identification of *Dehalococcoides* species responsible for in situ dechlorination of trichloroethene to ethene enhanced by hydrogen-releasing compounds. *Biotechnol. Appl. Biochem.* 51, 1-7
51. Amos, B.K., *et al.* (2008) Oxygen effect on *Dehalococcoides* viability and biomarker quantification. *Environ. Sci. Technol.* 42, 5718-5726
52. Scheutz, C., *et al.* (2008) Concurrent ethene generation and growth of *Dehalococcoides* containing vinyl chloride reductive dehalogenase genes during an enhanced reductive dechlorination field demonstration. *Environ. Sci. Technol.* 42, 9302-9309
53. Tas, N., *et al.* (2009) Tracking functional guilds: *Dehalococcoides* spp. in European river basins contaminated with hexachlorobenzene. *Appl. Environ. Microbiol.*, AEM.02829-02808
54. Moe, W.M., *et al.* (2009) *Dehalogenimonas lykanthroporepellens* gen. nov., sp nov., a reductively dehalogenating bacterium isolated from chlorinated solvent-contaminated groundwater. *Int. J. Syst. Evol. Microbiol.* 59, 2692-2697
55. Kittelmann, S., and Friedrich, M.W. (2008) Novel uncultured Chloroflexi dechlorinate perchloroethene to trans-dichloroethene in tidal flat sediments. *Environ. Microbiol.* 10, 1557-1570
56. Ahn, Y.B., *et al.* (2008) Biostimulation and bioaugmentation to enhance dechlorination of polychlorinated dibenzo-p-dioxins in contaminated sediments. *FEMS Microbiol. Ecol.* 66, 271-281
57. Sung, Y., *et al.* (2006) Quantitative PCR confirms purity of strain GT, a novel trichloroethene-to-ethene-respiring *Dehalococcoides* isolate. *Appl. Environ. Microbiol.* 72, 1980
58. Yan, J., *et al.* (2009) Isolation of novel bacteria within the Chloroflexi capable of reductive dechlorination of 1,2,3-trichloropropane. *Environ. Microbiol.* 11, 833-843

59. Spurgeon, S.L., *et al.* (2008) High throughput gene expression measurement with real time PCR in a microfluidic dynamic array. *PLoS ONE* 3, e1662
60. Morrison, T., *et al.* (2006) Nanoliter high throughput quantitative PCR. *Nucleic Acids Res.* 34, e123
61. Gelfand, D., and Kastury, K. (2009) High-throughput nanovolume qPCR: Next-generation platform combines microarray throughput with sensitivity. *Genet. Engin. Biotechnol. News* 29
62. He, Z.L., *et al.* (2007) GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J.* 1, 67-77
63. West, K.A., *et al.* (2008) Comparative genomics of *Dehalococcoides ethenogenes* 195 and an enrichment culture containing unsequenced *Dehalococcoides* strains. *Appl. Environ. Microbiol.* 74, 3533-3540
64. Johnson, D.R., *et al.* (2008) Temporal transcriptomic microarray analysis of *Dehalococcoides ethenogenes* strain 195 during the transition into stationary phase. *Appl. Environ. Microbiol.* 74, 2864-2872
65. Tang, Y.J.J., *et al.* (2009) Investigation of carbon metabolism in *Dehalococcoides ethenogenes* strain 195 by use of isotopomer and transcriptomic analyses. *J. Bacteriol.* 191, 5224-5231
66. Maron, P.A., *et al.* (2007) Metaproteomics: A new approach for studying functional microbial ecology. *Microb. Ecol.* 53, 486-493
67. Holliger, C., *et al.* (1998) Reductive dechlorination in the energy metabolism of anaerobic bacteria. *FEMS Microbiol. Rev.* 22, 383
68. Maphosa, F., *et al.* (2010) Exploiting the ecogenomics toolbox for environmental diagnostics of organohalide respiring bacteria. *TRENDS Biotechnol.* online publication
69. Holliger, C., *et al.* (1998) *Dehalobacter restrictus* gen. nov. and sp. nov., a strictly anaerobic bacterium that reductively dechlorinates tetra- and trichloroethene in an anaerobic respiration. *Arch. Microbiol.* 169, 313
70. Wild, A., *et al.* (1996) Isolation of an anaerobic bacterium which reductively dechlorinates tetrachloroethene and trichloroethene. *Biodegradation* 7, 507-511
71. Sun, B., *et al.* (2002) Microbial dehalorespiration with 1,1,1-trichloroethane. *Science* 298, 1023-1025
72. Grostern, A., and Edwards, E.A. (2006) Growth of *Dehalobacter* and *Dehalococcoides* spp. during degradation of chlorinated ethanes. *Appl. Environ. Microbiol.* 72, 428-436
73. Yan, T., *et al.* (2006) The reductive dechlorination of 2,3,4,5-tetrachlorobiphenyl in three different sediment cultures: Evidence for the involvement of phylogenetically similar *Dehalococcoides*-like bacterial populations. *Fems Microb. Ecol.* 55, 248
74. Ariel, G., *et al.* (2009) Chloroform respiration to dichloromethane by a *Dehalobacter* population. *Environ. Microbiol.* 9999

75. van Passel, M.W.J., *et al.* (2005) delta p-Web, an online tool to assess composition similarity of individual nucleic acid sequences. *Bioinformatics* 21, 3053-3055
76. van Passel, M.W.J., *et al.* (2005) An acquisition account of genomic islands based on genome signature comparisons. *BMC Genomics* 6, -
77. Aziz, R., *et al.* (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9, 75
78. Overbeek, R., *et al.* (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* 33, 5691-5702
79. Kanehisa, M., and Goto, S. (2000) KEGG: Kyoto Encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27-30
80. Lowe, T.M., and Eddy, S.R. (1997) tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955-964
81. Rutherford, K., *et al.* (2000) Artemis: sequence visualization and annotation. *Bioinformatics* 16, 944-945
82. Li, L., *et al.* (2003) OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13, 2178-2189
83. Edgar, R.C. (2004) MUSCLE: Multiple sequence alignment with improved accuracy and speed. 2004 *In Computational Systems Bioinformatics Conference, Proceedings*, 728-729
84. Altschul, S.F., *et al.* (1990) Basic local alignment search tool. *J. Mol. Biol.* 215, 403
85. Thompson, J.D., *et al.* (1994) Clustal-W - improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673-4680
86. Hall, B.G., and Barlow, M. (2006) Phylogenetic analysis as a tool in molecular epidemiology of infectious diseases. *Annals Epidemiol.* 16, 157-169
87. Kurtz, S., *et al.* (2004) Versatile and open software for comparing large genomes. *Genome Biol.* 5,
88. Siguier, P., *et al.* (2006) ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 34, D32-D36
89. Kuipers, P.O., *et al.* (1993) Characterization of the nisin gene cluster *nisABTCIPR* of *Lactococcus lactis*. *Eur. J. Biochem.* 216, 281-291
90. Rozen, S., and Skaletsky, H. (1999) Primer3 on the WWW for general users and for biologist programmers. *In Bioinformatics Methods and Protocols*, 365-386
91. Karlin, S. (2001) Detecting anomalous gene clusters and pathogenicity islands in diverse bacterial genomes. *Trends Microbiol.* 9, 335-343
92. Karlin, S., *et al.* (1994) Heterogeneity of genomes - Measures and values. *P.*

*Natl. Acad. Sci. USA* 91, 12837-12841

93. Schumacher, W., and Holliger, C. (1996) The proton/electron ration of the menaquinone-dependent electron transport from dihydrogen to tetrachloroethene in *Dehalobacter restrictus*. *J. Bacteriol.* 178, 2328-2333
94. Schumacher, W., *et al.* (1997) Redox chemistry of cobalamin and iron-sulfur cofactors in the tetrachloroethene reductase of *Dehalobacter restrictus*. *FEBS Lett.* 409, 421-425
95. Maillard, J., *et al.* (2003) Characterization of the corrinoid iron-sulfur protein tetrachloroethene reductive dehalogenase of *Dehalobacter restrictus*. *Appl. Environ. Microbiol.* 69, 4628-4638
96. Maillard, J., *et al.* (2005) Isolation and characterization of Tn-Dha1, a transposon containing the tetrachloroethene reductive dehalogenase of *Desulfitobacterium hafniense* strain TCE1. *Environ. Microbiol.* 7, 107-117
97. Korner, H., *et al.* (2003) Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol. Rev.* 27, 559-592
98. Smidt, H., *et al.* (2000) Transcriptional regulation of the cpr gene cluster in ortho-chlorophenol-respiring *Desulfitobacterium dehalogenans*. *J. Bacteriol.* 182, 5683-5691
99. Breitenstein, A., *et al.* (2001) Analysis of a 2,4,6-trichlorophenol dehalogenating enrichment culture and isolation of the dehalogenating member *Desulfitobacterium frappieri* strain TCP-A. *Arch. Microbiol.* 175, 133
100. Breitenstein, A., *et al.* (2007) Analysis of an anaerobic chemostat population stably dechlorinating 2,4,6-Trichlorophenol. *Engineering in Life Sciences* 7, 380-387
101. Rodionov, D.A., *et al.* (2003) Comparative Genomics of the vitamin B-12 metabolism and regulation in prokaryotes. *J. Biol. Chem.* 278, 41148-41159
102. He, J., *et al.* (2007) Influence of Vitamin B12 and cocultures on the growth of *Dehalococcoides* isolates in defined medium. *Appl. Environ. Microbiol.* 73, 2847-2853
103. Saïd El Fantroussi, H.N.S.N.A. (1998) Anaerobic dechlorinating bacteria. *Biotechnol. Prog.* 14, 167-188
104. Christiansen, N., *et al.* (1998) Purification and characterization of the 3-chloro-4-hydroxy-phenylacetate reductive dehalogenase of *Desulfitobacterium hafniense*. *FEBS Lett.* 436, 159-162
105. Gauthier, A., *et al.* (2006) Occurrence and expression of *crdA* and *cprA5* encoding chloroaromatic reductive dehalogenases in *Desulfitobacterium* strains. *Can. J. Microbiol.* 52, 47-55
106. Mesa, S., *et al.* (2006) A multitude of CRP/FNR-like transcription proteins in *Bradyrhizobium japonicum*. *Biochemical Society transactions* 34, 156-159



107. Madsen, T., and Licht, D. (1992) Isolation and characterization of an anaerobic chlorophenol-transforming bacterium. *Appl. Environ. Microbiol.* 58, 2874-2878
108. Amos, B.K., *et al.* (2007) Detection and quantification of *Geobacter lovleyi* strain SZ: Implications for bioremediation at tetrachloroethene- and uranium-impacted sites. *Appl. Environ. Microbiol.* 73, 6898-6904
109. Markowitz, V.M., *et al.* (2008) IMG/M: a data management and analysis system for metagenomes. *Nucleic Acids Res.* 36, D534-D538
110. McGinnis, S., and Madden, T.L. (2004) BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* 32, W20-W25
111. Christiansen, N., and Ahring, B.K. (1996) *Desulfitobacterium hafniense* sp. nov., an anaerobic, reductively dechlorinating bacterium. *Int. J. Syst. Bacteriol.* 46, 442-448
112. Mazon, H., *et al.* (2007) Transcriptional activation by CprK1 is regulated by protein structural changes induced by effector binding and redox state. *J. Biol. Chem.* 282, 11281-11290
113. Krasotkina, J., *et al.* (2001) Characterization of the B-12- and iron-sulfur-containing reductive dehalogenase from *Desulfitobacterium chlororespirans*. *J. Biol. Chem.* 276, 40991-40997
114. Smidt, H., *et al.* (2000) Halorespiring bacteria-molecular characterization and detection. *Enzyme. Microb. Technol.* 27, 812-820
115. McCarty, P.L. (1997) Breathing with chlorinated solvents. *Science* 276, 1521-1522
116. Regeard, C., *et al.* (2004) Development of degenerate and specific PCR primers for the detection and isolation of known and putative chloroethene reductive dehalogenase genes. *J. Microbiol. Methods* 56, 107 - 118
117. Adrian, L., *et al.* (2000) Bacterial dehalorespiration with chlorinated benzenes. *Nature* 408, 580-583
118. Fennell, D.E., *et al.* (2004) *Dehalococcoides ethenogenes* strain 195 reductively dechlorinates diverse chlorinated aromatic pollutants. *Environ. Sci. Technol.* 38, 2075-2081
119. He, J., *et al.* (2005) Isolation and characterization of *Dehalococcoides* sp. strain FL2, a trichloroethene (TCE)- and 1,2-dichloroethene-respiring anaerobe. *Environ. Microbiol.* 7, 1442 - 1450
120. Hölscher, T., *et al.* (2003) Reductive dehalogenation of chlorobenzene congeners in cell extracts of *Dehalococcoides* sp. strain CBDB1. *Appl. Environ. Microbiol.* 69, 2999 - 3001
121. Jayachandran, G., *et al.* (2003) Dehalorespiration with hexachlorobenzene and pentachlorobenzene by *Dehalococcoides* sp. strain CBDB1. *Arch. Microbiol.* 180, 411-416

122. Maymó-Gatell, X., *et al.* (1997) Isolation of a bacterium that reductively dechlorinates tetrachloroethene to ethene. *Science* 276, 1568
123. Lee, P.K.H., *et al.* (2006) Reductive dehalogenase gene expression as a biomarker for physiological activity of *Dehalococcoides* spp. *Appl. Environ. Microbiol.* 72, 6161-6168
124. Ritalahti, K.M., *et al.* (2006) Quantitative PCR targeting 16S rRNA and reductive dehalogenase genes simultaneously monitors multiple *Dehalococcoides* strains. *Appl. Environ. Microbiol.* 72, 2765-2774
125. Magnuson, J.K., *et al.* (2000) Trichloroethene reductive dehalogenase from *Dehalococcoides ethenogenes*: sequence of *tceA* and substrate range characterization. *Appl. Environ. Microbiol.* 66, 5141 - 5147
126. Hölscher, T., *et al.* (2004) Multiple nonidentical reductive-dehalogenase-homologous genes are common in *Dehalococcoides*. *Appl. Environ. Microbiol.* 70, 5290 - 5297
127. Sung, Y., *et al.* (2006) *Geobacter lovleyi* sp. nov. strain SZ, a novel metal-reducing and tetrachloroethene-dechlorinating bacterium. *Appl. Environ. Microb.* 72, 2775
128. Becker, J.G. (2006) A modeling study and implications of competition between *Dehalococcoides ethenogenes* and other tetrachloroethene-respiring bacteria. *Environ. Sci. Technol.* 40, 4473-4480
129. Duhamel, M., and Edwards, E.A. (2007) Growth and yields of dechlorinators, acetogens, and methanogens during reductive dechlorination of chlorinated ethenes and dihaloelimination of 1,2-dichloroethane. *Environ. Sci. Technol.* 41, 2303-2310
130. Zaan, B.v.d., *et al.* (2007) Molecular monitoring of a biological in situ treatment of chloroethene polluted site. In *Soil and Water Symposium*
131. Nubel, U., *et al.* (1996) Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. *J. Bacteriol.* 178, 5636-5643
132. Cupples, A.M., *et al.* (2003) Growth of a *Dehalococcoides*-like microorganism on vinyl chloride and cis-dichloroethene as electron acceptors as determined by competitive PCR. *Appl. Environ. Microbiol.* 69, 953
133. He, J., *et al.* (2003) Complete detoxification of vinyl chloride by an anaerobic enrichment culture and identification of the reductively dechlorinating population as a *Dehalococcoides* species. *Appl. Environ. Microbiol.* 69, 996
134. Huber, T., *et al.* (2004) Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* 20, 2317-2319
135. Adrian, L., *et al.* (2007) Growth of *Dehalococcoides* strains with chlorophenols as electron acceptors. *Environ. Sci. Technol.* 41, 2318
136. Cupples, A.M., *et al.* (2004) Vinyl chloride and cis-dichloroethene dechlorination kinetics and microorganism growth under substrate-limiting



- conditions. *Environ. Sci. Technol.* 38, 1102
137. Yu, S., *et al.* (2005) Kinetics and inhibition of reductive dechlorination of chlorinated ethylenes by two different mixed cultures. *Environ. Sci. Technol.* 39, 195
138. Shen, H., and Sewell, G.W. (2005) Reductive biotransformation of tetrachloroethene to ethene during anaerobic degradation of toluene: Experimental evidence and kinetics. *Environ. Sci. Technol.* 39, 9286-9294
139. Smatlak, C.R., *et al.* (1996) Comparative kinetics of hydrogen utilization for reductive dechlorination of tetrachloroethene and methanogenesis in an anaerobic enrichment culture. *Environ. Sci. Technol.* 30, 2850
140. Ballapragada, B.S., *et al.* (1997) Effect of hydrogen on reductive dechlorination of chlorinated ethenes. *Environ. Sci. Technol.* 31, 1728
141. Haston, Z.C., and McCarty, P.L. (1999) Chlorinated ethene half-velocity coefficients (Ks) for reductive dehalogenation. *Environ. Sci. Technol.* 33, 223
142. Garant, H., and Lynd, L. (1998) Applicability of competitive and noncompetitive kinetics to the reductive dechlorination of chlorinated ethenes. *Biotechnol. Bioen.* 57, 751-755
143. Fennell, D.E., and Gossett, J.M. (1998) Modeling the production of and competition for hydrogen in a dechlorinating culture. *Environ. Sci. Technol.* 32, 2450
144. Lendvay, J.M., *et al.* (2003) Bioreactive barriers: a comparison of bioaugmentation and biostimulation for chlorinated solvent remediation. *Environ. Sci. Technol.* 37, 1422
145. Behrens, S., *et al.* (2008) Monitoring abundance and expression of *Dehalococcoides* species chloroethene-reductive dehalogenases in a tetrachloroethene-dechlorinating flow column. *Appl. Environ. Microbiol.* 74, 5695-5703
146. Azizian, M.F., *et al.* (2008) Continuous-flow column study of reductive dehalogenation of PCE upon bioaugmentation with the Evanite enrichment culture. *J. Contaminant Hydrology* 100, 11-21
147. Chung, J., *et al.* (2008) Bioreduction of trichloroethene using a hydrogen-based membrane biofilm reactor. *Environ. Sci. Technol.* 42, 477-483
148. Kittelmann, S., and Friedrich, M.W. (2008) Identification of novel perchloroethene-respiring microorganisms in anoxic river sediment by RNA-based stable isotope probing. *Environ. Microbiol.* 10, 31-46
149. Regeard, C., *et al.* (2005) Indications for acquisition of reductive dehalogenase genes through horizontal gene transfer by *Dehalococcoides ethenogenes* strain 195. *Appl. Environ. Microbiol.* 71, 2955 - 2961
150. McMurdie, P.J., *et al.* (2007) Unusual codon bias in vinyl chloride reductase genes of *Dehalococcoides* species. *Appl. Environ. Microbiol.* 73, 2744-2747
151. Duhamel, M., *et al.* (2004) Characterization of a highly enriched *Dehalococcoides*-containing culture that grows on vinyl chloride and trichloroethene.

*Appl. Environ. Microbiol.* 70, 5538-5545

152. Aulenta, F., *et al.* (2006) Anaerobic bioremediation of groundwater containing a mixture of 1,1,2,2-tetrachloroethane and chloroethenes. *Biodegradation* 17, 193-206
153. Aulenta, F., *et al.* (2004) Detection and quantitative estimation of *Dehalococcoides* spp. in a dechlorinating bioreactor by a combination of fluorescent in situ hybridization (FISH) and kinetic analysis. *Appl. Microbiol. Biotechnol.* 64, 206
154. Neumann, A., *et al.* (1998) Tetrachloroethene dehalogenase from *Dehalospirillum multivorans*: Cloning, sequencing of the encoding genes, and expression of the *pceA* gene in *Escherichia coli*. *J. Bacteriol.* 180, 4140
155. Klein, D. (2002) Quantification using real-time PCR technology: applications and limitations. *Trends Mol. Medicine* 8, 257 - 260
156. Mackay, I.M. (2004) Real-time PCR in the microbiology laboratory. *Clinical Microbiology and Infection* 10, 190 - 212
157. Schena, L., *et al.* (2004) Real-time quantitative PCR: a new technology to detect and study phytopathogenic and antagonistic fungi. *European J. Plant Pathol.* 110, 893 - 908
158. Ding, C., and Cantor, C.R. (2004) Quantitative analysis of nucleic acids--the last few years of progress. *J. Biochem. Mol. Biol.* 37, 1 - 10
159. Lievens, B., *et al.* (2006) Real-time PCR for detection and quantification of fungal and oomycete tomato pathogens in plant and soil samples. *Plant Science* 171, 155 - 165
160. Schena, L., *et al.* (2006) Detection and quantification of *Phytophthora ramorum*, *P-kernoviae*, *P-citricola* and *P-quercina* in symptomatic leaves by multiplex real-time PCR. *Mol. Plant Pathol.* 7, 365 - 379
161. Tooley, P.W., *et al.* (2006) Real-time fluorescent polymerase chain reaction detection of *Phytophthora ramorum* and *Phytophthora pseudosyringae* using mitochondrial gene regions. *Phytopathol.* 96, 336 - 345
162. Brisson, M., *et al.* (2004) Optimization of multiplex reactions. *A-Z of quantitative PCR, edited by Bustin SA* 1st edition, 636 - 640
163. Klerks, M.M., *et al.* (2004) Comparison of real-time PCR methods for detection of *Salmonella enterica* and *Escherichia coli* O157: H7, and introduction of a general internal amplification control. *J. Microbiol. Meth.* 59, 337 - 349
164. Zijlstra, C., and Van Hoof, R.A. (2006) A multiplex real-time polymerase chain reaction (TaqMan) assay for the simultaneous detection of *Meloidogyne chitwoodi* and *M-fallax*. *Phytopathol.* 96, 1255 - 1262
165. van Doorn, R., *et al.* (2007) Quantitative multiplex detection of plant pathogens using a novel ligation probe-based system coupled with universal, high-throughput real-time PCR on OpenArrays<sup>TM</sup>. *BMC Genomics* 8, 276

166. Szemes, M., *et al.* (2005) Diagnostic application of padlock probes-multiplex detection of plant pathogens using universal microarrays. *Nucleic Acids Res.* 33, e70
167. van Doorn, R., *et al.* (2009) Robust detection and identification of multiple Oomycetes and Fungi in environmental samples by using a novel cleavable padlock probe-based ligation detection assay. *Appl. Environ. Microbiol.* 75, 4185-4193
168. Taş, N. (2009) *Dehalococcoides* spp. in river sediments: insights in functional diversity and dechlorination activity. *PhD thesis*
169. Nikolausz, M., *et al.* (2008) Single-nucleotide primer extension assay for detection and sequence typing of *Dehalococcoides* spp. *Appl. Environ. Microbiol.* 74, 300-304
170. Sleep, B.E., *et al.* (2006) Biological enhancement of tetrachloroethene dissolution and associated microbial community changes. *Environ. Sci. Technol.* 40, 3623-3633
171. Rahm, B.G., *et al.* (2006) Molecular characterization of microbial populations at two sites with differing reductive dechlorination abilities. *Biodegradation* 17, 523
172. Qvarnstrom, Y., *et al.* (2005) Comparison of real-time PCR protocols for differential laboratory diagnosis of amebiasis. *J. Clin. Microbiol.* 43, 5491-5497
173. Tian, B., *et al.* (2004) Comparison of Real-Time PCR assay methods in detection and quantification of b-actin genes in mouse tissues. *Preclinica* 2
174. Smalla, K., *et al.* (1993) Rapid DNA extraction protocol from soil for polymerase chain reaction-mediated amplification. *J. Appl. Bacteriol.* 74, 78-85
175. Wohlfarth, G., and Diekert, G. (1997) Anaerobic dehalogenases. *Curr. Opin. Biotechnol.* 8, 290-295
176. van Eekert, M.H., and Schraa, G. (2001) The potential of anaerobic bacteria to degrade chlorinated compounds. *Water Sci Technol* 44, 49-56
177. Abelson, P.H. (1990) Inefficient remediation of groundwater pollution. *Science* 250, 733
178. Okuda, I., *et al.* (1996) Physicochemical transport processes affecting the removal of residual DNAPL by nonionic surfactant solutions. *Environ. Sci. Technol.* 30, 1852-1860
179. James, C.A., and Strand, S.E. (2009) Phytoremediation of small organic contaminants using transgenic plants. *Curr. Opin. Biotechnol.* 20, 237-241
180. Salt, D.E., *et al.* (1995) Phytoremediation: A novel strategy for the removal of toxic metals from the environment using plants. *Nat. Biotech.* 13, 468-474
181. Hong-Bo, S., *et al.* (2010) Understanding molecular mechanisms for improving phytoremediation of heavy metal-contaminated soils. *Critic. Rev. Biotechnol.* 30, 23-30
182. Smets, B.F., and Pritchard, P.H. (2003) Elucidating the microbial component of natural attenuation. *Curr. Opin. Biotechnol.* 14, 283-288
183. Duhamel, M.A., *et al.* (2002) Comparison of anaerobic dechlorinating

- enrichment cultures maintained on tetrachloroethene, trichloroethene, cis-dichloroethene and vinyl chloride. *Water Res.* 36, 4193
184. Major, D., *et al.* (2002) Field demonstration of a successful bioaugmentation to achieve dechlorination of tetrachloroethene to ethene. *Environ. Sci. Technol.* 36, 5106
185. Singh, O.V. (2006) Proteomics and metabolomics: The molecular make-up of toxic aromatic pollutant bioremediation. *Proteomics* 6, 5481-5492
186. Qin, J., *et al.* (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59-65
187. Vogel, T.M., *et al.* (2009) TerraGenome: a consortium for the sequencing of a soil metagenome. *Nat. Rev. Micro.* 7, 252-252
188. Morris, R.M., *et al.* (2007) Comparative proteomics of *Dehalococcoides* spp. reveals strain-specific peptides associated with activity. *Appl. Environ. Microbiol.* 73, 320-326
189. Tyson, G.W., *et al.* (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 428, 37-43
190. Woyke, T., *et al.* (2009) Assembling the marine metagenome, One cell at a time. *PLoS ONE* 4, e5299
191. Denef, V.J., *et al.* (2010) Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. *P. Natl. Acad. Sci. USA* 107, 2383-2390
192. DeSantis, T., *et al.* (2007) High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. *Microb. Ecol.* 53, 371-383
193. Brodie, E.L., *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
194. Yergeau, E., *et al.* (2009) Environmental microarray analyses of Antarctic soil microbial communities. *ISME J.* 3, 340-351
195. Kuramae, E.E., *et al.* (2010) Microbial secondary succession in a chronosequence of chalk grasslands. *ISME J.* online publication
196. Hu, C.X., *et al.* (2009) Analytical strategies in lipidomics and applications in disease biomarker discovery. *J. Chromatogr. B* 877, 2836-2846
197. Morris, R.M., *et al.* (2006) Transcription and mass-spectroscopic proteomic studies of electron transport oxidoreductases in *Dehalococcoides ethenogenes*. *Environ. Microbiol.* 8, 1499-1509
198. Werner, J.J., *et al.* (2009) Absolute quantification of *Dehalococcoides* proteins: enzyme bioindicators of chlorinated ethene dehalorespiration. *Environ. Microbiol.* 11, 2687-2697
199. Koenigsberg, S.S., *et al.* (2005) Environmental biotechnology: A

bioremediation perspective. *Remediation J.* 15, 5-25

200. MacLean, D., *et al.* (2009) Application of 'next-generation' sequencing technologies to microbial genetics. *Nat. Rev. Micro.* advanced online publication



# Summary

## Summary

Organohalide respiring bacteria (OHRB) are efficient degraders of organohalides such as chlorinated ethenes, chlorophenols, and other halogenated aliphatic and aromatic hydrocarbons. Nevertheless, these organohalides appear to persist at various locations. The reason for this lack of degradation can be attributed to the absence of OHRB in sufficient numbers or improper physico-chemical conditions for their growth and activity. Hence, there is an urgent need for fast, robust and sensitive methods that allow for predicting and monitoring the bioremediation potential and activity of OHRB. The ecogenomic toolbox that capitalizes on genomics and other high throughput approaches offers completely new avenues to study OHRB in natural environments. To this end, we studied the genomic repertoire, as well as functional gene expression patterns, in two important Gram-positive OHRB. Moreover, we applied this and other accumulating knowledge in dechlorinating bioreactors and developed advanced detection systems for OHRB that were tested in various practical field applications.

Firstly, we provided advanced insights into the genome sequence of *b*-hexachlorocyclohexene-dehalogenating *Dehalobacter* sp. E1 that grows in strict coculture with *Sedimentibacter* sp. B4. The coculture metagenome and the genome of strain B4 were sequenced using 454 GS-FLX technology. The genome sequence of strain E1 (2.6 Mbp) contains 10 putative reductive dehalogenase (Rdh)-encoding gene clusters, including the *pceABCT* operon previously described in *Dehalobacter restrictus*. This suggested that strain E1 has a greater organohalide respiration potential than previously observed and we confirmed this experimentally by showing the dechlorination of trichloroethene by the coculture.

A second representative of OHRB that was studied in detail is *Desulfitobacterium hafniense* DCB-2, that is able to dehalogenate a diverse range of chloroaromatic compounds. Its genome sequence contains seven chlorophenol Rdh (*cprA*) -like genes, five of which (*cprA1-cprA5*) are predicted to encode functional proteins. We determined the transcriptional patterns for these five *cprA1-cprA5* genes using as substrates 2,3-dichlorophenol (2,3-DCP) for the meta-dechlorination and 2,4-dichlorophenol (2,4-DCP) for the ortho-dechlorination positions. In the presence of 2,4-DCP, *cprA1* and *cprA5*



were strongly induced, whereas 2,3-DCP induced *cprA3* and *cprA5*. This can be explained by distinct control at the transcriptional level of the various *cprA* genes, resulting in different specificities and responses to meta- and ortho- chlorophenols.

Molecular monitoring and model simulations were applied to determine in-situ degradation performance of an on-site dechlorinating bioreactor and its influence on a contamination plume. *Dehalococcoides* was the dominant dechlorinating microorganism as revealed by quantitative PCR (qPCR) targeting 16S rRNA- and chloroethene Rdh genes (*tceA*, *vcrA*, *bvcA*). The presence of all three Rdh genes indicated co-existence of several distinct *Dehalococcoides* populations in the bioreactor and groundwater, with reduction of cis-dichloroethene and vinyl chloride as the main dechlorinating activities. Saturation analysis showed that while performance of the bioreactor improved due to especially the addition of molasses, OHRB were fast-growing. However, once near-complete dehalogenation was achieved, *Dehalococcoides* only grew slowly and did not further increase in population density.

Finally, we report the development and application of a high-throughput, ligation based assay for simultaneous quantitative detection of multiple OHRB populations and their key Rdh genes. We designed nine taxonomic and ten functional gene ligation-based circularization probes and primer sets for probe amplification in Taqman qPCR. This assay combines the multiplex capabilities and specificity of ligation probes with high-throughput qPCR in the Biotrove OpenArray™, resulting in a flexible, quantitative multiplex detection system for OHRB. The applicability of this system was tested in various contaminated soils. Moreover, an outlook is provided on how the ecogenomics approaches described here and new tools such as proteomics and metabolomics currently being developed can be applied to advance both the understanding and the exploitation of soil bioremediation by OHRB.



# Nederlandse samenvatting

## **Het opsporen van organohalogeën-reducerende bacteriën: Ecogenomics benaderingen om de bioremediatie-capaciteit van de bodem te beoordelen.**

Organohalogeën-reducerende bacteriën (OHRB) zijn efficiënte afbrekers van organische chloorverbindingen, zoals gechloreerde ethenen, chloorfenolen en andere gehalogeneerde alifatische en aromatische koolwaterstoffen. Desondanks, lijken deze organische chloorverbindingen te volharden op verschillende locaties. De reden voor dit gebrek aan afbraak kan worden toegeschreven aan het ontbreken van OHRB in voldoende aantallen of aan verkeerde fysisch-chemische omstandigheden voor hun groei en activiteit. Derhalve is er een dringende behoefte aan snelle, robuuste en gevoelige methoden die het voorspellen van en het toezicht houden op het bioremediatie potentieel en de activiteit van OHRB mogelijk maken. De ecogenomics gereedschapskist die gebaseerd is op genomics en andere 'high throughput' benaderingen biedt geheel nieuwe wegen om OHRB te bestuderen in natuurlijke omgevingen.

Daartoe bestudeerden we het genomische repertoire en de functionele genexpressie patronen van twee belangrijke Gram-positieve OHRB. Tevens, hebben we deze en andere vergaarde kennis toegepast in dechlorerende bioreactoren en hebben we geavanceerde detectie systemen voor OHRB ontwikkeld en getest in verschillende praktische toepassingen.

Ten eerste, hebben we geavanceerde inzichten in de genoomsequentie van de  $\beta$ -hexachlorocyclohexaan-reducerende *Dehalobacter* sp. E1 verschaft, welke in strikte cocultuur met *Sedimentibacter* sp. B4 groeit. Het cocultuur metagenoom en het genoom van de B4 stam zijn gesequenced met behulp van 454 GS FLX-technologie. De genoomsequentie van stam E1 (2,6 Mbp) bevat 10 vermeende reductieve dehalogenase (*rdh*)-coderende genenclusters, met inbegrip van de eerder beschreven *pceABCT* operon van *Dehalobacter*

*restrictus*. Dit suggereerde dat de E1 stam een groter organohalogeën reductie potentieel heeft dan voorheen was waargenomen. We hebben dit experimenteel bevestigd door de dechlorering van trichlooretheen door de cocultuur aan te tonen.

Een tweede vertegenwoordiger van OHRB die in detail werd bestudeerd is *Desulfotobacterium hafniense* DCB-2, welke in staat is om een breed scala aan gechloroerde aromatische verbindingen te reduceren. De genomsequentie bevat zeven chloorfenol RDH (*cprA*) -achtige genen, waarvan er vijf (*cprA1-cprA5*) zijn voorspeld te coderen voor functionele eiwitten. We hebben de transcriptie patronen bepaald voor deze vijf genen met als substraten 2,3-dichloorfenol (2,3-DCP) voor de meta-dechlorering en 2,4-DCP) voor de ortho-dechlorering posities. In aanwezigheid van 2,4-DCP waren *cprA1* en *cprA5* sterk geïnduceerd, terwijl 2,3-DCP *cprA3* en *cprA5* induceerde. Dit kan verklaard worden door afzonderlijke controle op het transcriptie-niveau van de verschillende *cprA* genen, wat resulteert in verschillende specificiteiten en reacties op *meta*- en *ortho*-chloorfenolen.

Moleculaire monitoring en modelsimulaties werden toegepast om de in-situ afbraak prestaties van een on-site dechlorerende bioreactor te bepalen en zijn invloed op de vervuilingsspluim. *Dehalococcoides* was het dominante dechlorerende micro-organisme zoals bleek uit kwantitatieve PCR (qPCR) gericht op de 16S rRNA en chlooretheen Rdh-genen (*tceA*, *vcra*, *bvcA*). De aanwezigheid van alle drie de *rdh* genen duidde op het naast elkaar bestaan van verscheidene afzonderlijke *Dehalococcoides* populaties in de bioreactor en het grondwater, met reductie van cis-dichlooretheen en vinylchloride als de belangrijkste dechlorerende activiteiten. Verzadigingsanalyse toonde aan dat dat de OHRB snel-groeiend waren ten tijden dat de prestaties van de bioreactor verbeterden als gevolg van met name de toevoeging van melasse. Echter, zodra bijna volledige dehalogenatie werd bereikt, groeide *Dehalococcoides* alleen maar langzaam met als gevolg een niet verder

toenemende populatiedichtheid.

Ten slotte, informeren we over de ontwikkeling en toepassing van een ‘high-throughput’, ligatie gebaseerde test voor gelijktijdige kwantitatieve detectie van meerdere OHRB populaties en hun belangrijkste *rdh* genen. We ontworpen negen taxonomische en tien functionele genligatie-gebaseerde circularisatie probes en primer sets voor probe amplificatie in Taqman qPCR. Deze test combineert de multiplex mogelijkheden en specificiteit van de Ligation Probe met een ‘high-throughput’ qPCR in de Biotrove OpenArray™, wat resulteert in een flexibele, kwantitatieve multiplex detectie systeem voor OHRB. De toepasbaarheid van dit systeem werd getest in verschillende verontreinigde bodems. Tevens, geven we onze visie op welke wijze de hier beschreven ecogenomics benaderingen en nieuwe tools, zoals proteomics en metabolomics die momenteel ontwikkeld worden, kunnen worden toegepast om zowel het inzicht in en de exploitatie van de bodem bioremediatie door OHRB te vergroten.



## About the Author

Farai Maphosa was born in Zimbabwe on the 7<sup>th</sup> of March 1978. He did his Bachelors from 1997 to 2001 in Applied Biology and Biochemistry from the National University of Science and Technology, Zimbabwe. During this time he spent his college vacations working for a Dutch company involved in producing dried vegetables for the European market and got exposed to the Dutch soup delights among other things. He then did a one year traineeship with Nestle Zimbabwe where he specialized in production management and also completed a Diploma in Quality Control and Assurance. Following this he worked for a chemical company as a Quality Systems Manager working on implementing ISO 9001 quality standards and ISO14001 environmental and health and safety standards. It was during this time that he developed a keen interest towards environmental problems caused by chemical pollution. He then moved to the Netherlands in 2003 to study for an MSc in Cellular and Molecular Biotechnology at Wageningen University. He did his Msc thesis entitled “Molecular characterisation of dehalogenating microbial communities in groundwater samples” in the Molecular Ecology group of the Laboratory of Microbiology following on his interest in environmental issues. Immediately upon completion he continued to work as PhD student in the Laboratory of Microbiology on an Ecogenomics project –“The Bioremediation capacity of the soil”. The result of this research project is presented in this thesis. Since April 2009 he has been appointed as a post-doctoral researcher in the same Laboratory of Microbiology, Wageningen University, working on several collaborative projects involved in bioremediation of polluted soil and groundwater systems.



## List of Publications

**Farai Maphosa**, Willem M. de Vos, and Hauke Smidt (2010) Exploiting the ecogenomics toolbox for environmental diagnostics of organohalide respiring bacteria, *Trends in Biotechnology*, *In Press*

**Farai Maphosa**, Hauke Smidt, Willem M de Vos, Wilfred F.M. Röling, (2010) Microbial community- and metabolite dynamics of an anoxic dechlorinating bioreactor, *under review for Environmental Science and Technology*

**Farai Maphosa**, Mark W.J van Passel, Willem M. de Vos and Hauke Smidt (2010) Advanced insight in a hexachlorocyclohexane-dechlorinating coculture of *Dehalobacter* sp. and *Sedimentibacter* sp. based on metagenome analysis, *In Preparation for publication*

**Farai Maphosa**, Krisztina Gabor, Mark H.J. Sturme, Tian Ye, Willem M. de Vos and Hauke Smidt (2010) Differential expression of *Desulfitobacterium hafniense* DCB-2 chlorophenol reductive dehalogenase genes in response to *meta*- or *ortho*-chlorophenols, *In Preparation for publication*

**Farai Maphosa**, Ronald van Doorn, Willem M. de Vos, Cor Schoen, Hauke Smidt (2010) Chasing organohalide respirers: Quantitative multiplex detection of dechlorinating bacteria using a ligation probe-based detection system coupled to OpenArrays™, *In Preparation for publication*

## Acknowledgements

My first MSc lecture at Wageningen University was in the “beautifully old” former Laboratory of Microbiology building “up the hill”. As the saying goes first impressions last long, both the building and the lecturers of that late September 2003 afternoon made a positive impression on me, no wonder I came back to do my thesis in Molecular ecology (which was also the course I was following). For this impression I have Gosse Schraa (you had my vote for your teacher of the year award) and of course Hauke Smidt to thank. Having arrived late due to visa delays, I must also acknowledge Dr. Sonja Isken my MSc coordinator who advised that I take the Molecular Ecology course instead of a biochemistry course I wanted to do (it so turned out this was a very unpopular course with a high failure rate). Secondly when I had to make a choice for an MSc Thesis with an ambition to further do a PHD – Sonja your advice worked for me – “*...join a smaller group like Molecular Ecology and tell them you want a distinction*” So I did my MSC thesis with Hauke and Meta van Heusden and immediately continued to do my PHD. Hauke, thank you very much for the opportunities that you have given me right from 2003 when I came Wageningen. It has been great having you as a supervisor and mentor – thanks for helping me see the “half-full glass rather than the half-empty glass” when things got tough during the first 2 years. Your availability 24/7 is amazing. Willem thanks for all the timely advice and coaching – like you said “*you won’t see much of me but you are in good hands*” – Hauke has been great to help me through this journey.

I would like to thank Meta for teaching me all the techniques and tricks for working with organohalide respirers. You have also been a great friend since then. Special thanks to Wilfred Röling, for the collaborations especially the time I spent in your laboratory at Vrij Universiteit Amsterdam, and also for the modeling ideas that eventually resulted in Chapter 4. I also want to thank Ronald van Doorn, Cor Scheon, Richard van Hoof and Peter Bonants from Plant Research International for the collaboration on the work that resulted in Chapter 5. Ronald I enjoyed working with you on all those endless ligations and PCRs.

Much appreciation goes to you Floris van der Wal - for the bioreactor samples and data - I had many interested presentations entitled “*Cleaning under the*

*Supermarket.*” Marlea and Shakti from Bioclear many thanks for all the site samples and collaboration - my MSc students always had lots of interesting characterisations to do.

Thanks to all the MolEco’s past and present – it has been great coming in contact with each of you and you have been special in your own way. Past MolEcos Sergey, Vesela, Mirjana, David, Toshio, Rocio, Eline, Maaike, Meta, Susana, Mark Sturme, Mariana, Alcina, Carien, Stephanie, Mahmut, Markus and Neslihan you were there in the beginning and helped me fit along. Neslihan and Mark – I am grateful for all the discussions we had and particularly for helping make my one and only significant poster.

To the present Moleco’s its good to be still part of you – Erwin, Hans, Wilma, Detmer, Sahar, Josse, Sebastian, Tom, Milka, Su, Muriel, Ineke, Mauricio, Sukithar, Rafeal, Maria, Vincent, Jing, Saskia, Hans Snel –the only Moleco who has seen where I come from. Special thanks to Mark van Passel for all the genome data mining and pearl scripts. Hans Heilig – thanks for all the help and support and particularly the DNA extractions for genome sequencing. Tian thanks for all the hard work you did for the cprA expression. Petia and Odette its been great sharing this PHD journey together and now we have come to the completion – its been great to share the ups and downs with you. My dear office mates Thomas, Rozelin and Coline - thank you very much for listening and often just being there. I have had many laughs with you and you have often been of help in many ways. Thanks Thomas for the many discussions about everything, arranging the football and games, and thanks also for being my paranymph. I also am grateful to my MSC students - Paula Carvalho, Cai Hengzhe, Sanne Lazaroms and Thamari Chitavati for all your hardwork and good reports.

Also a special thanks all MicFys, Bacgen and Fungen guys. Wim van Doesburg, Farrakh Mehboob, and Miriam van Eekert I appreciate you for helping with the coculture. Kristina Gabor thanks for the collaboration that made Chapter 3. Mark Levisson and Marcel thanks for all the translations perfecting my “google translates.” Mark it was also very nice to have a chat in the Marijkeweg driveway after work every now and then. Wim Roelosen -from pc’s to sediment cores - you just amazing - thanks for everything

A big thanks to the Amazing Grace Parish (AGP) family, its been wonderful being part of you and to be involved in one way or the other in your lives as a friend and pastor. I cannot mention you all by name but you have all made my stay here special and fulfilling. For “...*of making many books there is no end; and much study is a weariness of the flesh... the conclusion of the whole matter is this: Fear God, and keep His commandments: for this is the whole duty of man.*” Among all the busy lab schedules it was a great joy to come fellowship and worship with all of you. Special thanks to Virginia – whom I have often given many extra things to do from my administrative responsibilities in AGP and for being my paranymph. Elton, Edna and Heather you been great friends to me and my family.

I also appreciate the Zimbabwean community - it was always heart warming to eat sadza neoxtail reminiscing about our motherland.

Bert and Anneke Geerken and Jos and Sarah, Irene and Dito, Marsha and Oma Geerken, my Dutch family - bedankt voor alles - for helping me settle down and understand my way around and for the many family functions you have included me and my family.

Baba naMai Maphosa and brothers Temba, Lindani and Tatenda, thank you for inspiring me to go all the way. Mum, now you can indeed give that long awaited thanksgiving speech – having not celebrated the Bachelor’s and Master degree, this now is the occasion for it. All those beautiful cakes and dresses you made and designed have not been for nothing. Above all thank you for just being Mum and Dad. Nyiyabonga.

My lovely wife Busi –my Proverbs 31 fulfillment –you have been strong and wonderful, thanks for being there throughout and holding fort alone on in the final months of writing, even with little baby Hillel and the “jongen”Anotida. You have given me support in every way. I love you lots and Thank you. I dedicate this work to you and the boys.

Lastly I give all the praise and glory to God who “*fathoms all things, ...counts the grains of the sand and calls the stars by name...*” and gives us all we know and understand.



Netherlands Research School for the  
Socio-Economic and Natural Sciences of the Environment

## CERTIFICATE

The Netherlands Research School for the  
Socio-Economic and Natural Sciences of the Environment  
(SENSE), declares that

*Farai Maphosa*

Born on: *7 March 1978* at: *Bulilima-Mangwe, Zimbabwe*

has successfully fulfilled all requirements of the  
Educational Programme of SENSE.

Place: *Wageningen* Date: *31 May 2010*

the Chairman of the  
SENSE board

Prof. dr. R. Leemans

the SENSE Director  
of Education

Dr. A. van Dommelen



The SENSE Research School declares that Mr. Farai Maphosa has successfully fulfilled all requirements of the Educational PhD Programme of SENSE with a work load of 39 ECTS, including the following activities:

SENSE PhD courses:

- Environmental Research in Context
- Research Context Activity: “Organizing workshop on ‘Ecogenomics’ as part of the SENSE Summer Symposium 2005 (23 June 2005 – Ede, the Netherlands)”
- Principles of Ecological Genomics
- Environmental Risk Assessment of Micropollutants

Other PhD and MSc courses:

- Dutch Beginners
- Techniques for Writing and Presenting a Scientific Paper
- Project and Time Management
- Basic and Advanced Statistics
- Systems Biology: Principles of -omics data analysis
- ARB – A software environment for Sequencing Data
- Career Assessment
- Career Perspectives
- Writing grant proposals

Research and Management Skills:

- Laboratory of Microbiology PhD Study Trip 2006, California Universities and Research Institutes, USA
- Visits to: Prof B. Schinks Laboratory, 2006, Konstanz, Germany
- Visit to: Max Planck Institute, 2006, Marburg, Germany



Oral Presentations:

- ° Chasing Halorespirers -Functional ecogenomics approaches towards dedicated degraders  
Ecogenomics Annual Meeting, 2006, Amsterdam, The Netherlands
- ° Ecogenomics Approaches to Assess the Bioremediation Capacity of Soils  
Max Planck Institute, 2006, Marburg, Germany
- ° Ecogenomics approaches to assess the bioremediation capacity of soils.  
PHD-Study Trip, 2006, DOE-Joint Genome Institute & Stanford University, California, USA
- ° In pursuit of Halorespirers –Ecogenomics approaches to assess the bioremediation capacity of soils contaminated with chlorinated ethenes– Soil and Water Conference, 2006, Zeist, The Netherlands
- ° The *Dehalobacter* - *Sedimentibacter* partnership in reductive dechlorination of polluted ground water.Ecogenomics Annual Meeting, 2007, Amsterdam, The Netherlands
- ° Cleaning Under the supermarket – Assessing the bioremediation capacity of soils contaminated with chlorinated ethenes. – Sensible Water, 2007, Leeuwarden, The Netherlands
- ° Chasing Halorespirers: Multiplex detection of dechlorinating bacteria using PRI-lock probes 3IMBE, 2008, Mallorca, Spain
- ° First genome sequence of a hexachlorocyclohexane-dechlorinating coculture of *Dehalobacter* sp. and *Sedimentibacter* sp. FEMS Congress, 2009, Gottenburg, Sweden

Mr. J. Feenstra  
SENSE Coordinator PhD Education and Research

Printing of this thesis was financially supported by:

-The Akkermansia Foundation

- Bioclear B.V., Groningen, The Netherlands



- PJ Milieu B.V., Nijkerk, The Netherlands



Design and Layout: Farai Maphosa

On the Cover : Dapi stain image of the *Dehalobacter* and *Sedimentibacter* coculture.

Printed by : GVO Printers & designers B.V Ponsen & Looijen





