Assessing the Bacterial Ecology of Organohalide Respiration for the Design of Bioremediation Strategies

THÈSE N^{O} 5379 (2012)

PRÉSENTÉE LE 5 JUILLET 2012

À LA FACULTÉ DE L'ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT
LABORATOIRE DE BIOTECHNOLOGIE ENVIRONNEMENTALE
PROGRAMME DOCTORAL EN ENVIRONNEMENT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Noam SHANI

acceptée sur proposition du jury:

Prof. F. Porté Agel, président du jury
Prof. C. Holliger, Dr P. Rossi, directeurs de thèse
Prof. H. Smidt, rapporteur
Prof. T. M. Vogel, rapporteur
Prof. J. Zeyer, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2012

Overview

Résumé 3

Summary 5

CHAPTER 1

General introduction 7

CHAPTER 2

Impact assessment of pumping parameters on groundwater bacterial communities 31

CHAPTER 3

PyroTRF-ID: a novel bioinformatics approach for the identification of terminal-restriction fragments using microbiome pyrosequencing data 45

CHAPTER 4

Accumulation of lower chlorinated ethenes in a contaminated aquifer due to competition between iron- and organohalide-respiring bacterial guilds 89

CHAPTER 5

Incomplete reductive dechlorination of tetrachloroethene as a result of hydrogeological characteristics of the Lyss aquifer 129

CHAPTER 6

Concluding remarks and outlook 167

Table of content 177

Remerciements 181

Curriculum vitae 183

Résumé

Parfois considérée comme une ressource naturelle non renouvelable au même titre que les énergies fossiles, l'eau souterraine est essentielle aux nombreuses activités humaines. Le développement industriel et la croissance de la population humaine ont progressivement mis en danger les réserves d'eau souterraine mondiales dès le début du 20^{e} siècle. A l'heure actuelle, les éthènes chlorés (CEs), dont le tetrachloroéthène (ou perchloroéthène, PCE) et le trichloroéthène (TCE) sont les représentants les plus communs, appartiennent à la classe des polluants halogénés les plus fréquemment rencontrés dans les eaux souterraines. De par leurs propriétés chimiques, ces deux composés sont récalcitrants à toute forme de dégradation en conditions oxiques et constituent de fait une source de pollution à long terme de ces environnements.

La possibilité d'une dégradation biologique de ces composés a été démontrée en conditions anoxiques. Les CEs peuvent ainsi être utilisés comme accepteurs finaux d'une chaîne de transporteurs d'électrons dans un processus respiratoire anaérobie (en anglais OrganoHalide Respiration, ou OHR). Le PCE est alors réduit de manière séquentielle en TCE, puis principalement en cis-dichloroéthène (cDCE), en chlorure de vinyle (VC), pour produire finalement de l'éthène (Eth) non toxique par l'abandon successif des atomes de chlore. Des stratégies de remédiation basées sur ce processus, telle que l'Atténuation Naturelle Contrôlée, ont vu le jour afin d'éliminer les CEs des aquifères contaminés de manière biologique. Cependant, un blocage de la séquence de déchloration des CEs a fréquemment été observé lors de la mise en œuvre de ces stratégies. La réduction incomplète produit dans ce cas une accumulation *in situ* de molécules-filles fortement toxiques (cDCE et VC). Dans l'écosystème que constitue un aquifère, les bactéries responsables de l'OHR interagissent de manière complexe avec les membres des autres guildes bactériennes afin de se pourvoir en sources d'électrons et de carbone. En même temps, elles entrent en compétition avec d'autres bactéries dont le métabolisme est centré sur l'utilisation d'accepteurs finaux d'électrons (TEAPs) différents. La complexité de ces interactions trophiques rend difficile la compréhension des mécanismes liés à l'accumulation de cDCE et de VC, ainsi que les tentatives de prédiction du résultat de la OHR et la conception de stratégies de remédiation. Cette thèse a pour objectif d'expliquer les raisons du blocage de la réduction des CEs et de l'accumulation de composés peu chlorés *in situ*. Une approche écologique a été développée afin d'appréhender l'écosystème de l'aquifère dans toute sa complexité. Cette approche a été appliquée à deux cas spécifiques d'aquifères contaminés présentant une accumulation de CEs peu chlorés.

La méthodologie proposée ici repose initialement sur une description précise de l'habitat et des communautés microbiennes de l'aquifère. Un questionnement initial s'est porté sur l'éventuel impact des paramètres de pompage de l'eau souterraine sur la structure des communautés bactériennes présentes dans l'échantillon. Les résultats des recherches ont montré que les paramètres liés aux caractéristiques du tuyau n'ont pas d'impact sur les communautés bactériennes apparentes. A l'inverse, l'étude a révélé un impact significatif du débit de pompage sur les communautés bactériennes extraites de l'aquifère. L'analyse de profiling des communautés (T-RFLP) combinée à une stratégie optimisée de l'échantillonnage de l'eau souterraine a permis d'obtenir une image détaillée de la structure des communautés bactériennes présentes

dans ces habitats. Cependant, cette stratégie ne livre pas l'identité des membres des communautés. Un outil bioinformatique appelé PyroTRF-ID a été développé pour combler cette lacune. Le software permet d'affilier avec succès des fragments de restriction terminaux (T-RFs) à des phylotypes précis en couplant les données de T-RFLP et de pyroséquençage. Une fonction additionnelle permet de tester *in silico* un choix d'enzymes de restriction afin de définir une procédure de T-RFLP appropriée.

Les outils techniques et les procédures analytiques développés précédemment ont permis d'étudier avec succès les raisons de l'accumulation de VC, et dans une moindre mesure de cDCE, dans un premier aquifère à porosité d'interstice présentant une faible conductivité hydraulique. Les résultats ont montré que la réduction du VC était supplantée par la réduction du Fe(III) dans certaines parties de l'aquifère. La combinaison des données fournies par PyroTRF-ID et par séquençage a mis en évidence des séquences d'organismes proches de "*Dehalococcoides*" sp., le seul genre actuellement connu pour sa capacité à réduire le cDCE et le VC, ainsi que des bactéries non cultivées du groupe "Lahn Cluster", appartenant à la classe des *Dehalococcoidetes* et décrites auparavant comme des microorganismes déchlorant le PCE jusqu'au cDCE. Une autre séquence majeure, négativement corrélée aux précédentes, a montré une forte similitude avec des organismes du genre *Rhodoferax*, contenant des bactéries ferri-réductrices. De plus, des indications d'une interaction mutualiste entre bactéries ferri-réductrices et sulfato-réductrices ont été réunies. Cette interaction est supposée jouer un rôle important dans l'aquifère étudié en abaissant localement le contenu en Fe(III).

L'importance de la structure et du fonctionnement hydrogéologiques de l'aquifère sur l'élimination des CEs a été démontrée lors de l'étude du second aquifère. Un processus OHR complet a été montré près la zone de source initiale de PCE avec production d'Eth. En aval de la zone, la dégradation était incomplète, comme l'a montré l'accumulation de cDCE, alors même que les conditions redox étaient similaires. De manière concomitante, les résultats ont révélé des communautés bactériennes radicalement différentes dans ces deux zones. La zone proche de la source contenait des populations bactériennes typiques de milieux fortement réduits, parmi lesquelles des bactéries connues pour leurs activités de dégradation des CEs, telles que *Dehalococcoides* sp. Au contraire, la zone aval était peuplée de populations bactériennes typiques d'environnements plus oxydés, telles que des bactéries aérobies, des nitrifiantes, et des dénitrifiantes. Les conditions dans cette zone étaient probablement faiblement oxydées de manière périodique par l'influence combinée de la recharge de l'aquifère, d'une configuration lithologique spécifique, ainsi que par une teneur en matière organique plus faible.

Enfin, sur la base de ces résultats, une méthodologie pour l'étude des raisons de l'accumulation des CEs faiblement chlorés dans les aquifères contaminés est proposée et discutée. Elle se base sur une description écologique multidisciplinaire de l'écosystème et permet de formuler des scénarii expliquant les raisons de l'accumulation des CEs basés sur un modèle conceptuel du fonctionnement de l'aquifère.

Mots-clés: déhalorespiration, TEAP, éthènes chlorés, contamination des eaux souterraines, éthènes chlorés, chlorure de vinyle, bioremédiation, Atténuation Naturelle Contrôlée, statistiques multivariées, PyroTRF-ID.

Summary

Groundwater is essential for human activities and is sometimes referred to as a non-renewable resource in the same way as oil and gas. During the past century, precious groundwater reserves worldwide have been threatened by anthropogenic release of chemical compounds. Among these, chlorinated ethenes (CEs) such as tetrachloroethene (PCE) and trichloroethene (TCE) belong to the most common class of groundwater contaminants. In oxic conditions, PCE and TCE are recalcitrant to any form of degradation and constitute a long-term source of groundwater contamination.

It has been reported that CEs can be degraded biologically in anoxic conditions, serving as electron acceptors of an anaerobic respiratory process called organohalide respiration (OHR). In this process, PCE is sequentially reduced to TCE, cis-dichloroethene (cDCE), vinyl chloride (VC), and finally to harmless ethene (Eth). Engineering strategies based on reductive dechlorination, such as monitored natural attenuation (MNA), have been designed for the bioremediation of aquifers contaminated with CEs. However, stalling of the sequence of dechlorination has often been observed, resulting from incomplete or impeded biodegradation of the highly toxic daughter molecules (cDCE and VC) and leading to their accumulation *in situ*. It was shown in laboratory experiments that OHR of CEs is more efficient in mixed cultures and that organohalide-respiring bacteria (OHRB) live in association with other microorganisms in microbial consortia. In the aquifer ecosystem, OHRB rely on complex interactions between members of bacterial communities for their electron donors and carbon supplies. On the other hand, they are in competition for these resources with other terminal electron-accepting processes (TEAPs). These complex interactions together with the intrinsic structural heterogeneity of aquifers make it difficult to understand the reasons for lower CEs accumulation, to predict the fate of OHR, and to design bioremediation strategies. This thesis aimed at elucidating the reasons for lower CEs accumulation *in situ*. An ecological approach considering the aquifer ecosystem in its whole complexity was developed and applied to the specific cases of two contaminated sites showing accumulation of lower CEs.

The proposed methodology relies on the precise depiction of both aquifer habitat and microbial communities interacting therein. However, the potential impact of the first technical step, namely the impact of the pumping parameters on groundwater samples used for the description of the bacterial communities, was not known at the start of this thesis work. Results of investigations addressing this topic showed that parameters related to the tubing characteristics did not impact the apparent bacterial community structures (BCS) in a laboratory experiment. However, the study revealed a significant impact of the pumping flow rate on apparent BCS extracted from the aquifer.

Terminal-restriction fragment length polymorphism (T-RFLP), coupled with an appropriate groundwater sampling strategy, enabled an accurate profiling of the BCS. However no indication was provided concerning the identities of the community members. A bioinformatics tool called PyroTRF-ID was developed to overcome this obstacle. The software enabled affiliating terminal-restriction fragments (T-RFs)

to precise phylotypes by coupling T-RFLP and pyrosequencing data. An additional function enables *in silico* assessment of restriction enzymes for designing appropriate T-RFLP procedures. The developed methodology was applied all along the thesis work.

Optimized tools and procedures were used to investigate the reasons for the accumulation of VC, and to a lesser extent cDCE, in the first aquifer showing relatively homogeneous lithological composition and slow groundwater fluxes. Statistical findings indicated that VC reduction was outcompeted by Fe(III) reduction in some sections of the aquifer. T-RFs showing significant correlations with VC reduction variables were identified by sequencing and with PyroTRF-ID. Results showed sequences closely affiliated to uncultured bacteria of the "Lahn Cluster" (LC) within the class *Dehalococcoidetes*, previously reported as PCE-to-cDCE dechlorinating microorganisms, and to the genus "*Dehalococcoides*". According to present knowledge, only members of this genus are able to reduce cDCE and VC. A major T-RF negatively correlated with the LC T-RF affiliated to the genus *Rhodoferax*, containing iron-reducing bacteria (IRB). Furthermore, indications were obtained of a mutualistic interaction between IRB and sulfate-reducing bacteria, potentially playing an important role by reducing the Fe(III) contents locally.

The importance of the aquifer hydrogeological structure and functioning was exemplified in the study of the second aquifer ecosystem, where discrepancies in the fate of OHR were observed. In the immediate vicinity of the source zone, complete OHR was occurring, with production of Eth, whereas cDCE was accumulating further downstream, in apparent similar redox conditions. The profiling and metagenomic techniques revealed drastically different bacterial communities in the zones displaying different OHR fate. The source zone displayed bacterial populations typically found in highly reduced environments and involved in OHR, such as *Dehalococcoides* sp. In contrast, the downstream zone showing an accumulation of cDCE displayed bacterial populations typically found in more oxidized environments, such as aerobic bacteria, nitrifiers, and denitrifiers. In this zone, the conditions were probably slightly oxidized periodically by the combination of aquifer recharge, a specific lithological configuration, and lower organic matter content.

Finally, and based on these considerations, a methodology for the investigation of the reasons for lower CEs accumulation in contaminated aquifers is proposed and discussed. This methodology follows on a multidisciplinary ecological approach for the description of the functioning of the aquifer ecosystem, and enables formulating scenarii of the reasons for lower CEs accumulation.

Keywords: organohalide respiration, TEAP, contaminated groundwater, chlorinated ethene, vinyl chloride, bioremediation strategy, monitored natural attenuation, multivariate statistics, PyroTRF-ID.

General introduction

1 General introduction

1.1 General aspects of groundwater pollution with chlorinated ethenes

1.1.1 General considerations

Groundwater represents about 97% of the existing freshwater resources, excluding the water accumulated in the form of ice (*1*). Used for many purposes by human societies, water is also an essential component of all life forms and activities. In Switzerland, it has been estimated that more than 80% of drinking and process water is obtained from groundwater sources (*2*). Invisible above ground, but seemingly constantly replenished, groundwater has long been considered as an inexhaustible resource (*1*). Since groundwater renewal in many groundwater systems is very slow (from a few years up to more than 100'000 years), groundwater is sometimes referred to as a non-renewable resource (*1*). An increasing demand related to the growth of the world population, to agricultural, and to industrial practices leads nowadays to the depletion of groundwater reserves (*3*). Based on these considerations, groundwater reservoirs are extremely precious. Besides the quantitative aspects of the pressure on this water resource, additional qualitative pressure is exerted. Indeed, numerous substances produced by human activities accumulate in groundwater as a result of agricultural practices, industrial activities, and waste disposal, including careless handling and storage, spills and leakages (*3*). In Europe, around 250'000 contaminated sites were reported as threatening human and ecosystem health in 2006, with potentially 3 million sites on which polluting activities have occurred (*4*). A similar situation exists in the USA where there are approximately 294'000 contaminated sites (*5*).

1.1.2 Naturally occurring organohalides

Organohalides, or halogenated hydrocarbons, are a large class of organic molecules with one or more covalently bound halogen atoms. Several natural sources of organohalides have been described. Some of them are abiotic, as a result of the activities of volcanoes and geophysical processes (*6–9*), whereas many others are of biological origin, in which case the compounds are being produced by physiological activities of Plants, Bacteria , Fungi, Insects, Corals, Sponges, and many others (for a comprehensive review on this topic, see (*10*)). Organohalides are thus present in a wide variety of environments. In 2003, Gribble et al. estimated the number of naturally-occurring organohalides to be larger than 3800 (*10*). Among all compounds, brominated and chlorinated organohalides are the most abundantly naturally produced (*11*).

Chlorinated ethenes (CEs) compose a relatively small subset of the large diversity of organohalide molecules. CEs include tetrachloroethene (or perchloroethene, PCE), trichloroethene (TCE), cis-1,2-

dichloroethene (cis-DCE), trans-1,2-dichloroethene (trans-DCE), 1,1-dichloroethene (DCE), and vinyl chloride (VC). Natural occurrence of PCE and TCE has been described as originating from volcanic activity (*7, 9*) and in oceans (*7, 12*) where marine algae have been shown to be potential PCE and TCE producers (*13*). VC, which has been believed to be a true xenobiotic, has been shown to be naturally produced in significant amounts in terrestrial environments as well (*14*). All CEs have demonstrated toxic and suspected carcinogenic effects on humans and animals (*15*), and all, with the exception of VC, are liquids that have significantly greater density than water. They have low water solubility with tDCE being the most and PCE the less soluble ones, respectively. General physico-chemical properties of CEs are summarized in Table 1.

1.1.3 Industrial use and abuse of chlorinated ethenes

Throughout the last century, anthropogenic production of chlorinated compounds has increased dramatically (*16*). Numerous industries and facilities have been using CEs for dry cleaning, metal degreasing, and for chemical synthesis processes (*17*). In the United States, PCE has become the most common dry cleaning solvent since its introduction in 1934, as a consequence of its higher cleaning ability, of the petroleum shortages during World War II and of its non-flammability properties (*18*). CEs were found equally in private household applications, as solvents and degreasers. CEs have been massively released into the environment during their 40 first years of use, when the consequences both for human health and the environment were not a topical issue and before the advent of major environmental laws and regulations. The widespread release of these compounds has resulted in extensive contamination of groundwater (*17, 19, 20*) and CEs are nowadays considered as the most frequently encountered organic contaminants in groundwater (*21*). However, there is no accurate information available on the amounts of CEs which were released into the environment (*16*) and the degree of contamination is estimated by the number of contaminated sites only. For instance, there are approximately 36,000 active dry cleaning facilities in the United States, and soil and groundwater are contaminated by dry cleaning solvents at about 75% of these facilities (*18*). In Switzerland, it has been estimated that today's generation will be in charge of cleaning-up 3'000 to 4'000 sites contaminated with chlorinated solvents in the next 25 years in order preserve water, soil and air quality (*22*).

Table 1. Physico-chemical properties of chlorinated ethenes [a].

Acronym	**PCE**	**TCE**	**1,1-DCE**	**cDCE**	**tDCE**	**VC**
Common name	Tetrachloroethene	Trichloroethene	1,1-dichloroethene	cis-dichloroethene	trans-dichloroethene	vinyl chloride
IUPAC name	1,1,2,2-tetrachloroethene	1,1,2-trichloroethene	1,1,-dichloroethene	cis-1,2-dichloroethene	trans-1,2-dichloroethene	chloroethene
Molecular weight [g/mol]	165.8	131.4	96,94	97.0	96,94	62.5
Density at 20°C [kg/L]	1.62	1.46	1,22	1.27	1,27	0.91
Solubility in water at 25°C [mg/L]	141	1090	2490	5090	6260	2800
Vapor pressure at 20°C [Pa]	1900	8600	79400	24000	40700	340000
Boiling point [°C]	121.2	87	32	60	48	-13
PNEC [mg/L][b]	0,05	0.15	0,0009	n.a.	n.a.	0.21
Limit in drinking water [mg/L][c]	0.04	0.02	-	0.05 [f]	0.05 [f]	0.0003
Requirement for groundwater [mg/L][d]	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

[a] Federal Office of the Environment (FOEN), Chloronet (http://www.bafu.admin.ch/chloronet/).
[b] PNEC: Predicted no effect concentration
[c] OMS
[d] Federal Office of the Environment (FOEN), OEaux legislation RS 214.801 (http://www.admin.ch).
[e] Sum of the concentrations of cDCE and tDCE.

1.1.4 Dense non-aqueous phase liquids (DNAPLs) and plume of dissolved contaminants in aquifers

Owing to their high density relative to water, as well as to their low solubility, PCE and TCE are immiscible in groundwater and form dense non-aqueous phase liquids (DNAPLs) (*23*). When released in the environment, DNAPLs sink in the subsurface respectively through the unsaturated and saturated zones, until they are trapped by layers of low permeability. Furthermore, portions of DNAPLs are retained during their migration through the subsurface due to hysteretic capillary forces (*24*) resulting in microscopic residues along the migration path (*25*). CEs entrapped as DNAPLs in low permeability aquifer matrix will act as a long-term source of dissolved phase contamination (*26*). In addition, the non-uniform and complex distribution of the DNAPLs (*19*) poses inherent large difficulties to locate both source and residues. Most often, DNAPLs in the subsurface are not found in large amounts, but rather in the form of dispersed residuals trapped in formation pore spaces (*27*). The presence of CEs is consequently typically inferred from dissolved concentrations of DNAPL constituents. In contrast with CEs in the form of DNAPLs, dissolved CEs are transported with groundwater, forming a plume and disseminating along the aquifer. The transport of the contaminants and the extension rate of the plume is influenced by groundwater advection, mechanical dispersion, molecular diffusion, and chemical partitioning between groundwater and porous media (*28*, *29*).

1.2 Organohalide respiration of chloroethenes

1.2.1 CEs can be degraded biologically

The fact that PCE and TCE are strongly oxidized molecules that are recalcitrant to oxidation, together with their high density and low solubility in water, make these molecules persistent in the subsurface for long periods of time, causing plumes of dissolved CEs to remain at concentrations that are many orders of magnitude above the level of concern (*30*). Because of their toxicity, persistence, ubiquitous distribution and abundance, contaminating CEs cause a global threat on groundwater reserves worldwide and subsequent actions must be undertaken to preserve this essential resource (*3*, *31*).

CEs were initially considered as totally recalcitrant to microbial biodegradation due to their persistence in aerobic environments on the one hand, and because their presence in the biosphere was considered as entirely anthropogenic on the other hand (30). Natural biodegradation of CEs was thus commonly viewed as a minor aspect when CE contaminations became a public concern in the 1980's (see (*30*) for a detailed review on this topic). In the mid 1980's however, observations that anoxic groundwater contaminated with PCE and TCE also contained compounds such as DCE and VC that probably were degradation products put forward the idea of microbial reductive dechlorination of CEs (*32–34*). In their extensive review on this topic, Bradley and Chapelle (2010) retrace the chronology of the discovery of the microorganisms which were shown to be involved in the reductive dechlorination process (*30*). Methanogens of the genus

Methanosarcina were first shown to be capable of reductive dechlorination in pure cultures without gaining energy from the reaction (*35*). Involvement of sulfate-reducers and homoacetogens was also reported (*35, 36*). Initial work on reductive dechlorination of CEs reported a slow transformation of these compounds which was thus assumed to be a cometabolic process carried out non-specifically by enzymatic systems dedicated to other metabolic processes (*34, 37*). Although it is generally considered as an inefficient mechanism of chlorinated ethene biodegradation in anoxic environments (*34*), some reports indicated that the global role of cometabolism in the reductive dechlorination of CEs in contaminated sites should not be underestimated (*37*).

1.2.2 Organohalide respiration (OHR) and organohalide-respiring bacteria (OHRB)

A major break-through was brought a decade after the finding that reductive dechlorination of CEs could be carried out by microorganisms. Indeed, in the early 1990's, bacteria conserving energy during reductive dechlorination have been discovered. These bacteria were able to utilize organohalides as terminal electron acceptors in a respiration process (referred to as organohalide respiration (OHR) hereafter). The first described organohalide-respiring bacterium (OHRB) was *Desulfomonile tiedjei* str. DCB-1 (*38*), which grew on 3-chlorobenzoate. Subsequently, CEs were also shown to be used as terminal electron acceptors by OHRB (*39–41*). Much higher rates of reductive dechlorination were reached by OHRB, as compared with the previously reported cometabolic processes (*39, 42*). The process of anaerobic organohalide respiration was elucidated quickly, showing that PCE is reductively dechlorinated in a sequential reaction to TCE, cDCE, VC, and finally to the harmless ethene (Figure 1) (*43*). In each step, chlorine is replaced with a hydrogen atom. The formation of the tDCE isomers is rarely observed (*44*).

The variety of described OHRB and organohalides used as terminal electron acceptors has increased since the initial findings. OHRB guild members have been discovered in a large variety of phyla, such as the *Firmicutes*, *Chloroflexi*, *ε-Proteobacteria* and *δ-Proteobacteria*. Interestingly, there is nowadays no evidence of a link between the phylogenetic affiliation of OHRB and their organohalide terminal electron acceptor (*45*), indicating that this respiration process is possibly a common feature among Prokaryotes.

OHRB which are known to use other electron acceptors than organohalides are referred to as facultative OHRB. The genera *Anaeromyxobacter*, *Desulfitobacterium*, and *Desulfomonile* include species which are all able to reduce organohalide compounds. All these species are equally able to use other electron acceptors (*46*). *Anaeomyxobacter* sp. is able to reduce bromo- and chlorophenols (*47*). *Desulfitobacterium* sp. can grow on reductive dechlorination of a wide range of organohalides, including chloro-, bromo-, and fluorophenols, as well as chloroethanes and -ethenes. Selected strains can reduce PCE to TCE or cDCE (for a complete review of the genus *Desulfitobacterium*, please refer to (*48*)). In addition to its ability of reductively dechlorinating 3-dichlorobenzoate (*38*), .*Desulfomonile tiedjei* is capable of respiratory reductive dechlorination of PCE and TCE (*49*). Other known facultative OHRB guild members are included in genera which in addition contain species that are unable to carry out the reductive dechlorination of organohalide

compounds. For instance, the species *Desulfovibrio dechloracetivorans* is able to reduce chlorophenols (*50*) and has been detected in high abundance in PCE-dechlorinating cultures from tidal flat sediments (*51*). Two *Desulfuromonas* species, namely *D. chloroethenica* and *D. michiganensis*, can reductively dechlorinate PCE and TCE (*52*, *53*). *Sulfurospirillum multivorans* and *S. halorespirans* can reduce PCE to cDCE as well (*54*, *55*). In addition, S. *multivorans* has been shown to grow in PCB-dechlorinating cultures (*56*). *Geobacter lovleyi* str. SZ is capable of reductively dechlorinating PCE and TCE (*57*), and *Geobacter thiogenes* is able to reductively dechlorinate trichloroacetic acid (*58*).

The second category of OHRB, and possibly the most intriguing, is composed of species which strictly require an organohalide molecule as terminal electron acceptor to support their growth. Among them, *Dehalobacter* sp. has first been shown to reductively dechlorinate PCE and TCE to cDCE (*59*). Other studies reported that members of this genus were able to grow on chloroethanes and polybrominated diphenyl ethers (*60*). The highest number of obligate OHRB is present in the *Chloroflexi* phylum. Among this phylum, "*Dehalococcoides"* sp. is undoubtedly the most widely studied. Suspected or demonstrated organohalides used as final electron acceptors by *Dehalococcoides* sp. include polychlorinated biphenyls (PCBs), dioxins, chlorophenols, chlorobenzenes, chloroethanes, and CEs (for a review, please refer to (*61*) and (*62*)). According to current knowledge, only *Dehalococcoides* strains were shown to mediate the last dechlorination steps of the CEs, from DCE to ethene, including strains BAV1, VS, GT, as well as *D. ethenogenes* str. 195 (*41*, *63*, *64*). The latter is the only bacterium known to date that can completely dechlorinate PCE to ethene, although the reduction of VC to ethene is cometabolic. More recently, a close relative of *Dehalococcoides* sp., *Dehalogenimonas lykanthroporepellens*, has been shown to reductively dechlorinate chloropropanes and chloroethanes (*65*). Other close relatives of *Dehalococcoides* sp. are represented by bacterium o-17 and "*Dehalobium chlorocoercia*" DF-1 and have been reported to reductively dechlorinate PCBs and chlorobenzenes (*66*, *67*). In addition, reductive dechlorination of CEs has been reported for uncultured *Chloroflexi* grouped in 2 distinct clusters in the class *Dehalococcoidetes*. Bacteria belonging to the Lahn Cluster (LC) have been reported to dechlorinate PCE to cDCE (*68*) in river sediments microcosms, whereas bacteria belonging to the Tidal Flat Cluster (TfC) have been reported to dechlorinate PCE to tDCE (*69*).

Electron donor requirements, fuelling the respiratory chain of the OHRB guild members, are very diverse (*70*, *71*). Complete reviews of this topic can be found elsewhere (*30*, *62*). Some OHRB are restricted to a single electron donor, such as *Dehalobacter restrictus* (*59*) and *Dehalococcoides* sp. (41), which both utilize exclusively H_2. In contrast, other OHRB such as *Sulfurospirillum* sp. and *Desulfitobacterium* sp. can use a broader spectrum of electron donors (*48*, *54*, *55*). *Desulfuromonas* sp. *(52, 53)* and *Geobacter lovleyi* (*57*) are the only strains that can utilize acetate as electron donor for PCE dechlorination. Furthermore, H_2 generally plays a central role in syntrophic interactions between bacteria oxidizing fermentation products and OHRB under field conditions (*72*), but acetate can act as an intermediate between the syntrophic and OHR guilds in a similar way (*45*). Growth and activity of obligate OHRB therefore depend on the activities of other non-

dechlorinating bacterial guilds, such as fermenters and syntrophs that hydrolyze and ferment organic compounds (*73*).

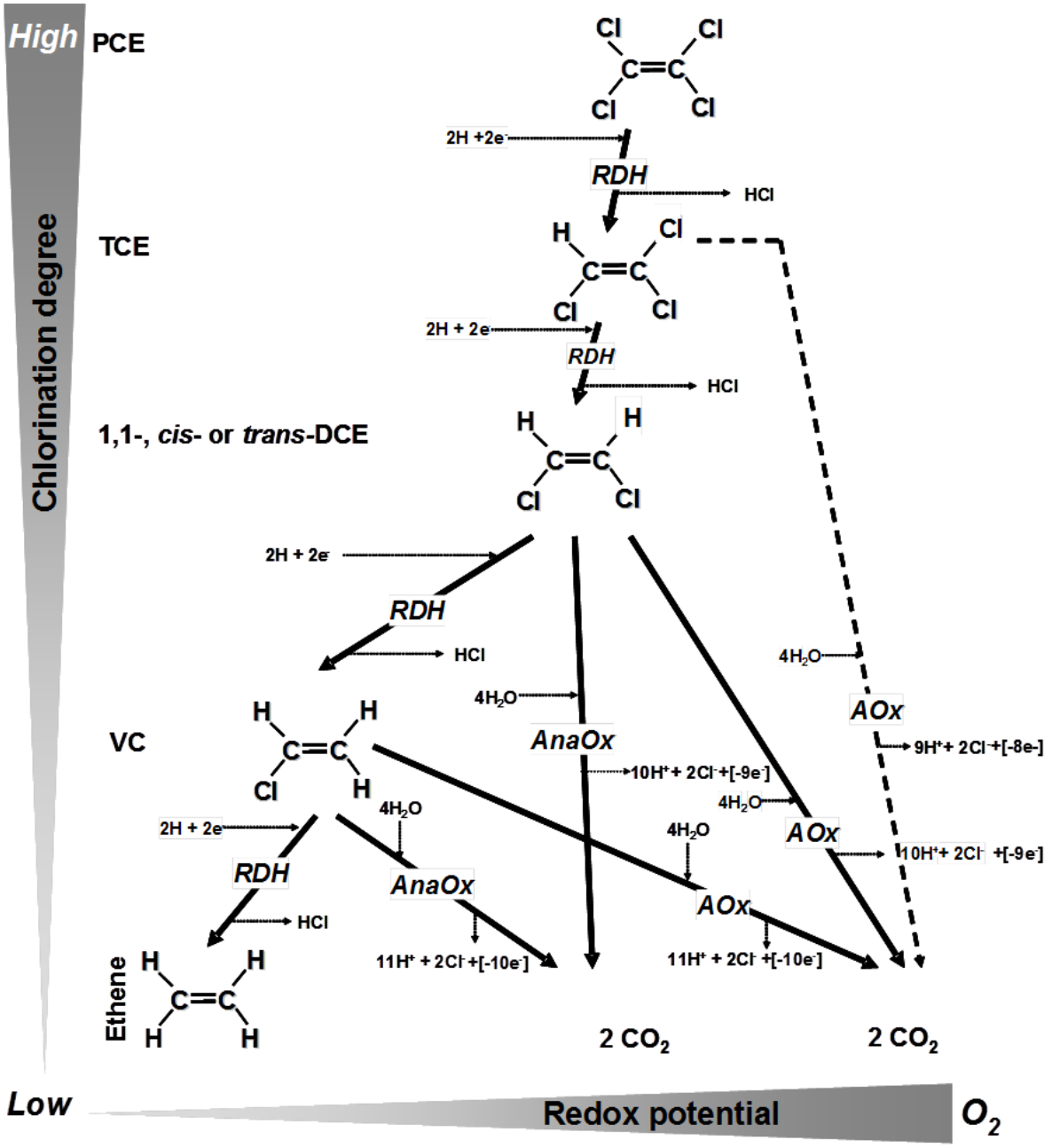


Figure 1. Known pathways of degradation of chlorinated ethenes (reproduced with the kind permission of Gwenaël Imfeld (*74*)). RDH: reductive dehalogenation; AnaOx: anaerobic oxidation; AOx: aerobic oxidation. The efficiency of reductive dechlorination decreases with decreasing chlorination degree. Conversely, the tendency for aerobic oxidation of chlorinated ethenes increases with decreasing number of chlorine atoms of the molecule. Both metabolic and cometabolic oxidation of lower chlorinated ethenes have been reported. However, mineralization of DCE and VC tends to increase with higher reduction potential

1.2.3 *In situ* interactions of the OHR guild with other microbial guilds

Most knowledge regarding the mechanism of CE dechlorination and the physiology of the involved microorganisms has been obtained from laboratory growth systems (*37*, *75*, *76*), whereas only limited information is available for dechlorinating guilds present in contaminated aquifers (*77–79*). There is good evidence that mixed dechlorinating cultures perform better than pure cultures and that OHRB live in close association with other microorganisms in microbial consortia (*80*). Reduction of CEs in aquifers takes place in a highly complex context. Microbial communities in bioremediation scenarii do not only rely on appropriate biocatalytic activities but on complex interactions between members of such communities. It was recently shown that an important bacterial richness was present in CE-contaminated microbial habitats, with members of the acetogens (*Acetobacterium* sp., *Spirochaetes* sp., and *Sporomusa* sp.) as well as members of known fermenters (*Bacteroidetes* sp., *Syntrophus* sp., and *Syntrophobacter* sp.) (*81*, *82*. Miller et al. (2007) showed a detailed analysis of the bacterial community structures (BCS) present in a contaminated aquifer and could identify seven phylogenetic groups, including *Alpha-*, *Beta-*, *Gamma-* and *Deltaproteobacteria*, *Nitrospira, Firmicutes* and *Cytophaga-Flexibacter-Bacteroidetes* using sequencing of the 16S rRNA gene pools (*83*). Members of *Spirochaetes*, *Bacteroidetes* (*84*), *Clostridiales* and *Syntrophobacterales* (*73*) and diverse methanogenic bacteria (*85*) were found together with dechlorinating members of the *Chloroflexi*. Furthermore, several lines of evidence support the hypothesis that sulfate reducers could be involved in the growth of the PCB-dechlorinating bacterial guild (*86*). Slowly growing filamentous bacteria belonging to the *Chloroflexi*-Subphylum I (including members of the genera *Levilinea*, *Belilinea*, *Leptolinea*, *Longilinea* and *Anaerolinea*) are known already to participate in the production of H_2 as the end product of the fermentation of saccharides and peptides (*87*). H_2 is a major electron donor for the OHR process. Therefore, the presence of these bacteria and their relative distribution within the same habitat were thought to be related directly to the OHR process and have an influence on the OHR guild activity. However, the presence of high concentrations of H_2 may influence positively the development of other bacterial guilds such as methanogens, homoacetogens, sulfate reducers, nitrate reducers, and metal reducers and may eventually result in the marginalization of OHRB via competition events. Laboratory experiments suggested that under low H_2 concentration the dechlorinating guild has the potential to outcompete other terminal electron-accepting processes (TEAPs). Furthermore, it was predicted from thermodynamic and kinetic considerations that hydrogenotrophic OHRB should outcompete hydrogenotrophic sulfate reducers, acetogens, methanogens, and even Fe(III) reducers in environments where hydrogen is the main source of electrons, which was confirmed by laboratory experiments and field observation (*88–91*). OHRB can thus function as a TEAP in tight syntrophic interactions with hydrogen-producing populations (*45*, *92*). Interactions between bacterial guilds are depicted in Figure 2.

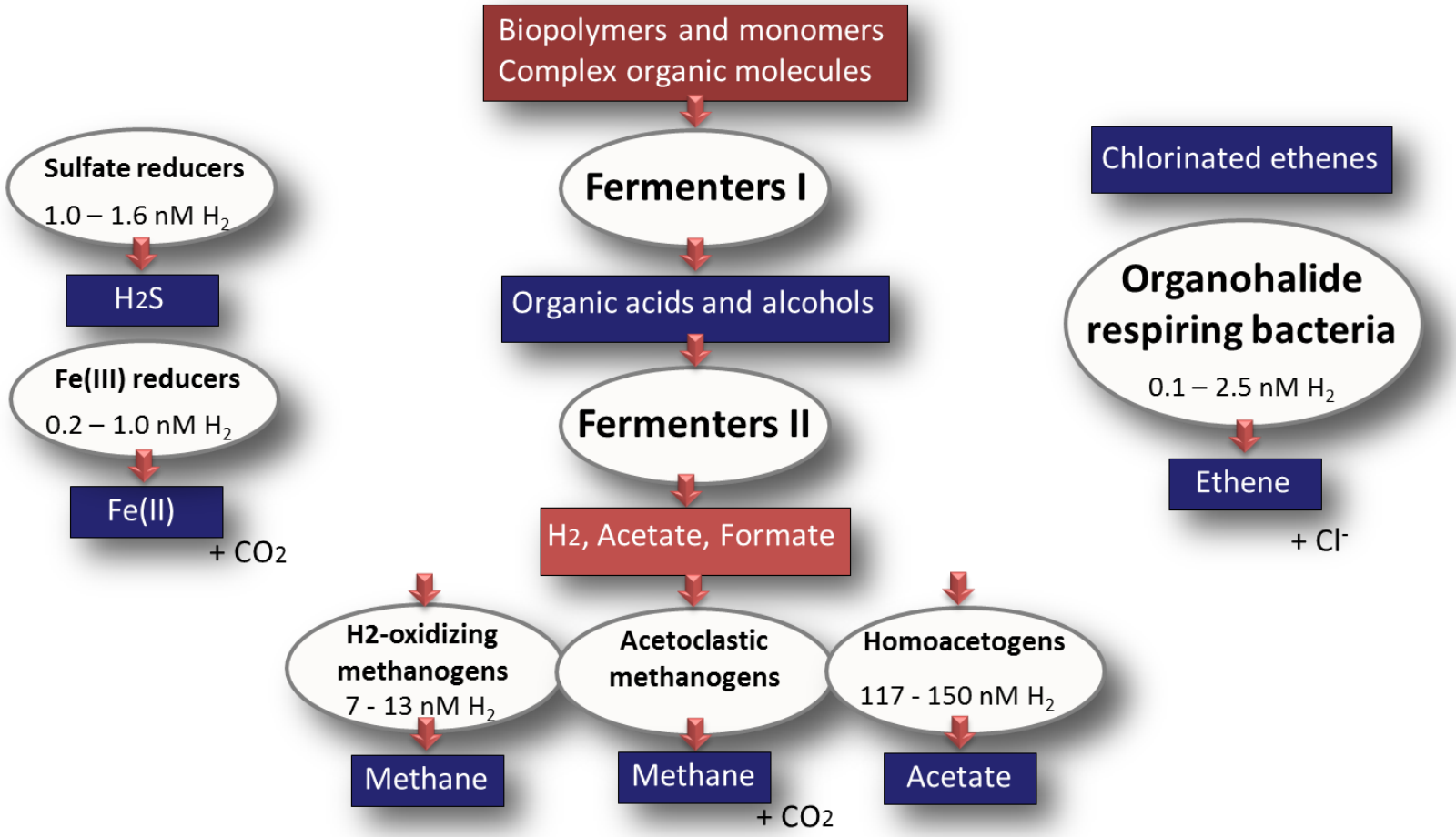


Figure 2. Simplified scheme representing the fluxes of carbon and energy across the main bacterial guilds in anoxic groundwater. The primary fermenters decompose complex organic molecules, releasing organic acids and alcohols, which are in turn used by secondary fermenters, or syntrophs. Among the products of the secondary fermentation, H_2 is the main donor feeding the other bacterial guilds present in the groundwater. H_2 concentrations measured in groundwater during the occurrence of these terminal electron accepting processes (reviewed in (*71*)) are indicated under the name of the bacterial guilds.

1.2.4 Oxidation of chlorinated ethenes

Cometabolic aerobic oxidation has been described for all CEs, though rarely for PCE (*44*). This process results from the fortuitous degradation of CEs by enzymes produced for the degradation of auxiliary growth substrates (*93*).

Metabolic oxidation of CEs has been described under aerobic and anaerobic conditions and generally proceeds at higher rates than cometabolic oxidation (*94*). In both aerobic and anaerobic oxidations, CEs are used as electron donors and are oxidized to CO_2 (*93*). Unlike reductive dechlorination, the ability of CEs to undergo oxidation increases with decreasing number of chlorine substituents (*93*). Aerobic metabolic oxidation has been shown efficient for both cDCE and VC (*95, 96*). Furthermore, it has been recently reported that aerobic metabolic oxidation could occur at very low O_2 concentrations, underlining the potential relevance of this process in contaminated sites (*96, 97*).

Anaerobic oxidation of CEs has first been reported by Vogel and McCarty (*32*). Subsequently, several studies have mentioned that anaerobic oxidation of cDCE and VC can take place under various oxidation-reduction conditions. The mechanism of anaerobic oxidation is still unclear, and the process is still a matter of debate (*30, 98*). It has recently been suggested that mineralization of CEs in anoxic conditions is an artifact and rather results from an aerobic oxidation in suboxic conditions (*96, 97*). A recent microcosm experiment however suggested that VC was oxidized under highly reducing conditions enabling its concomitant reductive dechlorination to ethene (*98*). These considerations are summarized in Table 2.

Table 2. Summary of the processes involved in the biodegradation of chlorinated ethenes (modified from Christ et al., (*19*))

	Metabolism			Cometabolism	
	Oxidation		Reduction	Oxidation	Reduction
	Anaerobic	Aerobic	Anaerobic	Aerobic	Anaerobic
Chlorinated ethenes	cDCE / VC	cDCE / VC / (TCE)	PCE / TCE / cDCE / VC	cDCE / VC	PCE / TCE / cDCE / VC
Relative degradation rates	Unknown	++	++++	+	++
Metabolic group	Fe(III) reducers Mn(IV) reducers Humic acid reducers	*Mycobacterium* sp. *Nocardioides* sp. *Pseudomonas* sp. *Polaromonas* sp.	OHRB	Organisms with broad range oxygenases	Sulfidogens Methanogens Acetogens
Frequency of active organisms in nature	Unknown	VC oxidizers versatile in oxic environments	Frequent in contaminated anoxic environments	High in oxic environments	High in anoxic environments
Favorable site conditions	Fe(III) reducing Mn(IV) reducing	Aerobic	Anaerobic, appropriate e^- donor present, no interfering TEAPs [a]	Aerobic, primary substrate present	Anaerobic, not e^- donor or e^- acceptor limited

[a] Terminal electron-accepting processes.

1.3 Existing remediation technologies and related problems

1.3.1 General aspects of groundwater remediation

Facing the immense challenge of cleaning up the large number of contaminated sites (*5*), several technologies have been developed for the remediation of contaminated groundwater. Classically, plumes of contaminated groundwater are managed by pump-and-treat programs. Other options consist in removing the contaminant by physico-chemical treatment, such as excavation, air sparging, chemical oxidation, and thermal treatment (for a review of physico-chemical approaches, please refer to (*19*)). However, these techniques often require long-term system operation (*100*). In addition, they are technically challenging, costly, and energy demanding. Furthermore, their efficiency is often questioned and rarely meets stringent groundwater cleanup objectives (*5*, *19*). For instance, aggressive contaminant mass removal of the source zone can mobilize and re-distribute the DNAPL. Other technologies, such as conventional air sparging, lack the ability to address sorbed-phase contaminants, often resulting in a rebound effect of contaminant

concentrations upon system shutdown (*19*). Skepticism and concern remain that application of such technologies may not substantially reduce risk and could potentially worsen site conditions.

Combining the above-mentioned techniques for the removal of the contaminating source zone with a biological treatment of the contaminant plume has been discussed elsewhere (*19*) and offers very interesting prospective in terms of both costs and efficiency. As defined by Bradley and Chapelle (2010) "*Biodegradation is the chemical breakdown of contaminant compounds by organisms, often microorganisms, and their associated metabolic processes*". This definition was extended for practical reasons by these authors and includes currently all forms of "*biologically-driven*" degradation categories. By extension, and always according to these authors: "*Any geochemical reaction that can occur abiotically, but is opportunistically catalyzed by microorganisms in order to meet metabolic and/or energetic requirements, would fall under this definition of biodegradation if a net transformation or breakdown of a contaminant compound results*" (*30*).

The word bioremediation is fairly new and appeared first in a peer-reviewed scientific publication in 1987 (*101*). Biostimulation and bioaugmentation are both derived from this term and are currently being used to define the engineered setting-up of conditions promoting the contaminant biodegradation. In the specific case of CE-contaminated aquifers, biostimulation consists in the injection of an electron source, usually an organic substrate, into the contaminated aquifer in order to modify redox conditions and stimulate the degradation process. The activity of the aquifer microbial communities can be efficiently enhanced, driving the aquifer into metal-reducing, sulfate-reducing and methanogenic conditions. In parallel, very high rates of OHR can be reached through the injection of these electron donors, which can provide equally broad-spectrum cometabolic activities. The process of bioaugmentation allows compensating for the lack of appropriate contaminant-degrading organisms which are necessary to conduct the degradation process to its end. Specific contaminant-degrading microorganisms, grown in laboratory conditions are thus injected directly into the groundwater system (*102*, *103*).

1.3.2 Monitored natural attenuation

In many polluted sites active remediation measures are not required or not possible. For instance, the contaminant may not induce any short- or mid-term hazard to drinking water reservoirs, or the configuration of the site may render active measures too complicated or too costly. Also, even though the source zone has been previously efficiently treated, the contaminant plume usually remains and may be too extended to be remediated actively at reasonable cost (*18*, *19*). In such cases monitored natural attenuation (MNA) offers an alternative, or can be efficiently combined with active remediation strategies, for instance as a post-treatment of source zone remediation. Another advantage of MNA is that it brings no or only little perturbation to the site (*105*). MNA is based on the natural biological activity of indigenous specialized microorganisms that degrade the contaminant to a sufficient degree when the geochemical conditions are favorable (*18*). MNA therefore includes a variety of physical, chemical, or biological processes that act without human intervention to reduce the mass, toxicity, mobility, volume, or concentration of contaminants in groundwater

(*30*). Although often referred to as a passive method, MNA requires a comprehensive description of the site characteristics, which in turn will be needed for making predictions about the fate of the contaminant, and a survey of the biodegradation process to ensure efficient contaminant removal. According to Bombach et al. (*5*), the major barrier to the acceptance of MNA as an effective remedial strategy is the difficulty in demonstrating that decrease in the pollutant concentration is indeed caused by biodegradation rather than due to dispersion, sorption, or volatilization.

1.3.3 The aquifer, a complex ecosystem

Aquifers are dynamic ecosystems, usually showing oligotrophic conditions, with complex interactions between physical, chemical and biotic components (*78*). The bacterial community subsists using a complex metabolic network, with input from trace organics within the environment or fixation of atmospheric gases (*105*). Aquifers are extremely heterogeneous, even at very small scale, showing entangled geochemical and grain size conditions (*106*). These complex habitats provide a large panel of living conditions and niches, which in turn allow for the balanced development of interrelated bacterial communities and their inherent activities (*107*). In turn, biological interactions have been shown to impose a spatial zonation in sometimes apparently homogenous geological structures (*108*). Hypothetically, the heterogeneous structure of the habitat offered by the aquifer is likely to provide optimal functioning conditions to the OHR guild members. Several observations suggested that OHR of CEs in groundwater systems could be the result of cooperation between microbial guild members rather than to the activity of individual species (see (*30*) for a review).

Spatial and temporal variations in microbial community structures in CE-contaminated aquifers are governed by factors that extend beyond chemical contaminants and TEAPs, possibly indicating that hydrogeological factors, such as aquifer recharge and heterogeneity in the groundwater fluxes, are also involved (*78*, *109*, *110*). However, hydrological features, such as the aquifer recharge and the inherent impact on the oligotrophic conditions is not fully understood. Because of the oligotrophic nature of many CE-contaminated aquifers, recharge events that deliver electron acceptors as well as soluble organic matter as a source of carbon and electron donors are likely to play an important role in the biological degradation processes (*110*, *111*).

1.3.4 Stalling OHR at cDCE or VC

To date, efforts in evaluating dechlorination processes in aquifers and in monitoring their *in-situ* bioremediation potential provided contrasting results. Indeed, even under very similar environmental conditions, microbial communities from different aquifers showed large discrepancies in their remediation activities (*34*, *112*, *113*). The main drawback of MNA relies in the frequent undesired accumulation of the toxic intermediates cDCE and/or VC which are even more toxic than the parent compound (*30*, *41*, *114*). This accumulation was observed in many aquifer systems where MNA was conducted or even where dechlorinating activities have been enhanced by the addition of an electron donor (*115*).

Although accumulation of cDCE and VC has often been observed under field conditions, the bacterial communities have to our knowledge never been investigated in full details in relation with stalling of OHR *in*

situ. Reasons for lower CEs (DCE and VC) accumulation have been investigated in microcosm experiments or enrichment cultures. For example, Hoelen & Reinhard (*90*) showed that SR affected the rate of lower CEs dechlorination in microcosm experiments. Addition of toluene stimulated H_2 production probably via fermentation, and increased the rate of dechlorination. Similar conclusions were obtained by Heimann et al. (*116*) in mixed anaerobic cultures. Wei et al. (*117*) reported that iron reduction was not inhibiting OHR in sediment batch experiments and enrichment cultures.

The undesired accumulation of these toxic intermediates is believed nowadays to result either from an incomplete sequence of dechlorination, or from a decreased efficiency of lower CEs dechlorination (*30*). Interestingly, the accumulation of cDCE and VC was shown to be fluctuating both in time and space, and was often the result of apparent heterogeneous local degradation conditions. The dependence of the reductive dechlorination of lower CEs on the OHR activity of *Dehalococcoides* sp. has been observed by Hendrickson et al. (*118*).These authors have shown that *Dehalococcoides* sp. was present in 21 contaminated sites where complete dechlorination to ethene was observed and that it was not found at the three sites were dechlorination was not proceeding beyond cDCE, suggesting that its absence at a site may prevent complete OHR. Contrariwise, Flynn et al. (*119*) showed that *Dehalococcoides* sp. was absent from enrichment cultures that dechlorinated PCE to ethene, indicating that simple molecular characterization based on this genus only may underestimate the indigenous OHR potential. Furthermore, studies have revealed the global distribution of OHRB guild members and their occurrence in polluted as well as pristine environments (*120*, *121*). These observations suggest that OHR has evolved as a response to the wide range of organohalides naturally produced on earth, and that the activity of the large variety of OHR enzymatic systems prevents these compounds to accumulate in nature (*45*). Consequently, anthropogenic release of organohalides in the environment would only cause local enrichment of already present OHRB. Furthermore, the apparent absence of OHRB at some sites could be more appropriately attributed to sub-optimal or inappropriate detection techniques.

Incomplete reduction of lower CEs has previously been explained by the lower efficiency of reductive dechlorination with decreasing chlorination degree (*122*). However, the energy available from CE reduction decreases only slightly with decreasing chlorination degree and remains favorable as compared with other TEAPs frequently encountered in groundwater, such as sulfate reduction (*72*, *123*).

According to present state of knowledge, incomplete or impeded reductive dechlorination can be explained by (i) the absence of bacterial species able to dechlorinate down to ethene (*118*, *124*, *125*), (ii) the depletion of the organic substrates used as electron donors by fermenting guilds (*18*, *125*) and (iii) the competition for the electron donors with other TEAPs (*72*, *127*, *128*).

1.4 Objectives and approach of the thesis

To date, the possible reasons for lower CEs accumulation have been assessed separately in laboratory experiments or field observations, and interpreted subsequently. The overall goal of this thesis was to elucidate the reasons for lower CEs accumulation. Specific objectives were (i) to develop an ecological and more holistic approach in considering the aquifer ecosystem in its full complexity, (ii) to assess the reasons for lower CEs accumulation with respects to the interactions of the OHRB guild with other TEAPs directly in the complex aquifer environment, and (iii) to assess the causes of discrepancies observed at a contaminated site in terms of OHR and extent of CE dechlorination.

Hydro-geo-chemical properties of the aquifer delineate the habitat of the bacterial community and provide crucial information on bacterial processes interacting in this environment. Both habitat and bacteriocenosis are indeed intimately interconnected. Bacterial communities are the main drivers of contaminant degradation and are reflecting the general state of the aquifer. Consequently, geological, hydrological, and chemical variables were analyzed and BCS were characterized using terminal-restriction fragment length polymorphism (T-RFLP). Both dynamic habitat and bacterial communities were combined using multivariate statistical tools to provide a comprehensive picture of the contaminated sites and infer the processes taking place, as well as their interactions. Finally, based on the results, tentative conceptual or predictive models were proposed.

Groundwater sampling does not enable getting a comprehensive picture of microbial communities of an aquifer. However, this is the most convenient and realistic way to get numerous samples to follow-up the evolution of BCS in space and time. In **CHAPTER 2**, the impact of groundwater pumping parameters on the characterization of BCS was investigated under laboratory and field conditions. Pumping flow rate did not impact BCS of homogeneous water pumped from a container, in contrast with BCS of groundwater pumped out of an aquifer. In **CHAPTER 3**, a bioinformatics tool, PyroTRF-ID, was developed to enable deeper characterization of bacterial communities with reduced costs and time by combining pyrosequencing and T-RFLP datasets originating from single samples. The tool was tested with different bacterial communities and allowed successfully identifying particular terminal-restriction fragments (T-RFs). In **CHAPTER 4**, the reasons for VC accumulation were investigated in a contaminated site displaying relatively homogeneous grain size distribution and slow groundwater fluxes. Correlations between BCS and environmental variables revealed the main TEAPs occurring at the site and pointed out a potential competition between VC- and Fe(III) reductions. In **CHAPTER 5**, the causes of spatial discrepancies in the fate of CEs were investigated in an aquifer displaying a complex geological structure and pointed out the usefulness of the developed approach for understanding the reasons for incomplete degradation. In **CHAPTER 6,** main aspects of the ecological approach are discussed and proposed for diagnosing contaminated sites and designing adapted bioremediation strategies.

References

1. **UNESCO** (2006) Non-renewable Groundwater Resources: A Guidebook on Socially-Sustainable Management for Water-Policy Makers.
2. **Swiss Federal Office of the Environment (FOEN)** - Groundwater protection. *http://www.bafu.admin.ch/grundwasser/07483/index.html?lang=en*
3. **Danielopol, D. L., Griebler, C., Gunatilaka, A., and Notenboom, J.** (2003) Present state and future prospects for groundwater ecosystems, *Environmental Conservation* 30 (2):104–130.
4. **European Environment Agency (EEA)** (2007) Progress in management of contaminated sites (CSI 015) - Assessment published Aug. 2007, European Environmental Agency (EEA), Copenhagen K, Denmark.
5. **Bombach, P., Richnow, H. H., Kästner, M., and Fischer, A.** (2010) Current approaches for the assessment of in situ biodegradation, *Applied Microbiology and Biotechnology* 86 (3):839–852.
6. **Isidorov, V. A., Zenkevich, I. G., and Ioffe, B. V.** (1990) Volatile organic compounds in solfataric gases, *Journal of atmospheric chemistry* 10 (3):329–340.
7. **Gribble, G. W.** (1994) The natural production of chlorinated compounds, *Environmental Science and Technology* 28 (7):310–319.
8. **Gribble, G. W.** (1998) Naturally occurring organohalogen compounds, *Accounts of Chemical Research* 31 (3):141–152.
9. **Jordan, A., Harnisch, J., Borchers, R., Le Guern, F., and Shinohara, H.** (2000) Volcanogenic halocarbons, *Environmental Science and Technology* 34 (6):1122–1124.
10. **Gribble, G. W.** (2003) The diversity of naturally produced organohalogens, *Chemosphere* 52 (2):289–297.
11. **Häggblom, M. M., and Bossert, I. D.** (2003) Dehalogenation: microbial processes and environmental applications, Springer.
12. **Gribble, G. W.** (1992) Naturally occurring organohalogen compounds - a survey, *Journal of Natural Products* 55 (10):1353–1395.
13. **Abrahamson, K., Ekdah, J., Collén, J., and Pedersen, K.** (1995) Marine algae - a source of trichloroethylene and perchloroethylene, *Limnology and Oceanography* 40 (7):1321–1326.
14. **Keppler, F., Borchers, R., Pracht, J., Rheinberger, S., and Schöler, H. F.** (2002) Natural formation of vinyl chloride in the terrestrial environment, *Environmental Science & Technology* 36 (11):2479–2483.
15. **Henschler, D.** (1994) Toxicity of chlorinated organic compounds: effects of the introduction of chlorine in organic molecules, *Angewandte Chemie International Edition in English* 33 (19):1920–1935.
16. **Stringer, R., and Johnston, P.** (2001) Chlorine and the environment: an overview of the chlorine industry, Springer.
17. **Fetzner, S.** (1998) Bacterial dehalogenation, *Applied Microbiology and Biotechnology* 50 (6):633–657.
18. **ITRC (Interstate Technology & Regulatory Council)** (2005) Overview of in situ bioremediation of chlorinated ethene DNAPL source zones, The Interstate Technology and Regulatory Council, Washington D.C.
19. **Christ, J. A., Ramsburg, C. A., Abriola, L. M., Pennell, K. D., and Loffler, F. E.** (2005) Coupling aggressive mass removal with microbial reductive dechlorination for remediation of DNAPL source zones: A review and assessment, *Environmental Health Perspectives* 113 (4):465–477.

20. **Aulenta, F., Canosa, A., Leccese, M., Petrangeli Papini, M., Majone, M., and Viotti, P.** (2007) Field study of in situ anaerobic bioremediation of a chlorinated solvent source zone, *Industrial & Engineering Chemistry Research* 46 (21):6812–6819.

21. **Holliger, C.** (1995) The anaerobic microbiology and biotreatment of chlorinated ethenes, *Current Opinion in Biotechnology* 6 (3):347–351.

22. **Swiss Federal Office of the Environment (FOEN)** - ChloroNet. Situation de départ et but de ChloroNet. *http://www.bafu.admin.ch/chloronet/06273/index.html?lang=fr*

23. **Mercer, J. W., and Cohen, R. M.** (1990) A review of immiscible fluids in the subsurface: Properties, models, characterization and remediation, *Journal of Contaminant Hydrology* 6 (2):107–163.

24. **Lenhard, R. J., Parker, J. C., and Kaluarachchi, J. J.** (1989) A model for hysteretic constitutive relations governing multiphase flow: 3. Refinements and numerical simulations, *Water Resources Research* 25 (7):1727–1736.

25. **Wilson, J. L., Conrad, S. H., Mason, W. R., Peplinski, W., and Hagan, E.** (1989) Laboratory investigation of residual liquid organics from spills, leaks, and the disposal of hazardous wastes in ground water. Final report, April 1986 - August 1989.

26. **Mackay, D. M., and Cherry, J. A.** (1989) Groundwater contamination: pump-and-treat remediation, *Environmental Science & Technology* 23 (6):630–636.

27. **Cohen, R. M., and Mercer, J.J.** (1993) DNAPL Site Evaluation - CRC Press Book.

28. **National Research Council** (2000) Natural Attenuation for Groundwater Remediation, Committee on Intrinsic Remediation, Water Science and Technology Board, The National Academies Press, Washington, D.C.

29. **ITRC** (2003) An introduction to characterizing sites contaminated with DNAPLs, Interstate Technology & Regulatory Council.

30. **Bradley, P. M., and Chapelle, F. H.** (2010) Biodegradation of chlorinated ethenes, In Stroo, H.F., and Ward, C.H. (Eds), *In Situ Remediation of Chlorinated Solvent Plumes*, pp 39–67. Springer Science+Business Media, NY; USA.

31. **Fetzner, S., and Lingens, F.** (1994) Bacterial dehalogenases: biochemistry, genetics, and biotechnological applications, *Microbiological Reviews* 58 (6):641–685.

32. **Vogel, T. M., and McCarty, P. L.** (1985) Biotransformation of tetrachloroethylene to trichloroethylene, dichloroethylene, vinyl chloride, and carbon dioxide under methanogenic conditions, *Applied and Environmental Microbiology* 49 (5):1080–1083.

33. **Barrio-Lage, G., Parsons, F. Z., Nassar, R. S., and Lorenzo, P. A.** (1986) Sequential dehalogenation of chlorinated ethenes, *Environmental Science & Technology* 20 (1):96–99.

34. **McCarty, P. L., Semprini, L., Norris, R., Hinchee, R., Brown, R., Wilson, J., Kampbell, D., Reinhard, M., Bouwer, E., and Borden, R.** (1994) Ground-water treatment for chlorinated solvents, *Handbook of bioremediation.* 87–116.

35. **Fathepure, B. Z., Nengu, J. P., and Boyd, S. A.** (1987) Anaerobic bacteria that dechlorinate perchloroethene, *Applied and Environmental Microbiology* 53 (11):2671–2674.

36. **Terzenbach, D. P., and Blaut, M.** (1994) Transformation of tetrachloroethylene to trichloroethylene by homoacetogenic bacteria, *FEMS Microbiology Letters* 123 (1-2):213–218.

37. **Holliger, C., Regeard, C., and Diekert, G.** (2003) Dehalogenation by anaerobic bacteria. In Häggblom, M.M., and Bossert, I.D. (Eds), *Dehalogenation: microbial processes and environmental applications*, pp 115–158, Kluwer Academic Publishers, Dordrecht, The Netherlands.

38. **Shelton, D. R., and Tiedje, J. M.** (1984) Isolation and partial characterization of bacteria in an anaerobic consortium that mineralizes 3-chlorobenzoic acid, *Applied and Environmental Microbiology* 48 (4):840-848.

39. **Holliger, C., Schraa, G., Stams, A. J. M., and Zehnder, A. J. B.** (1993) A highly purified culture couples the reductive dechlorination of tetrachloroethene to growth, *Applied and Environmental Microbiology* 59 (9):2991–2997.

40. **Krumholz, L. R., Sharp, R., and Fishbain, S. S.** (1996) A freshwater anaerobe coupling acetate oxidation to tetrachloroethylene dehalogenation, *Applied and Environmental Microbiology* 62 (4):4108–4113.

41. **Maymó-Gatell, X., Chien, Y., Gossett, J. M., and Zinder, S. H.** (1997) Isolation of a bacterium that reductively dechlorinates tetrachloroethene to ethene, *Science* 276 (5318):1568–1571.

42. **He, J. Z., Sung, Y., Dollhopf, M. E., Fathepure, B. Z., Tiedje, J. M., and Loffler, F. E.** (2002) Acetate versus hydrogen as direct electron donors to stimulate the microbial reductive dechlorination process at chloroethene- contaminated sites, *Environmental Science & Technology* 36 (18):3945–3952.

43. **Futagami, T., Goto, M., and Furukawa, K.** (2008) Biochemical and genetic bases of dehalorespiration, *Chemical Record* 8 (12):1–12.

44. **Tiehm, A., and Schmidt, K. R.** (2011) Sequential anaerobic/aerobic biodegradation of chloroethenes - aspects of field application, *Current Opinion in Biotechnology* 22 (3):415–421.

45. **Smidt, H., and de Vos, W. M.** (2004) Anaerobic microbial dehalogenation, *Annual Review of Microbiology* 58:43–73.

46. **Hiraishi, A.** (2008) Biodiversity of dehalorespiring bacteria with special emphasis on polychlorinated biphenyl/dioxin dechlorinators, *Microbes and Environments* 23 (1):1–12.

47. **Sanford, R. A., Cole, J. R., and Tiedje, J. M.** (2002) Characterization and description of *Anaeromyxobacter dehalogenans* gen. nov., sp. nov., an aryl-halorespiring facultative anaerobic Myxobacterium, *Applied and Environmental Microbiology* 68 (2):893–900.

48. **Villemur, R., Lanthier, M., Beaudet, R., and Lépine, F.** (2006) The *Desulfitobacterium* genus, *FEMS Microbiology Reviews* 30 (5):706–733.

49. **Fathepure, B. Z., Nengu, J. P., Boyd, S. A.** (1987) Anaerobic bacteria that dechlorinate perchloroethene, *Applied and Environmental Microbiology* 53 (11):2671-2674.

50. **Sun, B. L., Cole, J. R., Sanford, R. A., and Tiedje, J. M.** (2000) Isolation and characterization of *Desulfovibrio dechloracetivorans* sp. nov., a marine dechlorinating bacterium growing by coupling the oxidation of acetate to the reductive dechlorination of 2-chlorophenol, *Applied and Environmental Microbiology* 66 (6):2408–2413.

51. **Lee, J., Lee, T. K., Löffler, F. E., and Park, J.** (2011) Characterization of microbial community structure and population dynamics of tetrachloroethene-dechlorinating tidal mudflat communities, *Biodegradation* 22 (4):687–698.

52. **Krumholz, L. R.** (1997) *Desulfuromonas chloroethenica* sp. nov. uses tetrachloroethylene and trichloroethylene as electron acceptors, *International Journal of Systematic Bacteriol*ogy 47 (4):1262–1263.

53. **Sung, Y., Ritalahti, K. M., Sanford, R. A., Urbance, J. W., Flynn, S. J., Tiedje, J. M., and Löffler, F. E.** (2003) Characterization of two tetrachloroethene-reducing, acetate-oxidizing anaerobic bacteria and their description as *Desulfuromonas michiganensis* sp. nov, *Applied and Environmental Microbiology* 69 (5):2964–2974.

54. **Scholz-Muramatsu, H., Neumann, A., Messmer, M., Moore, E., and Diekert, G.** (1995) Isolation and characterization of *Dehalospirillum multivorans* gen. nov., sp. nov., a tetrachloroethene-utilizing, strictly anaerobic bacterium, *Archives of Microbiology* 163 (1):48–56.

55. **Luijten, M. L. G. C., de Weert, J., Smidt, H., Boschker, H. T. S., de Vos, W. M., Schraa, G., and Stams, A. J. M.** (2003) Description of *Sulfurospirillum halorespirans* sp. nov., an anaerobic, tetrachloroethene-respiring bacterium, and transfer of *Dehalospirillum multivorans* to the genus *Sulfurospirillum* as *Sulfurospirillum multivorans* comb. nov., *International Journal of Systematic and Evolutionary Microbiology* 53 (3):787–793.

56. **He, J., Robrock, K. R., and Alvarez-Cohen, L.** (2006) Microbial reductive debromination of polybrominated diphenyl ethers (PBDEs), *Environmental Science & Technol*ogy 40 (14):4429–4434.

57. **Sung, Y., Fletcher, K. F., Ritalaliti, K. M., Apkarian, R. P., Ramos-Hernandez, N., Sanford, R. A., Mesbah, N. M., and Loffler, F. E.** (2006) *Geobacter lovleyi* sp. nov. strain SZ, a novel metal-reducing and tetrachloroethene-dechlorinating bacterium, *Applied and Environmental Microbiology* 72 (4):2775–2782.

58. **de Wever, H., Cole, J. R., Fettig, M. R., Hogan, D. A., and Tiedje, J. M.** (2000) Reductive dehalogenation of trichloroacetic acid by *Trichlorobacter thiogenes* gen. nov., sp. nov., *Applied and Environmental Microbiology* 66 (6):2297–2301.

59. **Holliger, C., Hahn, D., Harmsen, H., Ludwig, W., Schumacher, W., Tindall, B., Vazquez, F., Weiss, N., and Zehnder, A. J. B.** (1998) *Dehalobacter restrictus* gen. nov. and sp. nov., a strictly anaerobic bacterium that reductively dechlorinates tetra- and trichloroethene in an anaerobic respiration, *Archives of Microbiology* 169 (4):313–321.

60. **Sun, B., Griffin, B. M., Ayala-del-Rio, H. L., Hashsham, S. A., and Tiedje, J. M.** (2002) Microbial dehalorespiration with 1,1,1-trichloroethane, *Science* 298 (5595):1023–1025.

61. **Taş, N., van Eekert, M. H. A., de Vos, W. M., and Smidt, H.** (2010) The little bacteria that can - diversity, genomics and ecophysiology of "*Dehalococcoides*" spp. in contaminated environments, *Microbial and Biotechnol*ogy 3 (4), 389–402.

62. **Yoshida, N., and Katayama, A.** (2010) Biodiversity of bacteria that dechlorinate aromatic chlorides and a new Candidate, *Dehalobacter* sp. Interdisciplinary Studies on Environmental Chemistry — Biological Responses to Contaminants: 65-76

63. **He, J., Ritalahti, K. M., Yang, K.-L., Koenigsberg, S. S., and Löffler, F. E.** (2003) Detoxification of vinyl chloride to ethene coupled to growth of an anaerobic bacterium, *Nature* 424 (3):62–65.

64. **Cupples, A. M., Spormann, A. M., and McCarty, P. L.** (2003) Growth of a *Dehalococcoides*-like microorganism on vinyl chloride and cis-dichloroethene as electron acceptors as determined by competitive PCR, *Applied and Environmental Microbiology* 69 (2):953–959.

65. **Moe, W. M., Yan, J., Nobre, M. F., da Costa, M. S., and Rainey, F. A.** (2009) *Dehalogenimonas lykanthroporepellens* gen. nov., sp. nov., a reductively dehalogenating bacterium isolated from chlorinated solvent-contaminated groundwater, *International Journal of Systematic and Evolutionary Microbiology* 59 (11):2692–2697.

66. **Cutter, L. A., Watts, J. E., Sowers, K. R., and May, H. D.** (2001) Identification of a microorganism that links its growth to the reductive dechlorination of 2,3,5,6-chlorobiphenyl, *Environmental Microbiology* 3 (11):699–709.

67. **Wu, Q., Milliken, C. E., Meier, G. P., Watts, J. E. M., Sowers, K. R., and May, H. D.** (2002) Dechlorination of chlorobenzenes by a culture containing bacterium DF-1, a PCB dechlorinating microorganism, *Environmental Science & Technology* 36 (15):3290–3294.

68. **Kittelmann, S., and Friedrich, M. W.** (2008) Identification of novel perchloroethene-respiring microorganisms in anoxic river sediment by RNA-based stable isotope probing, *Environmental Microbiology* 10 (1):31–46.

69. **Kittelmann, S., and Friedrich, M. W.** (2008) Novel uncultured *Chloroflexi* dechlorinate perchloroethene to trans-dichloroethene in tidal flat sediments, *Environmental Microbiology* 10 (6):1557–1570.

70. **Löffler, F., Cole, J. R., Ritalahti, K. M., and Tiedje, J. M.** (2003) Diversity of dechlorinating bacteria. In *Dehalogenation: microbial processes and environmental applications*, pp 53–88, Kluwer Academic publishers, Dordrecht, The Netherlands.

71. **Löffler, F. E., Sanford, R. A., and Ritalahti, K. M.** (2005) Enrichment, cultivation, and detection of reductively dechlorinating bacteria. In *Environmental Microbiology*, pp 77–111.

72. **Heimann, A., Jakobsen, R., and Blodau, C.** (2010) Energetic constraints on H_2-dependent terminal electron accepting processes in anoxic environments: a review of observations and model approaches, *Environmental Science & Technology* 44 (1):24–33.

73. **D'Angelo, E., and Nunez, A.** (2010) Effect of environmental conditions on polychlorinated biphenyl transformations and bacterial communities in a river sediment, *Journal of Soils and Sediments* 10 (6), 1186–1199.

74. **Imfeld, G.** (2009) Assessment of in situ biodegradation solvents in aquifers and constructed wetlands using an integrative approach. PhD thesis, Faculty of Science, University of Neuchâtel, Switzerland, 178 pp.

75. **Yang, Y., Pesaro, M., Sigler, W., and Zeyer, J.** (2005) Identification of microorganisms involved in reductive dehalogenation of chlorinated ethenes in an anaerobic microbial community, *Water Research* 39 (16):3954–66.

76. **Zhang, J. J., Joslyn, A. P., and Chiu, P. C.** (2006) 1,1-dichloroethene as a predominant intermediate of microbial trichloroethene reduction, *Environmental Science & Technology* 40 (6):1830–1836.

77. **Grandel, S., and Dahmke, A.** (2004) Monitored natural attenuation of chlorinated solvents: assessment of potential and limitations, *Biodegradation* 15 (6):371–386.

78. **Haack, S. K., Fogarty, L. R., West, T. G., Alm, E. W., McGuire, J. T., Long, D. T., Hyndman, D. W., and Forney, L. J.** (2004) Spatial and temporal changes in microbial community structure associated with recharge-influenced chemical gradients in a contaminated aquifer, *Environmental Microbiology* 6 (5):438–448.

79. **Löffler, F. E., and Edwards, E. A.** (2006) Harnessing microbial activities for environmental cleanup, *Current Opinion in Biotechnology* 17 (3):274–284.

80. **Aulenta, F., Majone, M., and Tandoi, V.** (2006) Enhanced anaerobic bioremediation of chlorinated solvents: environmental factors influencing microbial activity and their relevance under field conditions, *Journal of Chemical Technology and Biotechnology* 81 (9):1463–1474.

81. **Duhamel, M., and Edwards, E. A.** (2006) Microbial composition of chlorinated ethene-degrading cultures dominated by *Dehalococcoides*, *FEMS Microbiology Ecology* 58 (3):538–549.

82. **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, *Environmental Science & Technology* 41 (7):2303–2310.

83. **Miller, T. R., Franklin, M. P., and Halden, R. U.** (2007) Bacterial community analysis of shallow groundwater undergoing sequential anaerobic and aerobic chloroethene biotransformation, *FEMS Microbiology Ecology* 60 (2):299–311.

84. **Fagervold, S. K., May, H. D., and Sowers, K. R.** (2007) Microbial reductive dechlorination of aroclor 1260 in Baltimore Harbor sediment microcosms is catalyzed by three phylotypes within the phylum *Chloroflexi*, *Applied and Environmental Microbiology* 73 (9):3009–3018.

85. **Cabirol, N., Perrier, J., Jacob, F., Fouillet, B., and Chambon, P.** (1996) Role of methanogenic and sulfate-reducing bacteria in the reductive dechlorination of tetrachloroethylene in mixed culture, *Bulletin of Environmental Contamination and Toxicology* 56 (17):817–824.

86. **May, H. D., Miller, G. S., Kjellerup, B. V., and Sowers, K. R.** (2008) Dehalorespiration with polychlorinated biphenyls by an anaerobic ultramicrobacterium, *Applied and Environmental Microbiology* 74 (7):2089–2094.

87. **Yamada, T., and Sekiguchi, Y.** (2009) Cultivation of uncultured chloroflexi subphyla: Significance and ecophysiology of formerly incultured chloroflexi "Subphylum I" with natural and biotechnological relevance, *Microbes and Environments* 24 (3):205–216.

88. **Fennell, D. E., and Gossett, J. M.** (1998) Modeling the production of and competition for hydrogen in a dechlorinating culture, *Environmental Science & Technology* 32 (16):2450–2460.

89. **Löffler, F. E., Tiedje, J. M., and Sanford, R. A.** (1999) Fraction of electrons consumed in electron acceptor reduction and hydrogen thresholds as indicators of halorespiratory physiology, *Applied and Environmental Microbiology* 65 (9):4049–4056.

90. **Hoelen, T. P., and Reinhard, M.** (2004) Complete biological dehalogenation of chlorinated ethylenes in sulfate containing groundwater, *Biodegradation* 15 (6):395–403.

91. **Imfeld, G., Aragonés, C. E., Fetzer, I., Mészáros, E., Zeiger, S., Nijenhuis, I., Nikolausz, M., Delerce, S., and Richnow, H. H.** (2010) Characterization of microbial communities in the aqueous phase of a constructed model wetland treating 1,2-dichloroethene-contaminated groundwater, *FEMS Microbiology Ecology* 72 (1):74–88.

92. **Drzyzga, O., and Gottschal, J. C.** (2002) Tetrachloroethene dehalorespiration and growth of *Desulfitobactetium frappieri* TCE1 in strict dependence on the activity of Desulfovibrio fructosivorans, *Applied and Environmental Microbiology* 68 (2):642–649.

93. **Mattes, T. E., Alexander, A. K., and Coleman, N. V.** (2010) Aerobic biodegradation of the chloroethenes: pathways, enzymes,ecology and evolution, *FEMS Microbiology Reviews* 34 (4):445–475.

94. **Sayavedra-Soto, L. A., Gvakharia, B., Bottomley, P. J., Arp, D. J., and Dolan, M. E.** (2010) Nitrification and degradation of halogenated hydrocarbons - a tenuous balance for ammonia-oxidizing bacteria, *Applied Microbiology and Biotechnology* 86 (2):435–444.

95. **Giddings, C. G. S., Liu, F., and Gossett, J. M.** (2010) Microcosm assessment of *Polaromonas* sp. JS666 as a bioaugmentation agent for degradation of cis-1,2-dichloroethene in aerobic, subsurface environments, *Ground Water Monitoring & Remediation* 30 (2):106–113.

96. **Gossett, J. M.** (2010) Sustained aerobic oxidation of vinyl chloride at low oxygen concentrations, *Environmental Science & Technology* 44 (4):1405–1411.

97. **Bradley, P. M., and Chapelle, F. H.** (2011) Microbial mineralization of dichloroethene and vinyl chloride under hypoxic conditions, *Ground Water Monitoring & Remediation* 31 (6):39–49.

98. **Bradley, P. M., Chapelle, F. H., and Löffler, F. E.** (2008) Anoxic mineralization: Environmental reality or experimental artifact? *Ground Water Monitoring and Remediation* 28 (1):47–49.

99 **Smits, T. H. M., Assal, A., Hunkeler, D., and Holliger, C.** (2011) Anaerobic degradation of vinyl chloride in aquifer microcosms, *Journal of Environmental Quality* 40 (3):915–922.

100. **Suthersan, S. S., Payne, F. C., and Nelson, D. K.** (2004) Enhanced reductive dechlorination - A broader perspective, In *Contaminated Soils, Sediments and Water,* Science in the Real World vol.9.

101. **Sloan, R.** (1987) Bioremediation demonstrated at hazardous-waste site, *Oil Gas Journal* 85 (37):61–66.

102. **Adamson, D. T., McDade, J. M., and Hughes, J. B.** (2003) Inoculation of a DNAPL source zone to initiate reductive dechlorination of PCE, *Environmental Science & Technology* 37 (11):2525–2533.

103. **Lendvay, J. M., Löffler, F. E., Dollhopf, M., Aiello, M. R., Daniels, G., Fathepure, B. Z., Gebhard, M., Heine, R., Helton, R., Shi, J., Krajmalnik-Brown, R., Major Jr., C. L., Barcelona, M. J., Petrovskis, E., Hickey, R., Tiedje, J. M., and Adriaens, P.** (2003) Bioreactive barriers: a comparison of bioaugmentation and biostimulation for chlorinated solvent remediation, *Environmental Science & Technology* 37 (7):1422–1431.

104. **Megharaj, M., Ramakrishnan, B., Venkateswarlu, K., Sethunathan, N., and Naidu, R.** (2011) Bioremediation approaches for organic pollutants: a critical perspective, *Environment International* 37 (8):1362–1375.

105. **Barton, H. A., Taylor, M. R., and Pace, N. R.** (2004) Molecular phylogenetic analysis of a bacterial community in an oligotrophic cave environment, *Geomicrobiology Journal* 21 (1):11–20.

106. **Wilson, R., Thornton, S., and Mackay, D.** (2004) Challenges in monitoring the natural attenuation of spatially variable plumes, *Biodegradation* 15 (6):359–369.

107. **Goldscheider, N., Hunkeler, D., and Rossi, P.** (2006) Review: Microbial biocenoses in pristine aquifers and an assessment of investigative methods, *Hydrogeology Journal* 14 (6):926–941.

108. **Bethke, C. M., Ding, D., Jin, Q., and Sanford, R. A.** (2008) Origin of microbiological zoning in groundwater flows, *Geology* 36 (9):739–742.

109. **Vroblesky, D. A., and Chapelle, F. H.** (1994) Temporal and spatial changes of terminal electron-accepting processes in a petroleum hydrocarbon-contaminated aquifer and the significance for contaminant biodegradation, *Water Resources Research* 30 (5):1561–1570.

110. **McGuire, J. T., Long, D. T., and Hyndman, D. W.** (2005) Analysis of recharge-induced geochemical change in a contaminated aquifer, *Ground Water* 43 (4):518–530.

111. **Scow, K. M., and Hicks, K. A.** (2005) Natural attenuation and enhanced bioremediation of organic contaminants in groundwater, *Current Opinion in Biotechnology* 16 (3):246–253.

112. **Bouwer, E. J., Norris, R. D., Hinchee, R. E., Brown, R., McCarty, P. L., Semprini, L., Wilson, J. T., Kampbell, D. H., Reinhard, M., Bouwer, E. J., Borden, R. C., Vogel, T. M., Thomas, J. M., and Ward, C. H.** (1994) Bioremediation of chlorinated solvents using alternate electron acceptors. *In* Handbook of Remediation: 149-175.

113. B**radley, P. M., and Chapelle, F. H.** (1996) Anaerobic mineralization of vinyl chloride in Fe(III)-reducing, aquifer sediments, *Environmental Science & Technology* 30 (6):2084–2086.

114. **Middeldorp, P., Langenhoff, A., Gerritse, J., and Rijnaarts, H.** (2002) In situ biological soil remediation techniques, *Biotechnology for the Environment: Soil Remediation, Vol 3b 3B*, 73–90.

115. **Lorah, M. M., and Voytek, M. A.** (2004) Degradation of 1,1,2,2-tetrachloro ethane and accumulation of vinyl chloride in wetland sediment microcosms and in situ porewater: biogeochemical controls and associations with microbial communities, *Journal of Contaminant Hydrology* 70 (1-2):117–145.

116. **Heimann, A. C., Friis, A. K., and Jakobsen, R.** (2005) Effects of sulfate on anaerobic chloroethene degradation by an enriched culture under transient and steady-state hydrogen supply, *Water Research* 39 (15):3579–3586.

117. **Wei, N., and Finneran, K. T.** (2011) Influence of ferric iron on complete dechlorination of trichloroethylene (TCE) to ethene: Fe(III) reduction does not always inhibit complete dechlorination, *Environmental Science & Technology* 45 (17):7422–7430.

118. **Hendrickson, E. R., Payne, J. A., Young, R. M., Starr, M. G., Perry, M. P., Fahnestock, S., Ellis, D. E., and Ebersole, R. C.** (2002) Molecular analysis of *Dehalococcoides* 16S ribosomal DNA from chloroethene-contaminated sites throughout North America and Europe, *Applied and Environmental Microbiology* 68 (2):485–495.

119. **Flynn, S. J., Loffler, F. E., and Tiedje, J. M.** (2000) Microbial community changes associated with a shift from reductive dechlorination of PCE to reductive dechlorination of cis-DCE and VC, *Environmental Science & Technology* 34 (6), 1056–1061.

120. **Löffler, F. E., Sun, Q., Li, J., and Tiedje, J. M.** (2000) 16S rRNA gene-based detection of tetrachloroethene-dechlorinating *Desulfuromonas* and *Dehalococcoides* species, *Applied and Environmental Microbiology* 66 (4):1369–1374.

121. **Krzmarzick, M. J., Crary, B. B., Harding, J. J., Oyerinde, O. O., Leri, A. C., Myneni, S. C. B., and Novak, P. J.** (2012) Natural Niche for Organohalide-Respiring *Chloroflexi*, *Applied and Environmental Microbiology* 78 (2):393–401.

122. **Major, D. W., Hodgins, E. W., and Butler, B. J.** (1991) Field and laboratory evidence of in situ biotransformation of tetrachloroethene to ethene and ethane at a chemical transfer facility in North Toronto. In *On-site bioreclamation, Processes for xenobiotic and hydrocarbon treatment*, pp. 147–171, Butterworth-Heinemann, Boston.

123. **Dolfing, J.** (2000) Energetics of anaerobic degradation pathways of chlorinated aliphatic compounds, *Microbial Ecology* 40 (1):2–7.

124. **Fennell, D. E., Carroll, A. B., Gossett, J. M., and Zinder, S. H.** (2001) Assessment of Indigenous Reductive Dechlorinating Potential at a TCE-Contaminated Site Using Microcosms, Polymerase Chain Reaction Analysis, and Site Data, *Environmental Science & Technology* 35(9):1830–1839.

125. **Cichocka, D., Nikolausz, M., Haest, P. J., and Nijenhuis, I.** (2010) Tetrachloroethene conversion to ethene by a *Dehalococcoides*-containing enrichment culture from Bitterfeld, *FEMS Microbiology Ecology* 72 (2):297–310.

126. **Rectanus, H. V., Widdowson, M. A., Chapelle, F. H., Kelly, C. A., and Novak, J. T.** (2007) Investigation of reductive dechlorination supported by natural organic carbon, *Ground Water Monitoring & Remediation* 27 (4):53–62.

127. **Cord-Ruwisch, R., Seitz, H.-J., and Conrad, R.** (1988) The capacity of hydrogenotrophic anaerobic bacteria to compete for traces of hydrogen depends on the redox potential of the terminal electron acceptor, *Archives of Microbiology* 149 (4):350–357.

128. **Lovley, D. R., and Goodwin, S.** (1988) Hydrogen concentrations as an indicator of the predominant terminal electron-accepting reactions in aquatic sediments, *Geochimica et Cosmochimica Acta* 52 (12):2993–3003.

Impact assessment of pumping parameters on groundwater bacterial communities

2 Impact assessment of pumping parameters on groundwater bacterial communities

Abstract

Groundwater microbial community samples are traditionally collected using pumping techniques optimized for groundwater chemistry assessment although the impact of groundwater pumping parameters on apparent bacterial community structures (BCS) is not really known. We therefore studied the impact of pumping lift, flow regime and tubing material on BCS, which were analyzed by terminal-restriction fragment length polymorphism (T-RFLP). Ruzicka dissimilarity coefficients were calculated between T-RFLP profiles to assess disparities between BCS. Variations in pumping lift, flow regime and tubing material did not affect the apparent BCS in experiments using a homogenous water body under laboratory conditions showing that the conditions within the tube had no detectable effect on BCS. However, pumping groundwater from aquifer monitoring wells at different flow rates in the field revealed a significant impact on the apparent BCS. Water samples collected from fine sediment were the most affected by the pumping flow rate.

2.1 Introduction

Due to their mixed solid-liquid structure, aquifers are inherently very heterogeneous and complex ecosystems, often limited by carbon, energy and nutrient resources. Knowledge of biocenoses and more specifically of bacterial community structures (BCS) in aquifers is still relatively limited and little is known about the identity and abundances of key bacterial species (*1*). Presently, molecular biology provides highly sensitive and reproducible tools for the in-depth characterization of BCS (*2-4*). However, the power of these techniques contrasts with the scarce knowledge of the biases and variability introduced by sampling techniques.

Core samples (i.e. the solid matrix together with interstitial water) provide the most comprehensive depiction of BCS in aquifers (*5,6*). However, obtaining core samples is not only costly but also labor-intensive. Moreover, a specific location can only be sampled once due to the destructive nature of the technique. In addition, the method of obtaining core samples is known to alter local conditions by generating heat, causing disturbance in the sediment layering and producing ultrafine particles with subsequent increase of solute concentrations (*7-9*).

An alternative technique to acquire samples for microbial analysis consists of leaving a sterile substratum packed in a dedicated container to be colonized by the aquifer microorganisms within already constructed wells (*10*). However, it is quite uncertain that the conditions within the well are representative of the surrounding aquifer and no systematic studies have been conducted to assess this issue (*6*).

Microbiologists are often involved with ongoing hydrogeological studies or only have access to existing field test sites (*6*). Because of this and constraints of obtaining core samples, most research dealing with aquifer microbiology is based on the analysis of BCS of groundwater samples rather than core samples. Procedures to extract groundwater samples with representative chemical composition have been proposed and evaluated since the 1980s (*11-14*), and a comprehensive review of groundwater sampling and related biases has been reported (*15*). It has been shown that tubing materials alter the chemical composition of groundwater by sorption and volatilization of organic compounds (*16,17*) and that the pumping flow rate influences the amount of colloid-associated contaminants (*18,14*). Slow pumping flow rate (<500 mL/min) has been shown to be most effective at obtaining representative groundwater in terms of colloid and contaminant concentrations. Furthermore, peristaltic pumps have been shown to create pressure variations within the tubing, resulting in the volatilization of dissolved compounds which can be amplified by the sampling depth (lift) (*19*).

The above mentioned pumping procedures for chemical analysis of groundwater samples were implemented in the field of environmental microbiology with specific precautions such as disinfection of sampling devices and containers (*7*). However, these sampling techniques have not been verified for the investigation of *in-situ* groundwater microbial communities. In hydrogeological studies, a groundwater monitoring well is usually purged until groundwater physical-chemical parameters reach stable values before sampling is initiated (*11,13,15*). However, it has been shown previously that BCS only stabilized after pumping 36 well volumes,

whereas physical-chemical parameters stabilized after pumping five well volumes from a porous aquifer with a flow rate of 6 L/min (*20*).

Hypothetically, different pumping devices and mechanisms, pumping flow rate and tubing material could impact BCS in two ways. First, BCS may be altered during the transport through the pumping tube due to cell adsorption onto the tubing wall. This process depends on the surface properties of the tubing materials, such as hydrophobicity, lipophobicity, and net surface charge (*21*,*22*). Furthermore, cell lysis can be induced by rapid pressure changes. These effects would be more prone to occur when inappropriate tubing material is being used, when pumping depth increases, or when pumping is carried out using improper devices. Second, BCS may be modified by detachment of free or particle-associated cells from the sediment when groundwater is extracted from the aquifer using different pumping rates.

In this study we first assessed whether the conditions within the tubing were impacting BCS by varying parameters such as lift, flow regime, and tubing material. Second, we evaluated whether the groundwater pumping flow rate had a direct impact on the extraction of BCS from an aquifer. We assumed that if any influence was observed, it would depend on the grain size distribution. This assessment was performed in an aquifer equipped with wells installed at variable depths with different grain size distributions. For both laboratory and field sections of the study, BCS of groundwater samples were assessed using an optimized terminal-restriction fragment length polymorphism (T-RFLP) analysis (*23*) and compared using multivariate statistics.

2.2 Materials & Methods

2.2.1 Laboratory experiment

Batch experimental setup

The batch experimental setup for the assessment of the influence of lift, flow regime, and tubing material on BCS during water pumping through the tubing is depicted in Figure 1. It consisted of a 40-L polypropylene tank of opaque color to avoid the development of phototrophic organisms and covered with a lid. A manual sampler was inserted at the bottom of the tank to collect reference samples (R). Samples (S) were collected via a 10-m long tube inserted through a hole in the lid. The inlet of this tube was located at the bottom of the tank, close to the location where reference samples were collected. An overhead stirrer was installed to ensure a homogeneous well-mixed system. The tank was carefully disinfected with sodium hypochlorite and rinsed with deionized water prior to use. It was filled with industrial water from the tap used at Ecole Polytechnique Fédérale de Lausanne (EPFL) for different purposes where drinking water is not needed. This water originates from Lake of Geneva 600 m offshore below the thermocline at a depth of about 60 m. Water samples from the tank were collected through the tubes using a peristaltic pump (Type P2.52, Eijkelkamp, Giesbeek, The Netherlands).

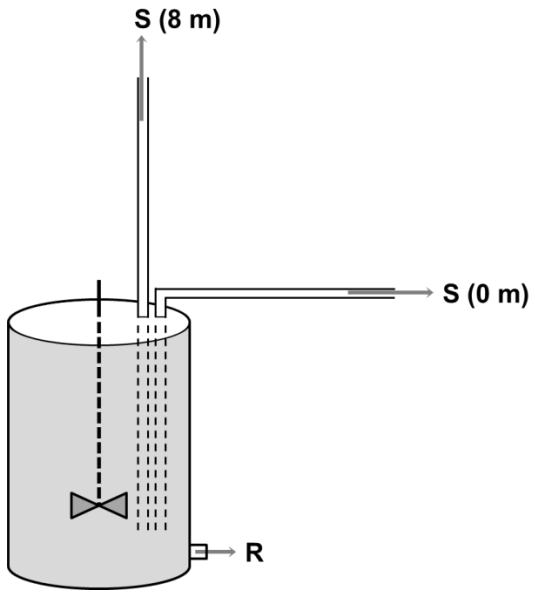


Figure 1. Experimental setup used for testing the influence of lift, flow regime, and tubing material on BCS during pumping.Water was stored in a 40-L tank and homogenized with an overhead stirrer. Samples (S) were collected at 0 and 8 meters high (indicated in brackets) using a peristaltic pump and references (R) were collected manually at the bottom of the tank.

Test conditions

The influence of lift was tested by simulating two depths. For this purpose, height was changed by placing the pump at 0 or 8 m high. Laminar and turbulent regimes were tested with flow rates of 100 mL/min and 1000 mL/min, respectively. This guaranteed to have Reynolds numbers Re <2000 for laminar and Re >4000 for turbulent flow, respectively ($Re = 4Q/\pi d\nu$; Q: flow rate, d: inner diameter of tubing (4 mm), ν: kinematic viscosity) (*24*,*25*). Assuming a water temperature of 10°C, the kinematic viscosity was considered to be 1.307 x 10^{-6} m^2/s. Two tubing materials, polytetrafluoroethylene and low density polyethylene (respectively PTFE and LDPE, Semadeni AG, Switzerland), were chosen because they are frequently used for groundwater sampling and still display different hydrophobicity and lipophobicity properties. All combinations of the three parameters were tested in a total of eight experimental setups (Table 1).

In order to maintain comparable experimental conditions, new 10-m tubes were cut from the same PTFE and LDPE tubing stocks for each test. Tubes were rinsed with 500 mL deionized water (four times their volume) then flushed with one tubing volume with test water. For each experimental condition, samples (500 mL) were collected in triplicate and one reference sample (500 mL) was collected directly from the tank without passing through the tube. The time between two successive experiments was less than 10 min. All samples were stored in 1-L polypropylene Nalgene® bottles (Thermo Fisher Scientific, USA) in the dark at 4°C until DNA extraction, which was done within one day after sampling.

Table 1. Experimental setup to test the impact of lift, flow regime, and tubing material on BCS during pumping.

Experiment n°	I	II	III	IV	V	VI	VII	VIII
Lift [m]	0	8	0	8	0	8	0	8
Flow regime[1]	La	La	Tu	Tu	La	La	Tu	Tu
Tubing material[2]	PTFE	PTFE	PTFE	PTFE	LDPE	LDPE	LDPE	LDPE

[1] Flow regime was either laminar (La) or turbulent (Tu).
[2] The tubing was composed of either polytetrafluoroethylene (PTFE) or low density polyethylene (LDPE).

2.2.2 Field experiment

Test site

The impact of the pumping flow rate on apparent groundwater bacterial communities when sampling monitoring wells was evaluated in an aquifer located in Lyss (Switzerland). The aquifer is composed of two superimposed geological layers of Quaternary depositional origin. The lower layer is composed of coarse gravel with a high hydraulic conductivity (1.10^{-3} ms^{-1}). The upper geological layer is composed of finer material such as fine sand and silt with a low hydraulic conductivity (5.10^{-5} ms^{-1}). The boundary between the two layers is gradual and gives rise to a vertical gradient of grain size distribution. The aquifer is constantly recharged with water originating from the nearby Lyssbach River located ca. 20 m upstream of the test site.

Sampling wells

Four sampling wells (SW1 to SW4) consisting of pre-packed screened piezometers with 1.91 cm internal diameter (3/4" PrePack Well, AMS, USA) were used. They were screened to access different depths of the aquifer that have different hydrogeological properties (Table 2).

Table 2. Hydrogeological characteristics of the Lyss aquifer (Switzerland) at different sampling depths.

Sampling well	Sampling depth (mbs)[1]	Lithology at sampling depth
SW1	3.5 - 4.4	Silt
SW2	4.4 - 5.3	Silty sand
SW3	5.4 - 6.3	Sandy gravel
SW4	6.6 - 7.4	Coarse gravel

[1] mbs = meters below surface

Sampling procedure

Groundwater samples (1 L) were collected using a peristaltic pump (Type P2.52, Eijkelkamp, Giesbeek, The Netherlands) and PTFE tubes (inner diameter 4 mm, Semadeni SA, Switzerland). Each well was purged first at a flow rate of 100 mL/min for at least one well volume until physical-chemical parameters measured on-site (temperature, pH, electrical conductivity, oxidation-reduction potential) reached stable readings. The influence of pumping flow rate on BCS extracted from the aquifer was then tested as follows. The first groundwater sample was collected from each well at a flow rate of 100 mL/min. One day later, a second groundwater sample was collected from the same wells at a flow rate of 1000 mL/min following an identical purging procedure as described above.

To exclude potential temporal bias, variability of BCS was assessed in additional groundwater samples collected from the same wells after a nine-day interval with the same sampling procedure. All groundwater samples were stored in 1-L polypropylene Nalgene® bottles (Thermo Fisher Scientific, USA) in the dark and at 4°C until DNA extraction, which was done within one day after sampling.

2.2.3 Molecular analyses of microbial communities

All groundwater samples were filtered through 0.2-µm autoclaved polycarbonate membranes (Isopore™ Membrane Filters, Millipore, USA) on a filtration system (Filter Funnel Manifolds, Pall Corporation, USA) under a laminar flow hood. DNA from laboratory experiments was extracted as follows. Following filtration, each membrane was washed in a 15-mL Falcon-like tube containing 10 mL of sterile 0.9 g/L NaCl aqueous solution. The tube was shaken on a Vortex Genie II (Scientific Industries Inc., USA) fitted with an adapted holder at maximal speed for 10 min. The membrane was then removed and the tube was centrifuged at 4500 x g for 10 min. The supernatant was discarded and the cell pellet was re-suspended in 1 mL of DNA extraction buffer (5 M Guanidine thiocyanate, 100 mM EDTA, 1% Na-lauroylsarcosine, 1% PVP K30, 150 mM sodium phosphate buffer pH = 8.0). The entire tube volume was transferred into a 2-mL sterile screw cap tube already filled with extracting zirconia-silica beads (MP Biomedicals, USA). Twenty-five µL of freshly prepared 50 mM dithiothreitol (DTT, AppliChem, Germany) was added and the tube was processed in a bead-beater (Bio101, USA) at 6 m/s for 30 s. The tube was centrifuged at 13'000 x g for 2 min and the supernatant was processed for DNA extraction in a Maxwell® 16 Genomic DNA Purification robot (Promega, Madison, USA) fitted with Maxwell® LEV Blood DNA Kit (Promega, Madison, USA) according to the manufacturer's protocol, except that the initial solution of the third cartridge well was replaced by 250 µL of a lysing solution consisting of 60% DNA IQ Casework Pro kit lysis Buffer (Promega, Madison, USA, ref # A8261) and 40% pure ethanol.

DNA from field experiment was extracted using the PowerSoil™ DNA Extraction Kit (MoBio Laboratories, USA) according to the manufacturer's instructions, except that the samples were processed in a bead-beater at 4.5 m/s for 30 s after the addition of solution C1. Extracted DNA was quantified and its quality assessed with a

Nanodrop spectrophotometer (Nanodrop®, ND-1000, Thermo Fisher Scientific, USA). BCS were analyzed using terminal-restriction fragment length polymorphism (T-RFLP). The analytical procedure as well as the numerical treatment of the data was carried out according to (*23*) with the following modifications: (i) 30 µl PCR reactions were composed of 3 µL 10X Y buffer (PeqLab Biotechnologie GmbH, Germany), 2.4 µL 10 mM dNTPs (PeqLab Biotechnologie GmbH, Germany), 1.5 µL of 10 µM FAM-Eub27f forward primer (FAM-5'-GAGTTTGATCMTGGCTCAG-3') (*26*) and unlabeled Univ518r reverse primer (5'-ATTACCGCGGCTGCTGG-3') (*27*), 6 µL 5X enhancer P (PeqLab Biotechnologie GmbH, Germany), 1.5 U PeqGold Taq polymerase (PeqLab Biotechnologie GmbH, Germany) and 0.2 ng/µl template DNA (final concentration), completed with autoclaved and UV-treated Milli-Q water (Millipore, USA); (ii) PCR products were purified using the MSB® Spin PCRapace (Stratec Molecular GmbH, Germany) according to the manufacturer's instructions and digested with the restriction enzyme *Hae*III. For each sample, T-RFLP profiles were obtained in triplicate.

2.2.4 Multivariate statistical analyses

Multivariate statistical analyses were carried out with the R software (*28*) and the package Vegan (*29*). Ruzicka dissimilarity (RD) was calculated between triplicate T-RFLP profiles. RD is an asymmetrical index indicating a statistical distance between two species profiles (T-RFLP profiles in the present study) based on the abundance of each species (T-RF) (*30*). Profiles displaying a RD higher than 15% with their replicates were considered as outliers and were removed from the analysis. Eventually, the profile displaying the lowest dissimilarity with its replicates was selected for further analysis. Kruskal-Wallis one-way ANOVA on ranks was used to test for significant differences between all sets of dissimilarities at the $p = 0.05$ level. For the field experiment, one-way ANOVA with Tukey's test was carried out in order to assess for significant differences between sets of dissimilarities at the $p = 0.05$ level.

2.3 Results

2.3.1 Controlled laboratory experiment with stable BCS of a homogenous water body

The first set of experiments tested the impact of tubing material, flow regime, and lift on BCS. Ruzicka dissimilarities (RDs) were calculated (i) between samples of each experimental condition (Table 1), as well as between samples and the corresponding reference sample in order to assess the variability induced by the tested sampling parameters, (ii) between the triplicate water samples collected successively for each sampling

condition in order to assess the variability induced by the analytical procedure (from DNA extraction to T-RFLP electrophoresis), (iii) between successive reference samples in order to assess the temporal stability of BCS during the experiments, and finally (iv) between all reference samples from the eight successive runs in order to assess the variability induced by the fluctuations of the bacterial communities within the tank in the course of the eight successive experiments. The variability observed between the eight experimental sampling conditions (average RD = 0.16 ± 0.04) was not significantly different from the variability observed between each sample and the corresponding reference sample (0.13 ± 0.03), between sample triplicates (0.17 ± 0.07), between the successive reference samples that were collected at less than ten minutes interval (0.14 ± 0.04), and between all reference samples (0.14 ± 0.03) (Figure 2). This clearly indicated that the BCS were not significantly different before and after the passage of the water through the sampling tubing.

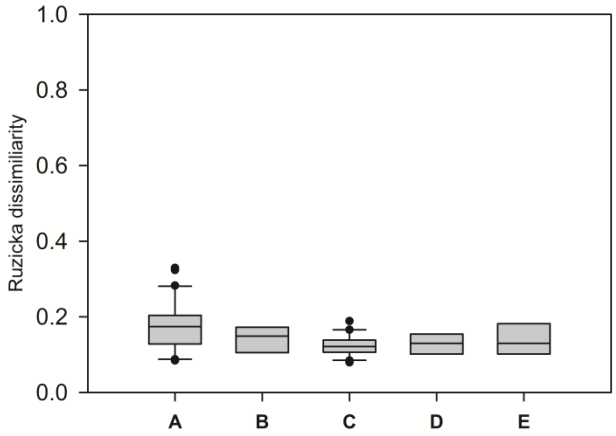


Figure 2. Ruzicka dissimilarities (RDs) calculated between BCS of water samples collected under eight experimental conditions. (A) RDs calculated between sample triplicates; (B) RDs calculated between successive reference samples; (C) RDs calculated between the experimental sampling conditions; (D) RDs calculated between samples and their corresponding reference; (E) RDs calculated between all reference samples.

2.3.2 Field experiment

Groundwater samples were collected from four wells equipped with identical pre-packed piezometers, screened at increasing depths, giving access to specific lithological conditions (Table 2). Natural temporal variations of BCS were assessed first. The temporal variability of BCS was found to increase with depth (Figure 3A). RDs ranged from 0.33 ± 0.04 for well SW1 to 0.75 ± 0.03 for well SW4, with intermediate values for wells SW2 and SW3, with RDs = 0.50 ± 0.04 and 0.64 ± 0.04, respectively. The RDs followed the geological layering of the aquifer, with the lowest RD value in the well SW1, screened at a depth of 3.5 to 4.4 m, giving access to a low permeability zone. The highest RDs were calculated for well SW4, screened at a depth of 6.6 to 7.4 m, which

gives access to a highly permeable geological medium. The differences found between the calculated RDs for each well were highly significant (Tukey $P < 0.001$).

The impact of pumping flow rate on the apparent BCS was assessed by computing RDs between groundwater samples collected with two different pumping flow rates (100 and 1000 mL/min). All RDs computed for field samples were by far higher than those obtained in the laboratory experiments. Values were 0.76 ± 0.01, 0.69 ± 0.01, 0.39 ± 0.03, and 0.75 ± 0.03 for samples collected at wells SW1, SW2, SW3 and SW4, respectively (Figure 3B). These results indicated that pumping rate had a significant impact when sampling geological layers with fine sediment material. This effect was less pronounced in layers composed of coarse material, although the temporal variability of BCS at nine days interval was respectively larger (SW3) and at the same level (SW4) as the variability induced by the change of pumping rate.

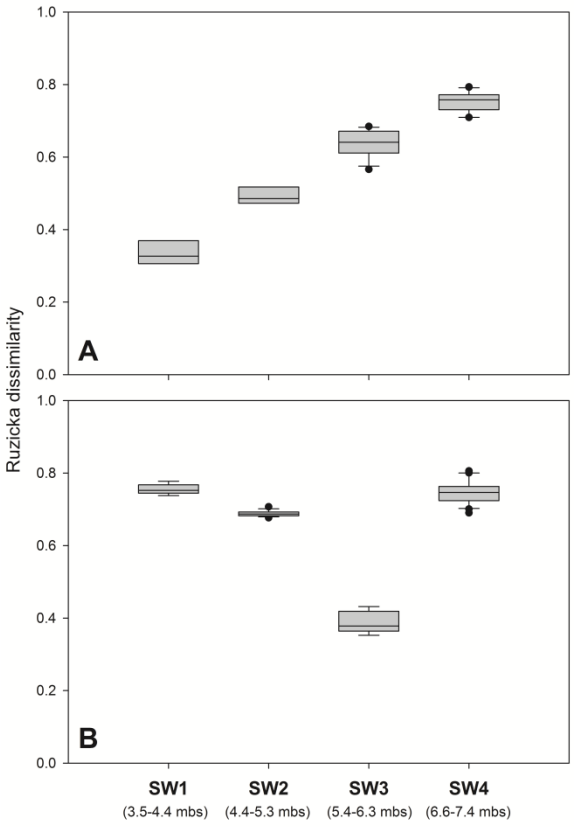


Figure 3. Ruzicka dissimilarities (RDs) calculated for groundwater samples collected from sampling wells SW1, SW2, SW3 and SW4. (A) RDs calculated for samples collected using an identical flow rate at a nine-day interval. (B) RDs calculated for samples collected using two flow rates at a one-day interval (100 and 1000 mL min^{-1}); mbs: depth range of the piezometer screen, in "meters below surface".

2.4 Discussion

Results of this study showed that variations of lift, flow regime, and tubing material were not resulting in any statistically significant difference in the calculated RDs when a stable BCS was pumped from a well-mixed water tank set under laboratory conditions. The dissimilarities induced by the PCR amplification step, source of the highest measured variability of the analytical T-RFLP method used here (0.120 ± 0.143, (*23*)), were in the same range as the dissimilarities between samples of the laboratory experiment. In other words, if any variation of BCS occurred between reference samples and water coming out of the pumping tubes as a consequence of the impact of lift, flow regime, or tubing material, it can be considered to be below the detection limit of the employed molecular fingerprinting technique. It has to be noted that the latter is enabling both a highly sensitive and a robust molecular diversity characterization of bacterial communities (*31,2,3*) and it was hence assumed to detect small variations in BCS. Modern molecular techniques for the characterization of microbial communities rely on DNA extraction and PCR amplification steps, which are sources of biases and inaccuracy (*32,33*). Variations induced by the DNA extraction step were minimized using an automated procedure, but PCR still represents a bottleneck of the overall accuracy of these techniques.

The impact of different pumping rates on BCS under field conditions has only rarely been investigated. A single study showed that BCS were impacted to some extent when using a high flow rate (up to 12 L/min) (*20*) which is by far higher than the rate imposed classically to porous media and superior to the tests which were carried out in the present study. In wells SW1 and SW2, pumping flow rates of 100 and 1000 mL/min resulted in very different BCS (Figure 3B). Both piezometers provided access to the upper zone of the aquifer, composed of fine sediments with low hydraulic conductivity. Discrepancies in the BCS obtained this way can be explained by at least three joint mechanisms. Firstly, the higher flow rate imposed in a zone of low conductivity has been shown to increase the amount of very fine suspended particulate matter in the water sample (*18,14,15*). Visual inspection of the 0.2 µm polycarbonate filters indeed revealed the presence of various amounts of particulate matter retrieved from the pumped piezometers. Secondly, a higher abundance of bacterial cells is located in fine eutrophic sediments (*5*), and more cells are therefore susceptible to be displaced from the sediment surface at higher pumping rate (*20*). Finally, increasing the flow rate can increase the amount of water originating from geological zones that were not reached using a low flow rate. Porous aquifer ecosystems are extremely heterogeneous in terms of lithological composition and grain size distribution, even at a very small scale. Preferential groundwater fluxes within the aquifer are known already to take place as a consequence of differences in the hydraulic conductivity (*5,34,1,20*). Wells SW3 and SW4 gave access to a section of the aquifer composed of coarser sediment and a much higher hydraulic conductivity than the section targeted by wells SW1 and SW2. Based on the results obtained with groundwater from wells SW1 to SW3, the impact of pumping flow rate on BCS was decreasing while the sediment went coarser. In well SW4, the variations induced by the pumping flow rate could not be distinguished from the natural temporal variations of BCS. Reducing the time

interval between two successive samplings with identical pumping flow rate could allow overcoming this lack of resolution.

2.5 Conclusion

The present study showed that selected pumping parameters (lift, flow regime and tubing material) had no detectable impact on BCS when sampling a homogenous water body under laboratory conditions. However, the pumping flow rate had a strong impact on the apparent groundwater BCS when sampling porous aquifer groundwater from *in situ* wells, especially when the targeted sampling zones were composed of fine sediments showing low permeability. This impact was masked partially by rapid natural temporal fluctuations of the autochthonous BCS in coarser sediment.

Acknowledgements

We thank Dr. Jürgen Abrecht (Geotest AG) for test site sampling authorization, Prof. Kiwi Lioubow and Dr. Christophe Roussel for providing the room for the laboratory experiment. This work was supported by the Swiss National Science Foundation Grant # 3100A0-120627.

A modified version of this chapter was submitted to Groundwater Monitoring & Remediation, March 2012 according to the following title:

Noam Shani[1], **Vasantha Aravinthan**[1†], **Pierre Rossi**[2], **Guillaume Laulan**[1], **Pascale Ducommun**[3], **Christof Holliger**[1] (2012) Impact assessment of pumping parameters on groundwater bacterial communities.

1. Ecole Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering, Laboratory for Environmental Biotechnology (LBE), Lausanne, Switzerland

2. Ecole Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering, Central Environmental Laboratory (CEL), Lausanne, Switzerland

3. Centre for Hydrogeology and Geothermics (CHYN), University of Neuchâtel, Neuchâtel, Switzerland

† Present address: Australian Centre for Sustainable Catchments, University of Southern Queensland (USQ), Qld 4350, Australia

References

1. **Goldscheider, N., Hunkeler, D., and Rossi, P.** (2006) Review: Microbial biocenoses in pristine aquifers and an assessment of investigative methods, *Hydrogeology Journal 14*:926–941.
2. **Thies, J. E.** (2007) Soil microbial community analysis using terminal restriction fragment length polymorphisms, *Soil Science Society of America Journal 71:*579–591.
3. **Courtney, K. C., Bainard, L. D., Sikes, B. A., Koch, A. M., Maherali, H., Klironomos, J. N., and Hart, M. M.** (2012) Determining a minimum detection threshold in terminal restriction fragment length polymorphism analysis, *Journal of Microbiological Methods* 88 (1):14–18.
4. **Ekblom, R., and Galindo, J.** (2011) Applications of next generation sequencing in molecular ecology of non-model organisms, *Heredity (Edinb)* 107 (1):1–15.
5. **Alfreider, A., Krossbacher, M., and Psenner, R.** (1997) Groundwater samples do not reflect bacterial densities and activity in subsurface systems, *Water Research* 31 (4):832–840.
6. **Lehman, R. M.** (2007) Understanding of aquifer microbiology is tightly linked to sampling approaches, *Geomicrobiology Journal* 24 (3-4):331–341.
7. **Hirsch, P. & Rades-Rohkohl, E.** (1988) Some special problems in the determination of viable counts of groundwater microorganisms, *Microbial Ecology* 16 (1):99–113.
8. **Morin, R. H., LeBlanc, D. R., and Teasdale, W. E.** (1988) A statistical evaluation of formation disturbance produced by well-casing installation methods, *Ground Water* 26 (2):207–217.
9. **Kim, K.** (2003) Long-term disturbance of ground water chemistry following well installation, *Ground Water* 41 (6):780–789.
10. **Flynn, T. M., Sanford, R. A., and Bethke, C. M.** (2008) Attached and suspended microbial communities in a pristine confined aquifer, *Water Resources Research* 44 (7).
11. **Gibb, J. P., Schuller, R. M., and Griffin, R. A.** (1981) Procedures for the collection of representative water quality data from monitoring wells, *Cooperative Report 7*: Illinois State Water and Geological Surveys, Champaigns.
12. **Barcelona, M. J., Helfrich, J. A., Garske, E. E., and Gibb, J. P.** (1984) A Laboratory evaluation of groundwater sampling mechanisms, *Ground Water Monitoring and Remediation* 4 (2):32–41.
13. **Barcelona, M. J., Gibb, J. P., Helfrich, J. A., and Garske, E. E.** (1985) Practical guide for ground-water sampling, State Water Survey Contract Report #374 to RSKERL-Ada.
14. **Puls, R. W., and Paul, C. J.** (1995) Low-flow purging and sampling of ground-water monitoring wells with dedicated systems, *Ground Water Monitoring and Remediation* 15 (1):116–123.
15. **Barcelona, M. J., Varljen, M. D., Puls, R. W., and Kaminski, D.** (2005) Ground water purging and sampling methods: History vs. hysteria, *Ground Water Monitoring and Remediation* 25 (1):52–62.
16. **Ho, J. S. Y.** (1983) Effect of sampling variables on recovery of volatile organics in water, *Journal of the American Water Works Association* 75 (11):583–6.
17. **Barcelona, M. J., Helfrich, J. A., and Garske, E. E.** (1985) Sampling tubing effects on groundwater samples, *Analytical Chemistry* 57 (2):460–464.
18. **Backhus, D. A., Ryan, J. N., Groher, D. M., Macfarlane, J. K., and Gschwend, P. M.** (1993) Sampling colloids and colloid-associated contaminants in ground-water, *Ground Water* 31 (3):466–479.
19. **Parker, L. V.** (1994) The effects of ground water sampling devices on water quality: A litterature review, *Ground Water Monitoring and Remediation* 14 (2):130–141.

20. **Kwon, M. J., Sanford, R. A., Park, J., Kirk, M. F., and Bethke, C. M.** (2008) Microbiological response to well pumping, *Ground Water* 46 (2):286–294.

21. **Vanloosdrecht, M. C. M., Lyklema, J., Norde, W., Schraa, G., and Zehnder, A. J. B.** (1987) Electrophoretic mobility and hydrophobicity as a measure to predict the initial steps of bacterial adhesion, *Applied and Environmental Microbiology* 53 (8):1898–1901.

22. **Goulter, R. M., Gentle, I. R., and Dykes, G. A.** (2009) Issues in determining factors influencing bacterial attachment: a review using the attachment of Escherichia coli to abiotic surfaces as an example, *Letters in Applied Microbiology* 49 (1):1–7.

23. **Rossi, P., Gillet, F., Rohrbach, E., Diaby, N., and Holliger, C.** (2009) Statistical assessment of variability of terminal restriction fragment length polymorphism analysis applied to complex microbial communities, *Applied and Environmental Microbiology* 75 (22):7268–7270.

24. **Chadwick, A. J., Morfett, J. C., and Borthwick, M.** (2004) Hydraulics in civil and environmental engineering, Spon Press.

25. **Nalluri, C., Featherstone, R. E., and Marriott, M.** (2009) *Nalluri & Featherstone's civil engineering hydraulics: essential theory with worked examples*: Wiley-Blackwell.

26. **Lane, D. J.** (1991) 16S/23S rRNA sequencing. In *Nucleic acid techniques in bacterial systematics*, pp 115-175, John Wiley & Sons, New York.

27. **Muyzer, G., Dewaal, E. C., and Uitterlinden, A. G.** (1993) Profiling of complex microbial populations by denaturing gel electrophoresis analysis of polymerase chain reaction amplified genes coding for 16S ribosomal RNA. *Applied and Environmental Microbiology* 59 (3):695-700.

28. **R Development Core Team** (2008), R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.

29. **Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G. L., Solymos, P., Henry, M., Stevens, M. H. H., and Wagner, H.** (2009) vegan: Community Ecology Package. R package version 1.16-22., *http://R-Forge.R-project.org/projects/vegan/*.

30. **Borcard, D., Gillet, F., and Legendre, P.** (2011) Numerical Ecology with R, Springer.

31. **Tiedje, J. M., Asuming-Brempong, S., Nusslein, K., Marsh, T. L., and Flynn, S. J.** (1999) Opening the black box of soil microbial diversity, *Applied Soil Ecology* 13 (2):109–122.

32. **Delmont, T. O., Robe, R., Cecillon, S., Clark, I. M., Constancias, F., Simonet, P., Hirsch, P. R., and Vogel, T. M.** (2011) Acessing the soil metagenome for studies of microbial diversity, *Applied and Environmental Microbiology* 77 (4):1315-1324.

33. **von Wintzingerode, F., Göbel, U. B., and Stackebrandt, E.** (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis, *FEMS Microbiology Reviews* 21(3):213-229

34. **Chapelle, F. H.** (2001) Ground-water microbiology and geochemistry, John Wiley & Sons, New York.

PyroTRF-ID: a novel bioinformatics approach for the identification of terminal-restriction fragments using microbiome pyrosequencing data

3 PyroTRF-ID: a novel bioinformatics approach for the identification of terminal-restriction fragments using microbiome pyrosequencing data

(The work presented in this chapter was written in collaboration with David Weissbrodt, Laboratory for Environmental Biotechnology (LBE), EPFL, Lausanne, Switzerland)

Abstract

In molecular microbial ecology, pyrosequencing is gradually supplanting classical fingerprinting techniques such as terminal-restriction fragment length polymorphism (T-RFLP) combined with cloning-sequencing for enhanced characterization of microbiomes. Here, the PyroTRF-ID software was developed as a high-throughput approach to combine pyrosequencing and T-RFLP for the description of microbial communities with optimized lab and computational efforts. In contrast to existing bioinformatics methods for phylogenetic affiliation of T-RFs, the proposed strategy aims at conserving the entire microbial information contained in the samples taken from the investigated environments.

PyroTRF-ID was encoded on the Vital-IT high performance computing cluster of the Swiss Institute for Bioinformatics for mapping and massive digital T-RFLP profiling of full pyrosequencing datasets, for comparing digital and experimental fingerprints obtained from the same DNA extracts, and for identifying contributions of phylotypes to T-RFs. The method was used to characterize bacterial communities in groundwater samples from aquifers contaminated by chloroethenes and in aerobic granular sludge biofilm from wastewater treatment systems. Each DNA extract was subjected to amplification of 500 bp fragments of 16S rRNA gene pools, T-RFLP with the *Hae*III endonuclease, 454 tag encoded FLX amplicon pyrosequencing and PyroTRF-ID analysis. Greengenes was selected as mapping database.

PyroTRF-ID was efficient for high-throughput mapping and digital T-RFLP profiling of pyrosequencing datasets. After denoising, 20 min were required to reprocess at least 15 datasets of 6’000 to 35’000 reads of 500 bp. Digital and experimental profiles were aligned with maximum cross-correlation coefficients of 0.71 and 0.92 for high-complexity groundwater and low-complexity synthetic wastewater environments, respectively. On average, 63±23% respectively 61±12% of all experimental T-RFs (73±16 and 47±14 peaks per sample) were affiliated to phylotypes. Bacterial dynamics were then optimally elucidated by T-RFLP.

PyroTRF-ID enables high-throughput matching of pyrosequencing and T-RFLP datasets, and affiliation of T-RFs to precise phylotypes. This methodology is efficient for optimizing laboratory and computational efforts for high-resolved description of microbial community dynamics in various systems such as the ones investigated in environmental and medical sciences.

3.1 Introduction

Molecular microbial ecology has become an important discipline in natural and medical sciences. Research on the structure, dynamics, and evolution of microbial communities in environmental, human, and engineered systems such as soils, groundwater, human or animal bodies, and wastewater treatment plants, can provide substantial scientific knowledge for understanding the underlying microbial processes, for predicting their behavior, and for controlling, favoring or suppressing target populations with dedicated actions (*1*, *2*).

Different bioanalytical methods have been successively developed for the assessment of microbial community structures via profiling or metagenomic approaches (*3*). Terminal-restriction fragment length polymorphism (T-RFLP) has been widely used over the last decade for culture-independent assessment of complex microbial community structures (*4*, *5*). Standardized robust and highly reproducible high-throughput T-RFLP has become efficient for high-quality and high-resolution community fingerprinting since its implementation in automated multicapillary gel electrophoresis devices (*6*, *7*). Cloning and sequencing methods have been developed for empirical taxonomic affiliation of the most important terminal-restriction fragments (T-RF) and have been progressively optimized for cost-efficient isolation and sequencing of large numbers of clones of 16S rRNA encoding genes (*8*, *9*). Cloning and sequencing however remains a time-consuming approach and only leads to insufficient characterization of the genetic diversity (*10*).

After initial development in the early 2000, high-throughput DNA pyrosequencing technologies have recently been intensively applied for deeper and ultrafast analyses of microbiomes with reduced sequencing costs (*11*–*15*). Metagenomics projects have however resulted in novel significant requirements in resource and expertise for generation, storage, processing and interpretation of massive datasets as well as for cloud computing (*16*–*20*). Overall, 'omics' technologies have opened the bioinformatics challenge of designing tailored computing solutions for enhanced production of scientific knowledge from massive raw datasets. According to Field et al. (*21*), the advent of free and open source software, of the *BioLinux* strategy and of the e-Science vision facilitated the access for microbial ecologists to high performance computing solutions for deciphering signals of the surrounding microbial world. Different open-source software have for instance been proposed for processing pyrosequencing data such as MG-RAST (*22*), PyroNoise (*23*), Mothur (*24*), VAMPS (*25*) and QIIME (*26*). To date, no uniform procedure has however been defined by microbial ecologists as standard approach for the processing of massive microbiome datasets.

While pyrosequencing has progressively replaced the traditional combination of T-RFLP and cloning-sequencing approaches, recent studies have demonstrated that the concomitant use of both strategies can be useful and relevant (*27*). For instance, in an investigation of the effect of antibiotic treatment on human indigenous microbiota, Jakobsson et al. (*28*) have used T-RFLP for obtaining a rapid overview of shifts in predominant operational taxonomic units (OTU) and pyrosequencing for in-depth information about changes in relative abundances of thousands of underlying microbial populations. Camarinha-Silva et al. (*29*) have used T-

RFLP and pyrosequencing in an elegant combination for high-throughput profiling of nasal microbial communities. The authors have first assessed best suited combinations of endonucleases for optimal segregation of T-RFs by *in silico* restriction of reference sequences of a set of representative species of the nasal environment with the REPK software (*30*). They have processed samples using concomitantly T-RFLP with the optimal combination of restriction enzymes and pyrosequencing, and have assigned phylotypes to predominant T-RFs by comparing the relative abundances obtained using both techniques. T-RFLP has then been applied for screening for discrepancies in abundant species-related OTUs over 100 volunteers.

Overall, the combination of routine T-RFLP and massive pyrosequencing strategies could represent an efficient compromise between the laboratory efforts required for the analysis of bacterial communities and the financial and infrastructural costs related to datasets processing. Multiple web-based computational procedures have previously been proposed for *in silico* digestion of either artificial communities synthesized from sequences retrieved from public databases or of clone libraries, for prediction of T-RFLP fingerprints and for anticipating phylogenetic affiliation of T-RFs (*31*, *32*). However, inferring the phylogenetic affiliation of T-RFs of an experimental system based on existing databases is conceptually questionable, since it presupposes that the members of the bacterial communities to be analyzed have already been sequenced. In the current study, we propose a novel high-throughput bioinformatics approach to assign phylogenetic affiliations to experimental T-RFs by preserving the original microbial complexity of the samples taken from the investigated environments. For this purpose, the PyroTRF-ID software was developed to compare experimental T-RFLP profiles (eT-RFLP) with digital T-RFLP profiles (dT-RFLP) reconstructed from full pyrosequencing datasets obtained from the same DNA extracts. Main functionalities were encoded for denoising of pyrosequencing datasets, for sequence mapping against reference public databases, for massive dT-RFLP profiling based on the pyrosequencing data, for identifying microbial contributions to dT-RFs, for aligning the eT-RFLP and dT-RFLP profiles and for affiliating eT-RFs to phylotypes present in the samples. An additional functionality was programmed for the screening of restriction enzymes for optimizing the resolution and representativeness of T-RFLP profiles. The PyroTRF-ID procedure was validated using bacterial communities from three different natural and engineered systems, namely groundwater from aquifers contaminated by chloroethenes as well as flocculent activated sludge and aerobic granular sludge biofilms from wastewater treatment systems operated for full biological nutrient removal. The effect of different algorithms for the processing of pyrosequencing datasets and the comparison between eT-RFLP and pyrosequencing-based dT-RFLP profiling is discussed.

3.2 Materials & Methods

3.2.1 Samples

Samples from three different environments were analyzed. The first set of samples consisted of groundwater from two different chloroethene-contaminated aquifers (GRW01 to GRW10) that have been previously described by Aeppli et al. (*33*) and Ducommun et al. (in prep.). The second one consisted of aerobic granular sludge biofilms (AGS01 to AGS07) from anaerobic-aerobic sequencing batch reactors operated for full biological nutrient removal from an acetate-based synthetic wastewater similar to the ones described in Ebrahimi et al. (*34*). The third one was composed of one flocculent activated sludge sample (FLS01) taken from a wastewater treatment plant removing all nutrients biologically. This flocculent sludge served as inoculum for the aerobic granular sludge sequencing-batch reactors.

3.2.2 DNA extraction

Groundwater samples were filtered through 0.2-µm autoclaved polycarbonate membranes (Isopore™ Membrane Filters, Millipore) with a mobile filtration system (Filter Funnel Manifolds, Pall Corporation). DNA was extracted using the PowerSoil™ DNA Extraction Kit (Mo-Bio Laboratories, Inc.) following the manufacturer instructions, except that the samples were processed in a bead-beater (Fastprep FP120, Bio101) at 4.5 m s^{-1} for 30 s after the addition of solution C1.

DNA from flocculent and granular sludge samples was extracted with the automated Maxwell 16 Tissue DNA Purification System (Promega, Duebendorf, Switzerland). An aliquot of 100 mg of grinded granular sludge was preliminarily digested during 1 h at 37°C in 500 µL of a solution composed of 5 mg mL^{-1} lysozyme in a Tris-HCl:EDTA buffer at pH 7.5 ($T_{10}E_{0.1}$ n/n). The DNA extracts were resuspended in 300 µL of $T_{10}E_{0.1}$ buffer.

Extracted DNA was quantified with a ND-1000 Nanodrop® spectrophotometer (Thermo Fisher Scientific, USA) and stored at -20°C until analysis.

3.2.3 Experimental T-RFLP

eT-RFLP analysis of the groundwater bacterial communities as well as their numerical treatment were modified from the protocol described by Rossi et al. (*35*). PCR products were generated in triplicate with a combination of universal labeled FAM-Eub27f forward primer (FAM-5'-GAGTTTGATCMTGGCTCAG-3') (*36*) and unlabeled Univ518r reverse primer (5'-ATTACCGCGGCTGCTGG-3') (*37*) according to the following conditions: 30 µL PCR reactions were composed of 3 µL 10X Y buffer, 2.4 µL 10 mM dNTPs, 1.5 µL of each primer at 10 µM, 6 µL 5X enhancer P, 1.5 U PeqGold Taq polymerase and 0.2 ng/µl template DNA (final

concentration), completed with autoclaved and UV-treated Milli-Q water (Millipore, USA). PCR amplifications were conducted in a T3000 Thermocycler (Biometra GmbH, Germany) as follows: initial denaturing step at 94°C (4.5 min), followed by 25 cycles of 0.5 min denaturation at 94°C, 1 min annealing at 56°C, 1 min elongation at 72°C and a final elongation step of 10 min at 72°C. PCR products were examined in a 1.5% agarose gel to confirm the specificity of the amplification reaction. The PCR products were purified with the purification kit Montage® PCR Centrifugal Filter Devices (Millipore, USA) according to the manufacturer's instructions.

Bacterial communities from granular sludge reactors were analyzed with adaptation of the T-RFLP procedure described by Ebrahimi et al. (*34*). PCR amplification of 500-bp gene fragments was performed in 50 µL reaction mixes: 10 µL Go Taq Buffer 5x (Promega, Switzerland), 4 µL dNTPs 2.5 mM, 2 µL of each 10 µM FAM-8f and 518r primer, 0.25 µL Go Taq polymerase (Promega, Switzerland) and 0.1 ng/µL template DNA (final concentration) completed with 30.75 µL ultrapure DNAse- and RNAse-free water (Qiagen, Switzerland). The PCR program was the same as described above, except that 10-min initial denaturation, 1-min within-cycle denaturation and 30 cycles were required for optimal amplification.

For all types of biomass samples, 200 ng of purified amplicons were digested with 0.5 units of the *Hae*III endonuclease (5'-GG^CC-3' enzyme's recognition site; Promega, Switzerland) at 37°C for 3 h.

The T-RFLP analytes were prepared by mixing 1 µL of digestion product with 8.5 µL of HiDi formamid and 0.5 µL of GeneScan 600-LIZ internal size standard (Applied Biosysems, USA), and were denaturated for 2 min at 95°C. The T-RFs were separated by capillary gel electrophoresis and detected by laser fluorescence in an ABI 3130xl DNA capillary sequencer (Applied Biosystems, USA) equipped with 50 cm long capillaries (80 µm inner diameter) filled with a fluid POP-7 gel matrix (dye set G5). The resulting eT-RFLP profiles were generated between 50 and 500 bp according to Rossi et al. (*35*), which have reported lack of precision in sizing for eT-RFs below 50 bp. The eT-RFLP profiles were aligned using Treeflap (http://www.sci.monash.edu.au/wsc/staff/walsh/treeflap.xls) and expressed as relative contributions of OTUs. For each groundwater sample, which exhibited numerous low abundant populations, the final bacterial community datasets were constructed as follows: Ruzicka dissimilarities were calculated between replicates of eT-RFLP profiles with R (*38*) and the additional package Vegan (*39*); the profile at the centroid (i.e. displaying the lowest dissimilarity with its replicates) was selected for each sample to build the final community profiles. For granular sludge biofilm samples which were characterized by less complex communities, triplicates were periodically measured and resulted in a mean relative standard coefficient of 6% over the analytical method.

3.2.4 Cloning and sequencing

Identification of eT-RFs was carried out using a cloning and sequencing method. To construct the clone libraries, total 16S rRNA gene pool of sample DNA extracts was amplified using the same PCR procedures as described above, except that the 8f primer was not labeled. The PCR products were purified with the purification kit Montage® PCR Centrifugal Filter Devices (Millipore, USA), ligated into pGEM®-T Easy Vector (Promega, USA) and transformed into *E.coli* XL1-Blue competent cells (Agilent Technologies, USA). The T-RFLP procedure was then applied on each clone separately in order to screen for T-RFs. In parallel, the 16S rRNA gene was extracted from the plasmid using a PCR amplification with primers T7 and SP6 (Promega, USA) and purified as described above. A sequencing reaction was carried out on each purified PCR product in a T3000 Thermocycler (Biometra GmbH, Germany) as follows: 10 µL reaction mixtures contained 2 µL of Terminator mix V3.1 (Applied Biosystems Inc., USA), 2 µL of 5X BigDye V3.1 Sequencing buffer (Applied Biosystems Inc., USA), 1.6 µL of primer T7 at 1 µM, 200 ng DNA template and completed with autoclaved and UV-treated Milli-Q water (Millipore, USA). The sequencing reaction was carried out with the following program: 30 cycles of 10 s denaturation at 94°C, 5 s of primer annealing at 50°C, and 4 min of elongation at 60°C. The products were sequenced in an ABI 3130xl DNA capillary sequencer (Applied Biosystems, USA) equipped with 50 cm long capillaries (80 µm inner diameter) and POP 7 electrophoresis matrix (dye set G5, Applied Biosystems Inc., USA). Sequences were aligned in BioEdit (*40*), primers were removed, and sequences were tested for chimeras using Bellerophon (*41*).

3.2.5 Pyrosequencing

DNA from selected samples was sent for bacterial tag encoded FLX amplicon pyrosequencing (bTEFAP) analysis. For all samples, partial amplification of the V3 region of the 16S rRNA gene was performed with the same unlabeled 8f and 518r primers as the ones used for T-RFLP analysis. Amplicons were analyzed in two 454 GS-FLX Titanium Genome Sequencing System devices (Roche, Switzerland). A first set of amplicons from groundwater and granular sludge samples was sent to Research and Testing Laboratory LLC (Lubbock, TX, USA) for bTEFAP 16S diversity analysis on at least 3’000 pyrosequencing reads following the procedure published by Sun et al. (*15*). This method was called in the present chapter the low reads amount method (LowRA). A second set of amplicons from groundwater samples was analyzed by GATC Biotech AG (Konstanz, Germany) following an analog procedure but targeting at least 10’000 pyrosequencing reads. This method was called the high reads amount method (HighRA).

3.2.6 Description of PyroTRF-ID

The PyroTRF-ID software for identification of T-RFs from pyrosequencing datasets was written in python language and was running on the Vital-IT high performance computing cluster (HPCC) of the Swiss Institute of Bioinformatics (Switzerland). The flowchart description of PyroTRF-ID is depicted in Figure 1, and software functionalities are described hereafter. The central procedure of the program is available in *Additional file 1*.

Input files

Input 454 tag-encoded pyrosequencing datasets are introduced in the program either in raw standard flowgram (.sff) or as pre-denoised fast-all format (.fasta). Input datasets of eT-RFLP profiles are provided in coma-separated-values format (.csv).

Denoising

The raw sff files containing the whole pyrosequencing information are first decomposed in sff.txt, fasta and quality (.qual) files using the Mothur software (*24*). QIIME scripts for denoising sequences (*26*) integrated into the software enable treating the files generated by Mothur. Briefly, the script *split_libraries.py* is used first to remove tags and primers. Sequences are then filtered based on two criteria: (i) a sequence length ranging from the minimum (default value of 300 bp) and maximum 500-bp amplicon length set by the user, and (ii) a Phil's-Read-Editor (PHRED) sequencing quality score above 20 according to Ewing and Green (*42*). Denoising for the removal of classical 454 pyrosequencing flowgram errors such as homopolymers (*23*) is carried out with the script *denoise_wrapper.py*. Denoised sequences are then processed using the script *inflate_denoiser_output.py* in order to generate clusters of similar sequences and to define one representative sequence for each cluster (centroid) as well as non-clustering sequences (singletons). A new file is created containing singletons and cluster centroids inflated according to the original cluster size at the species level 97% similarity conventionally used in the microbial ecology community (*43*). The denoising process can optionally be disabled by the user if the dT-RFLP analysis has to be run on raw data.

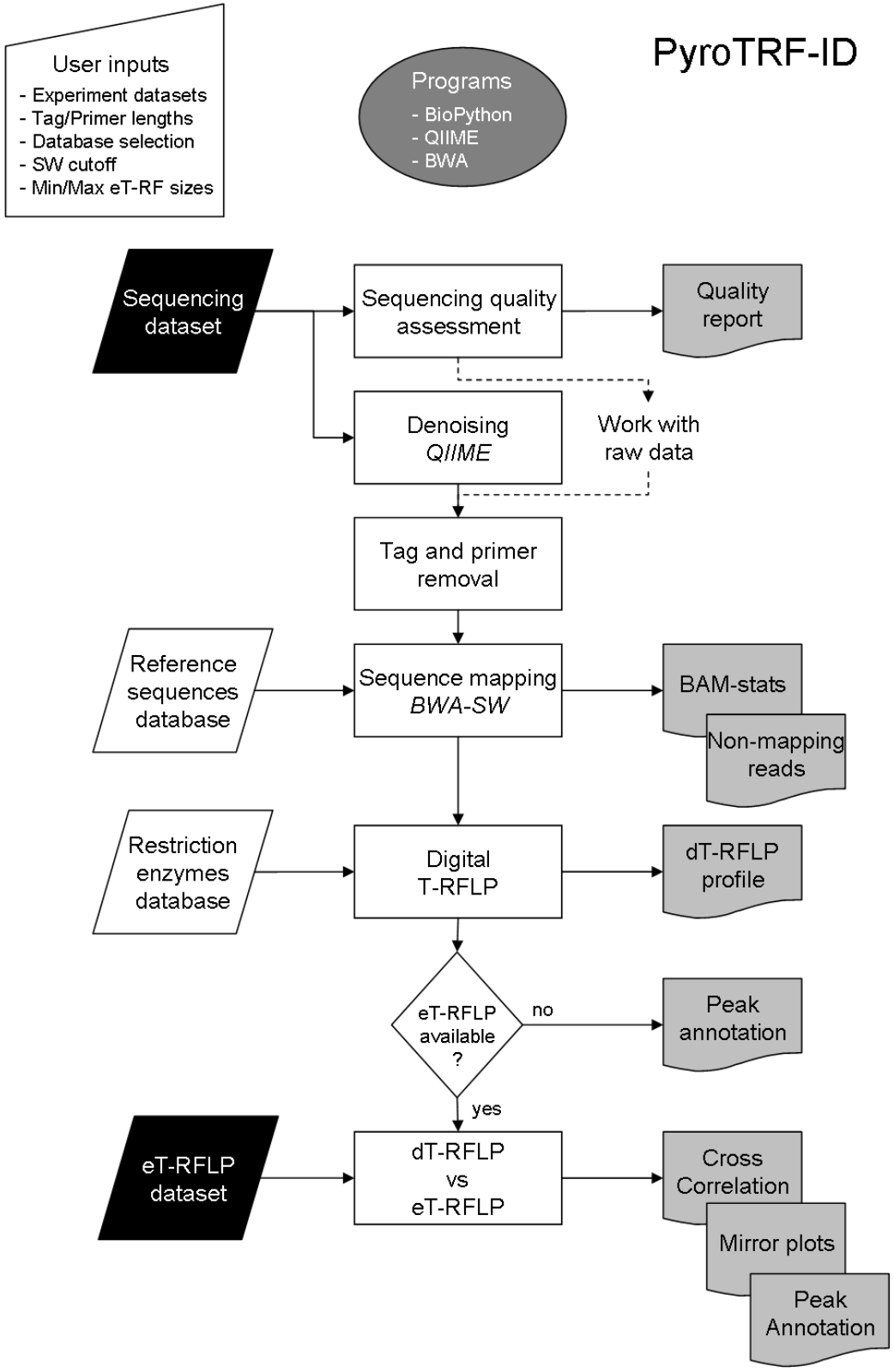


Figure 1. Data workflow in the PyroTRF-ID software. Experimental pyrosequencing and T-RFLP input datasets (*black parallelograms*), reference input databases (*white parallelograms*), data processing (*white rectangles*), output files (*grey sheets*).

Mapping

Mapping of sequences is performed using the Burrows-Wheeler Aligner's Smith-Waterman Alignment (BWA-SW) algorithm for long reads above 200 bp (*44*) against the Greengenes database (*45*). When required, additional public reference sequence databases can be installed and selected prior to launching the PyroTRF-ID software. The Smith-Waterman (SW) score is used as mapping quality criterion. In the determination of the SW score, nucleotides and gaps are taken into account. The highest SW mapping score that can be obtained for a read is the length of the read itself. Sequences with Smith-Waterman (SW) mapping scores below 150 (default value) are removed from the pipeline. The threshold can optionally be modified by the user. Normalized SW scores to sequence lengths are computed as estimates of percentages of identity in order to allow comparison of the sequence mapping quality. For datasets of sequences between 300 and 450 bp, the SW threshold at 150 corresponds to a normalized cutoff at 50% and 33% identity. Two files are generated consecutive to mapping. The first one provides general mapping statistics related to the Sequence Alignment Map converted in Binary format (SAM/BAM files) for each sample. The second one provides the list of unmapped sequences, which are removed from the PyroTRF-ID pipeline.

Generation of dT-RFLP profiles

Remaining sequences are digested *in silico* using selected endonucleases available in the Bio.Restriction biopython database. In the present study, dT-RFLP profiles were produced by *in silico* digestion by default with the *Hae*III restriction endonuclease to allow comparison with corresponding eT-RFLP profiles generated with the same restriction endonuclease. dT-RFLP profiles are generated for each sample considering both the size of the dT-RFs and their relative abundance in the sample. Sequences containing no restriction site for *Hae*III are discarded. A raw dT-RFLP profile plot is generated as output file. A second dT-RFLP profile is generated after removal of T-RFs below 50 bp, in order to allow comparison with eT-RFLP profiles.

Comparison of eT-RFLP and dT-RFLP profiles

The generated dT-RFLP profiles are compared by PyroTRF-ID with the eT-RFLP profiles provided by the user. eT-RFs can be subject to experimental shifts compared with theoretical values due to intrinsic sequence characteristics, such as length, composition, G+C nucleotide content and secondary structure (*46*), as well as because of the different electrophoretic mobility of the fluorophores used for labeling the internal standard and the experimental DNA (*47*). To overcome this discrepancy, the software evaluates the most probable shift between the dT-RFLP and the eT-RFLP profiles by computing estimates of cross-correlation between both profiles. A plot showing the results of the cross-correlation is generated in order to help the user assessing the optimal shift to apply for aligning the dT-RFLP profile to the eT-RFLP profile. By default, PyroTRF-ID shifts the dT-RFLP profile based on the lag related to the best cross-correlation. However, the user can optionally

define a specific lag to apply. After shifting the dT-RFLP data, a mirror plot is generated allowing visual comparison of the dT-RFLP and eT-RFLP profiles.

Assignment of affiliation to dT-RFs

Peak annotation files are generated in Excel format (.xls), listing all dT-RFs composing each dT-RFLP profile, together with their shifted and non-shifted lengths. Closest phylogenetic affiliations are provided together with the number of reads and their relative contribution to the dT-RF, as well as with the absolute and normalized SW mapping scores, and the reference code of each sequence. If the user does not provide any eT-RFLP data, the peak annotation file is directly obtained after dT-RFLP processing without indication of T-RF shift and without removing dT-RFs below 50 bp.

Screening of restriction enzymes

PyroTRF-ID can be used to test restriction endonucleases on pyrosequencing data for designing eT-RFLP procedures intended for the description of bacterial communities from a specific environment using a defined set of primers. To do so, an option allows the user to select single or combinations of enzymes available in the Bio.Restriction biopython database. Only pyrosequencing data should be fed as input file to PyroTRF-ID, as only a dT-RFLP profile is generated.

Overall, richness and Shannon's H' diversity indices were calculated in a way to quantify the impact of the pyrosequencing data processing parameters on the resulting dT-RFLP profiles. The aim was to optimize the cross-correlation between dT-RFLP and the reference eT-RFLP profiles. Overestimated dT-RFLP structures were characterized by richness and diversity values far from the ones of eT-RFLP profiles.

Testing and applications of PyroTRF-ID

After initial tests with eighteen pyrosequencing datasets originating from groundwater, flocculent activated sludge and aerobic granular sludge biofilm samples, the impacts of the data processing steps were assessed along the PyroTRF-ID pipeline with the groundwater sample GRW01, the flocculent sludge sample FLS01 and the aerobic granular sludge sample AGS01. Three different combinations of algorithms were tested for the processing of sequences (Table 1), and their respective impact on the final dT-RFLP profiles was compared. The optimal standard PyroTRF-ID procedure was validated based on this assessment.

The optimal procedure was then applied for the comparison of PyroTRF-ID results obtained from groundwater and wastewater environments. Finally, restriction endonucleases commonly used in T-RFLP analyses of bacterial communities from similar environments as the ones studied here were selected for comparison of profiling resolutions. dT-RFLP profiles were generated with the endonucleases *Alu*I (5'-AG^CT-3'), *Hha*I (5'-GCG^C-3'), *Msp*I (5'-C^CGG-3'), *Rsa*I (5'-GT^AC-3'), *Taq*I (5'-T^CGA-3'), and *Hae*III. Visual observation,

richness and diversity indices as well as density plots were used to analyze the distributions of T-RFs obtained with the different restriction enzymes.

Table 1. – Combinations of algorithms for the processing of pyrosequencing data for construction of dT-RFLP profiles tested in PyroTRF-ID

Pyrosequencing data processing procedure	Processing algorithms				
	PHRED-filtering of sequences	Denoising	Filtering by SW mapping score	Restriction of sequences[a]	Restriction of reference sequences[b]
1) standard dT-RFLP[c]	**Yes (20)**[f]	**Yes**	**Yes (150)**[g]	**Yes**	No
2) filtered dT-RFLP[d]	**Yes (20)**	No	**Yes (150)**	**Yes**	No
3) raw dT-RFLP[e]	**Yes (20)**	No	No (0)[h]	**Ye**s	No

[a] The processed sequences are digested *in silico* with the restriction enzyme.
[b] The sequences from the reference database mapping with the sequences are digested *in silico* with the restriction enzyme.
[c] The first combination with denoising was defined as the standard PyroTRF-ID procedure.
[d] In the second combination, only a filtering method at the mapping stage was considered.
[e] In the third combination, raw datasets of sequences obtained after PHRED-filtering of the pyrosequencer outputs were digested without post-processing.
[f] PHRED score = -10 log P_{error} with $P_{error} = 10^{-PHRED/10}$ as the probability that a base was called incorrectly. For all trials, the raw pyrosequencing datasets directly obtained from the pyrosequencer were systematically filtered after the pyrosequencing PHRED quality score. Only sequences with a related PHRED score higher than 20 were conserved. This corresponds to a P_{error} of 1/100.
[g] A SW mapping score of 150 was set as cutoff. In the case when sequences were preliminarily denoised, it was nevertheless observed that no denoised sequence was rejected at the mapping stage.
[h] Processing without filtering by the SW mapping score was done by setting a cutoff of 0.

3.3 Results

3.3.1 PyroTRF-ID software and performance

The PyroTRF-ID software was successfully encoded on the Vital-IT HPCC of the Swiss Institute for Bioinformatics. PyroTRF-ID was efficient for high-throughput denoising, mapping and digital T-RFLP profiling of pyrosequencing datasets. After the denoising step which can last between 1 and 15 h depending on the size of the dataset, 20 min were required by PyroTRF-ID to process at least 20 pyrosequencing datasets of 6’000 to 35’000 reads of 500 bp in parallel.

3.3.2 Pyrosequencing quality assurance and read length limitation

The principal quality outputs given by PyroTRF-ID are presented in Figure 2 for the two low reads amount (LowRA) and high reads amount (HighRA) pyrosequencing methods used in this study. 6'380 and 32'480 reads were obtained on average per method, respectively. The sequences were related to PHRED score above the quality criterion of 20 for lengths up to at least 400 bp (Figure 2a). For sequence length above 450 bp, the quality score decreased drastically below 20, and the related sequences were discarded. For the LowRA and HighRA methods, the median number of reads (800 and 2'750) was related to a PHRED score of 30 and 35, respectively (Figure 2b). With both methods, more than 99% of reads were related to a PHRED score above 20. Figure 2c was useful to set a sequence length cutoff for the following processing steps. Only sequences longer than 300 bp were conserved for further analysis in order to produce dT-RFLP profiles that were based on similar DNA fragments as the eT-RFLP profiles. For instance, the massive pyrosequencing datasets obtained with the HighRA method were predominantly composed of short reads below 300 bp (69% of a total of 24'810 reads in this example), and therefore these reads were discarded. After discarding this fraction, a large amount of high quality sequences was still available for PyroTRF-ID analysis (7'641 reads, 31%), and was even larger than the number of high quality sequences remaining with the LowRA method (2'804 reads, 46%).

3.3.3 Denoising

Denoising was performed in order to correct for classical 454 analytical errors and chimeras with the following cutoff values: a 20 minimum PHRED quality score, a 300-bp minimum sequence length, and a 500-bp maximum sequence length corresponding to theoretical amplicons of ca. 500 bp (*E. coli* standard) produced by PCR. The denoising process generated a subset of representative sequences with at least 3% dissimilarity amounting to 17±1% and 43±9% of the number of reads present in the raw datasets obtained with the HighRA and LowRA methods, respectively. The figures presented in *Additional file 2* were generated in order to determine if the minimum sequence length cutoff set of could have impacted on bacterial community profiles. Discarding sequences shorter than 300 bp did not lead to a reduced amount of detected species. Bacterial community profiles generated without and with minimum sequence length cutoff were exhibiting determination coefficients R^2 between 0.80 and 0.99 depending on the sample type and on the reference mapping database used. A reduced amount of bacterial relatives were significantly affected by the sequence length cutoff. For instance, the highest difference in relative abundances (18%) exhibited by the genus *Geobacter* was related to a high proportion of short reads below 200 bp in the dataset GRW01.

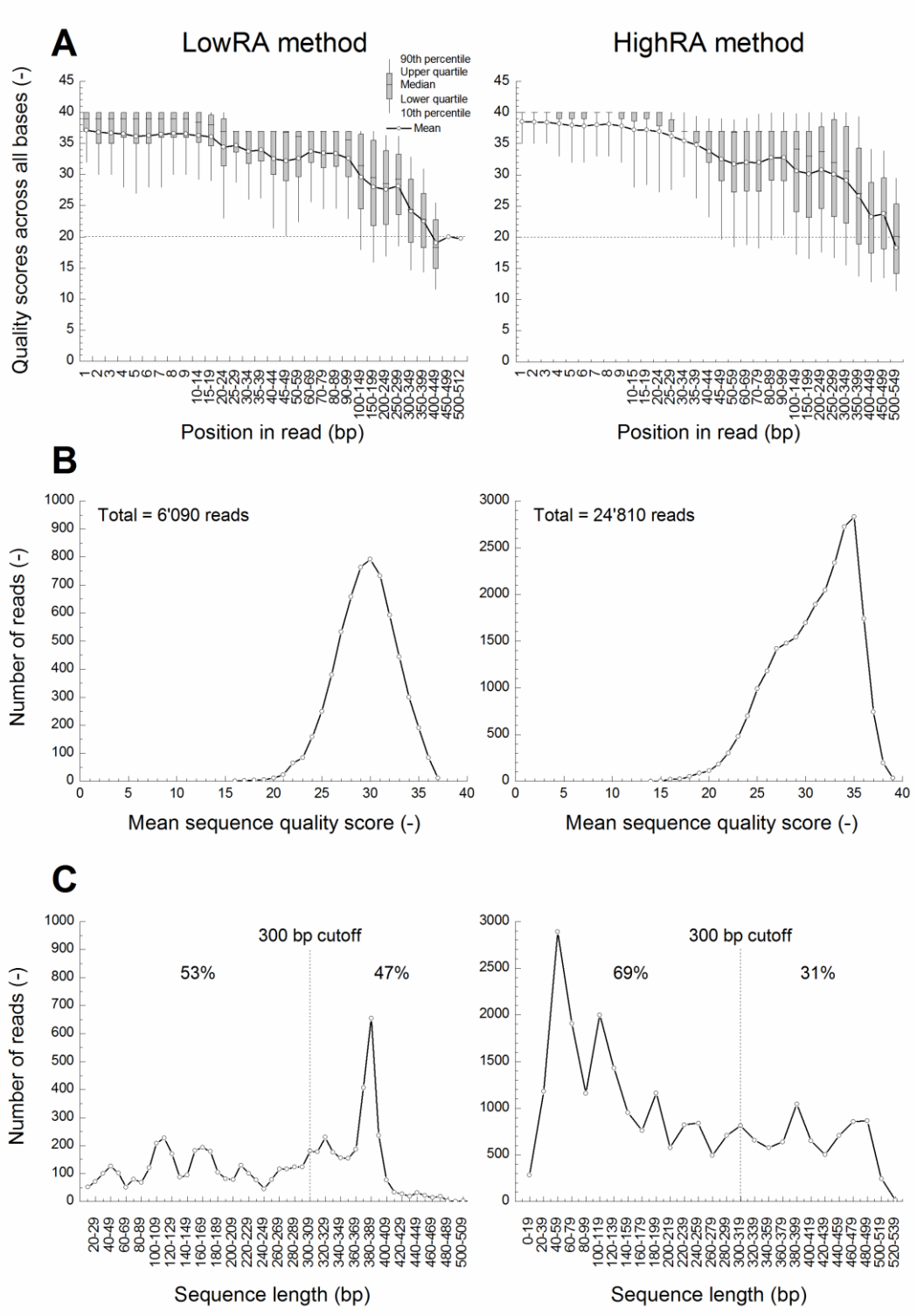


Figure 2. Quality plots generated for samples pyrosequenced with the LowRA (>3'000 expected reads) and HighRA methods (>10'000 expected reads). Sequence quality PHRED scores over all bases (**A**): PHRED scores are defined as the logarithm of the base-calling error probability $P_{error} = 10^{-PHRED}$ and PHRED = -10 log P_{error}. Box plots represent the distribution of reads quality at each sequence length. The black curve represents the mean sequence quality in function of the sequence length. Distribution of the mean sequence quality PHRED score over the pyrosequencing reads (**B**). Distribution of sequence lengths over all pyrosequencing reads (**C**). Only sequences between 300 and 500 bp were kept for digital T-RFLP analysis.

3.3.4 Mapping

The mapping process was the time-limiting step in the PyroTRF-ID pipeline. Twenty minutes were required for mapping the largest datasets against the Greengenes database. After PHRED-filtering, the remaining raw sequences had maximum lengths around 400-450 bp. Maximal SW mapping scores therefore amounted to around 400-450. Normalized SW-scores to the read lengths were comprised between 0.63 and 1.00 for predominant affiliations (Tables 2 and 3).

3.3.5 Digital T-RFLP

dT-RFLP profiles were successfully generated by PyroTRF-ID from extensive pyrosequencing datasets. Figure 3 presents full standard dT-RFLP fingerprints over 500 bp obtained with PyroTRF-ID from denoised bacterial pyrosequencing datasets representing the bacterial community structures present in contaminated groundwater (GRW01), in flocculent activated sludge (FLS01) and in aerobic granular sludge (AGS01). Results are shown for *in silico* enzymatic restriction with *Hae*III and *Msp*I endonucleases. Considering the number of peaks present in the fingerprints, the dT-RFLP profile of the acetate-fed granular sludge community exhibited lower complexity than the ones obtained with groundwater and flocculent activated sludge. With *Hae*III, 70±19 and 59±11 T-RFs were present in the dT-RFLP profiles of groundwater samples GRW01-GRW06 (HighRA method) and GRW07-GRW10 (LowRA method), 78 in the dT-RFLP profile of flocculent activated sludge (FLS01), and 42±6 in the dT-RFLP profile of granular sludge (AGS01-AGS07), respectively. For all samples, only a reduced amount of dT-RFs above 400 bp was obtained because of the low pyrosequencing quality at sequence lengths between 400-500 bp. *In silico* restriction with the *Msp*I endonuclease resulted in more dispersed profiles.

Table 2. Phylogenetic annotation of identified T-RFs

eTRF[a] (bp)	dTRF[a] (bp)	dTRF shifted[b] (bp)	Counts[c] (-)	Relative contribution to T-RF[d] (%)	Phylogenetic affiliation[e]	Reference OTU[f]	Reference accession number	SW mapping score[g] (-)	Normalized SW mapping score[h] (-)
Aerobic granular sludge biofilms from wastewater treatment reactors									
n.a.	39	34	550	70.6	F: *Xanthomonadaceae*	4015	GQ396926	386	0.960
			(276)	(35.0)	(G: *Thermomonas*)	(4045)	(EU834762)	(452)	(0.983)
			(128)	(16.0)	(G: *Pseudoxanthomonas*)	(4035)	(EU834761)	(385)	(0.955)
			112	14.3	O: *Flavobacteriales*	1151	AY468464	434	1.000
			46	5.9	F: *Rhodobacteraceae*	2718	AY212706	448	1.000
			37	4.8	S: *Rhodocyclus tenuis*	3160	AB200295	363	0.917
			18	2.3	O: *Sphingobacteriales*	1229	GU454872	394	0.990
			5	0.6	C: *Gammaproteobacteria*	3370	AY098896	403	0.906
			4	0.5	O: *Rhizobiales*	2549	EU429497	360	0.981
			4	0.5	O: *Myxococcales*	3246	DQ228369	302	0.765
			1	0.1	O: *Bacteroidales*	991	EU104248	180	0.636
194	198	193	10	90.9	G: *Acidovorax*	3011	AJ864847	384	1.000
			1	9.1	F: *Xanthomonadaceae*	4035	EF027004	303	0.819
214	219	214	769	99.6	S: *Rhodocyclus tenuis*	3160	AB200295	371	0.949
			1	0.1	G: *Methyloversatilis*	3158	DQ066958	368	0.958
			1	0.1	G: *Dechloromonas*	3156	DQ413103	321	0.988
			1	0.1	G: *Nitrosomonas*	3136	EU937892	278	0.753
220	225	220	50	92.6	O: *Rhizobiales*	2580	NR025302	448	1.000
			(31)	(57.0)	(G: *Aminobacter*)				
			2	3.7	S: *Rhodocyclus tenuis*	3160	AB200295	206	0.703
			1	1.9	F: *Hyphomonadaceae*	2656	AF236001	229	0.636
			1	1.9	P: *Firmicutes*	2235	DQ413080	284	1.000
216	221	216	10	34.5	S: *Rhodocyclus tenuis*	3160	AF502230	296	0.773
			8	27.6	G: *Nitrosomonas*	3136	GU183579	364	0.948
			6	20.7	C: *Anaerolineae*	1317	EU104216	202	0.598
			3	10.3	G: *Methyloversatilis*	3158	CU922545	360	0.909
			1	3.4	G: *Aminobacter*	2580	L20802	281	0.829
			1	3.4	G: *Dechloromonas*	3156	DQ413103	273	0.898
223	228	223	44	72.1	F: *Intrasporangiaceae* (G: *Tetrasphaera*)	418	AF255629	373	0.961
			15	24.6	F: *Hyphomonadaceae*	2656	AF236001	298	0.674
			1	1.6	F: *Microbacteriaceae*	441	GQ009478	228	0.544
			1	1.6	O: *Acidimicrobiales*	268	GQ009478	153	0.447
239	243	238	275	98.9	C: *Gammaproteobacteria*	3370	EU529737	446	0.982
			2	0.7	G: *Leptospira*	4092	AB476706	350	0.926
			1	0.4	P: *Armatimonadetes*	975	EU332819	275	0.846
249	253	249	9	100.0	S: *Rhodocyclus tenuis*	3160	AB200295	228	0.752
255	258	253	7	100.0	O: *Sphingobacteriales*	1171	FJ793188	355	0.989
260	263	258	16	94.1	G: *Nitrospira*	2360	GQ487996	389	0.982
			1	5.9	O: *Sphingobacteriales*	1171	FJ536916	251	0.640
260	264	259	38	97.4	O: *Sphingobacteriales*	1170	EU104185	267	0.706
			1	2.6	G: *Nitrospira*	2360	GQ487996	319	0.788
297	302	297	26	100.0	G: *Herpetosiphon*	1359	NC009972	339	0.867
307	311	306	38	97.4	P: *Armatimonadetes*	975	CU921283	218	0.472
			1	2.6	O: *Sphingobacteriales*	1171	EU104210	196	0.525
321	323	318	17	100.0	G: *Cytophaga*	1208	EU104191	367	0.968
393	397	392	33	100.0	G: *Bdellovibrio*	3173	CU466777	262	0.663

Table 2 – *Continued*

eTRF[a] (bp)	dTRF[a] (bp)	dTRF shifted[b] (bp)	Counts[c] (-)	Relative contribution to T-RF[d] (%)	Phylogeny[e]	Reference OTU[f]	Reference accession number	SW mapping score[g] (-)	Normalized SW mapping score[h] (-)
Groundwater samples from chloroethene-contaminated aquifers									
63	69	64	93	85.3	F: *Methylococcaceae*	3686	AB354618	432	0.915
			14	12.8	F: *Crenotrichaceae*	3681	GU454947	290	0.816
			1	0.9	F: *Ectothiorhodospiraceae*	3510	AM902494	168	0.542
			1	0.9	P: OP3	2388	GQ356152	187	0.488
165	168	163	143	100.0	G: *Dehalococcoides*	1368	EF059529	448	0.953
190	193	191	12	54.6	F: *Desulfobulbaceae*	3177	AJ389624	379	0.945
			4	13.6	F: *Sphingomonadaceae*	2880	AY785128	263	0.555
			2	9.1	F: *Erythrobacteraceae*	2872	DQ811848	343	0.771
			2	9.1	C: *Alphaproteobacteria*	2451	AY921822	337	0.926
			1	4.6	F: *Rhodospirillaceae*	2793	AY625147	294	0.679
			1	4.6	F: *Rhodobiaceae*	2641	AB374390	328	0.877
198	201	196	140	98.6	G: *Desulfovibrio*	3215	FJ810587	473	1.000
			2	1.4	F: *Comamonadaceae*	3039	FN428768	311	0.814
210	214	209	233	98.3	F: *Dehalococcoidaceae*	1367	EU679418	262	0.665
			2	0.8	O: *Burkhorderiales*	3009	AM777991	367	0.927
			1	0.4	F: *Spirochaetaceae*	4130	EU073764	295	0.848
			1	0.4	P: TM7	4379	DQ404736	277	0.723
216	221	216	1010	90.9	F: *Gallionellaceae*	3080	EU802012	353	0.869
			94	8.5	G: *Rhodoferax*	3050	DQ628925	369	0.920
			3	0.3	G: *Methylotenera*	3093	AY212692	291	0.744
			1	0.1	G: *Methyloversatilis*	3158	GQ340363	296	0.765
			1	0.1	F: *Clostridiaceae*	2005	AJ863357	338	0.833
			1	0.1	C: *Anaerolineae*	1315	AB179693	229	0.511
			1	0.1	C: *Actinobacteria*	949	EU644115	372	0.907
243	247	243	389	99.7	F: *Dehalococcoidaceae*	1367	EU679418	255	0.631
			1	0.3	F: *Anaerolinaceae*	1321	AB447642	253	0.806

[a] Experimental (eT-RF) and digital T-RFs (dT-RF).
[b] Digital T-RF obtained after having shifted the digital dataset with the most probable average cross-correlation lag.
[c] Number of reads of the target phylotype that contribute to the T-RF.
[d] Diverse bacterial affiliates can contribute to the same T-RF.
[e] Phylogenetic affiliation of the T-RF (K: kingdom, P: phylum, C: class, O: order, F: family, G: genus, S: species). Only the last identified phylogenetic branch is presented here.
[f] Reference operational taxonomic unit (OTU) from the Greengenes public database related with the best SW mapping score. In the Greengenes taxonomy, OTU refer to terminal levels at which sequences are classified.
[g] Best SW mapping score obtained. SW scores consider nucleotide positions and gaps. The highest SW mapping score that can be obtained for a read is the length of the read itself.
[h] SW mapping score normalized by the read length, as an estimation of the percentage of identity.
[i] After having observed the presence of the dT-RF 34 bp, we returned to the raw eT-RFLP data and found an important eT-RF at 32 bp. However, Rossi et al. (*35*) considered that T-RFs below 50 bp are inconsistent and lacks of precision in sizing. This peak was therefore primarily not taken into account in the original eT-RFLP profiles.

Table 3. T-RF diversity for single phylogenetic descriptions

Phylogenetic[a] affiliation	dTRF (bp)	dTRF shifted[b] (bp)	Counts[c] (-)	Relative contribution to T-RF[d] (%)	Reference OTU[e]	Reference accession number	SW mapping score[f] (-)	Normalized SW mapping score[g] (-)
Flocculent and aerobic granular sludge samples from wastewater treatment systems								
G: *Rhodocyclus tenuis*	39	34	37	4.8	3160	AB200295	363	0.917
	199	194	1	25.0	3160	AB200295	248	0.648
	205	200	3	100.0	3160	AF204247	314	0.858
	210	205	1	100.0	3160	AF204247	211	0.699
	218	213	11	91.7	3160	AB200295	356	0.942
	219	**214**	**769**	**99.6**	**3160**	**AB200295**	**371**	**0.949**
	220	215	6	37.5	3160	AF502230	318	0.817
	221	216	1	7.7	3160	AF502230	276	0.865
	225	220	2	3.7	3160	AB200295	206	0.703
	252	247	3	100.0	3160	AB200295	305	0.762
	253	248	9	100.0	3160	AB200295	228	0.752
	257	252	1	20.0	3160	AF502230	241	0.660
Groundwater samples from aquifers contaminated with chloroethenes								
G: *Dehalococcoides*	166	161	1	100.0	1368	EF059529	290	0.775
	168	**163**	**143**	**100.0**	**1368**	**EF059529**	**241**	**0.717**
	169	164	2	100.0	1368	EF059529	331	0.768
	170	165	2	100.0	1368	EF059529	241	0.693
	171	166	1	50.0	1368	EF059529	303	0.783
	173	168	1	100.0	1368	EF059529	241	0.717
	176	171	1	100.0	1369	DQ833317	211	0.687
	179	174	1	100.0	1369	DQ833317	193	0.629
	188	183	4	66.7	1369	DQ833340	464	0.947

[a] Phylogenetic affiliation of the T-RF (K: kingdom, P: phylum, C: class, O: order, F: family, G: genus, S: species). Only the last identified phylogenetic branch is presented here.
[b] Digital T-RF obtained after having shifted the digital dataset with the most probable average cross-correlation lag.
[c] Number of reads of the target phylotype that contribute to the T-RF.
[d] Diverse bacterial affiliates can contribute to the same T-RF.
[e] Reference OTU from the Greengenes public database obtained after mapping.
[f] Best SW mapping score obtained.
[g] SW mapping score normalized by the read length.

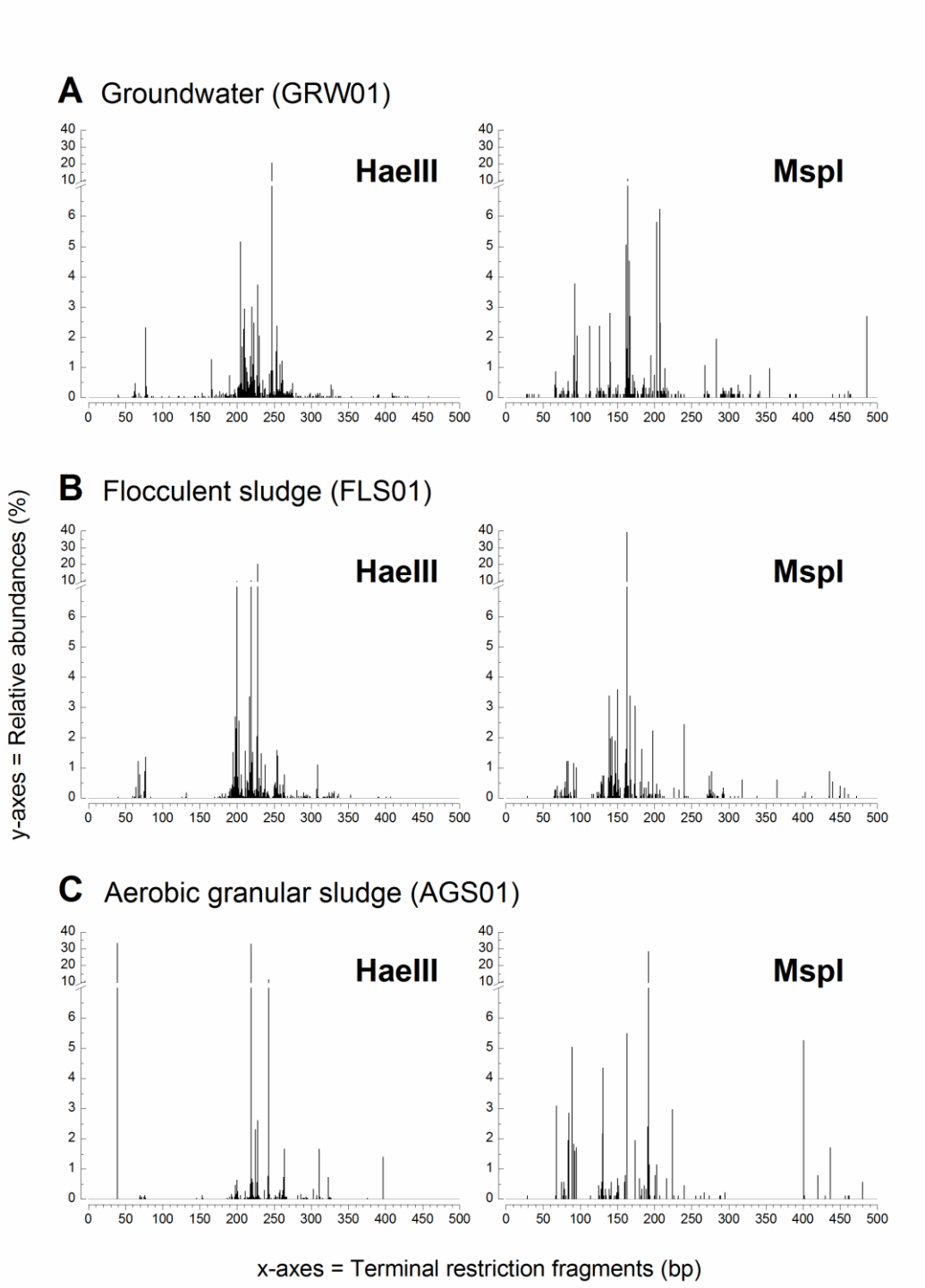


Figure 3. Examples of full digital T-RFLP profiles obtained with the restriction enzymes *Hae*III and *Msp*I for contaminated groundwater (GRW01), flocculent activated sludge (FLS01) and aerobic granular sludge (AGS01) samples.

3.3.6 Mirror plots and cross-correlations between digital and experimental T-RFLP profiles

Figure 4 presents one example of a mirror plot generated by PyroTRF-ID computed with raw and denoised pyrosequencing datasets obtained for a complex groundwater bacterial community (GRW01). Further examples of mirror plots are available in *Additional file 3* for the different sample types. dT-RFLP profiles generated from raw pyrosequencing datasets were highly affected by the background noise generated by the 454 pyrosequencing analysis (Figure 4a). The digital profiles were affected by Gaussian dispersions of neighbor peaks around the most dominant T-RFs, and exhibiting identical bacterial affiliations (data not shown). Denoised dT-RFLP profiles displayed enhanced relative abundances of dominant peaks and had a higher cross-correlation with eT-RFLP profiles (Figure 4b). From selection of representative sequences for clusters of at least 97% similar reads, all neighbor peaks were thus integrated in the dominant centroid T-RFs. For sample GRW01, the cross-correlation between dT-RFLP and eT-RFLP profiles increased from 0.43 to 0.62 considering raw or denoised pyrosequencing data, respectively.

The dT-RFLP profiles exhibited an average shift of 4 to 6 bp compared to eT-RFLP profiles. From *in silico* restriction of a set of 150 clones from an internal cloning-sequencing databank using the TRiFLe software (*48*), we confirmed that *in silico* T-RFs were on average 4±1 bp (min 3 bp – max 6 bp) longer than the experimental ones (data not shown). After matching at optimal lag (*Additional file 4*) maximum cross-correlation coefficients between denoised dT-RFLP and eT-RFLP profiles ranged from 0.55±0.14 and 0.67±0.05 for the groundwater samples (HighRA and LowRA method, respectively) to 0.82±0.10 for the flocculent and granular sludge samples (LowRA method).

3.3.7 Impact of sequence processing steps, pyrosequencing methods, and sample types

Richness and diversity indices were used to express the impacts of data processing steps, pyrosequencing methods and sample types on the final dT-RFLP profiles (Figure 5). Improved matching with reference eT-RFLP profiles occurred when the richness and diversity equivalents of dT-RFLP profiles approached the characteristics of eT-RFLP profiles.

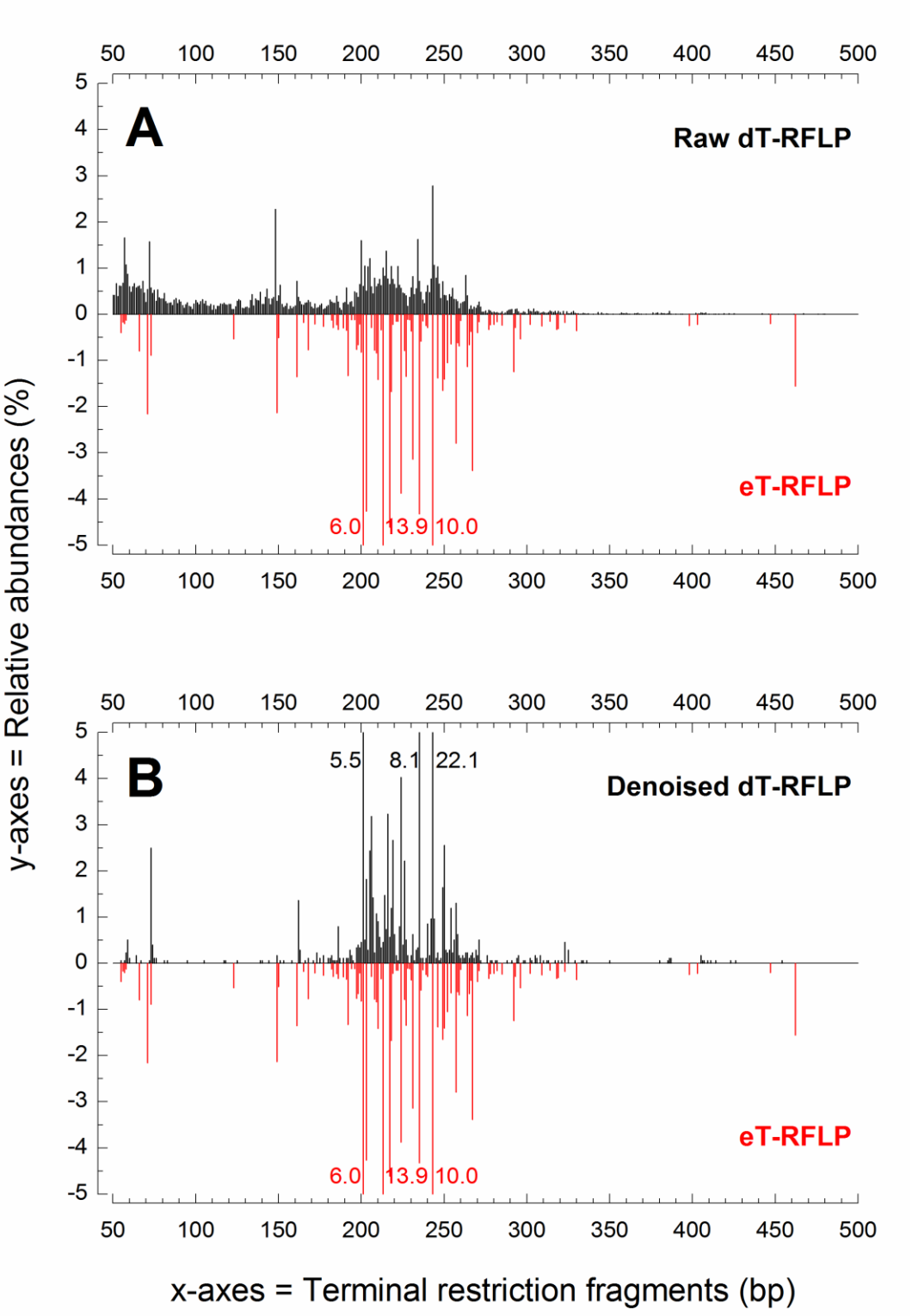


Figure 4. Mirror plot displaying the cross-correlation between digital and experimental T-RFLP profiles. This mirror plot was generated for a complex bacterial community of a groundwater environment contaminated with chloroethenes (GRW01). Comparison of mirror plots constructed with raw (**A**) and denoised sequences (**B**). Relative abundances are displayed up to 5% absolute values. For those T-RFs exceeding these limits, the actual relative abundance is displayed beside the peak.

The noisy raw dT-RFLP profiles were composed of 2.4 to 7.4 times more T-RFs than the eT-RFLP profiles, and exhibited on average 1.3 times higher diversity equivalents, depending on the pyrosequencing method and on the sample type. Filtering the reads with a SW mapping score cutoff of 150 did not significantly help reducing the noise in the dT-RFLP profiles. The richness and diversity equivalents were only reduced by a factor of 1.1. Denoising was successful for suppressing the background noise, and for optimizing the cross-correlation between dT-RFLP and eT-RFLP profiles. In comparison to raw profiles, denoising resulted in an improvement of richness and diversity equivalents by factors of 1.3 to 2.6 and 1.5, respectively.

The dT-RFLP profiles of the groundwater samples GRW01-GRW06 which were sequenced with the HighRA method were composed of 4 to 7.4 times more T-RFs than their eT-RFLP correspondents. For the second set of groundwater samples GRW07-GRW10 sequenced with the LowRA method, the ratios of dT-RFs to eT-RFs were between 2.4 and 5.2. After denoising, both sets of groundwater-related dT-RFLP profiles exhibited similar richness, and similar richness ratios with eT-RFLP data.

In addition, eT-RFLP profiles of the complex groundwater and flocculent activated sludge bacterial communities displayed 1.6 and 1.3 times higher richness and diversity than the acetate-fed granular sludge communities. The denoised dT-RFLP profiles were around two times richer and more diverse for the high-complexity samples (factors of 1.9 and 1.8, respectively). After denoising, the ratios of richness and of diversity between dT-RFLP and eT-RFLP profiles were close for all types of high-complexity (2.5±0.6 and 1.0±0.3, respectively) and low-complexity samples (2.1±0.5 and 0.8±0.2, respectively).

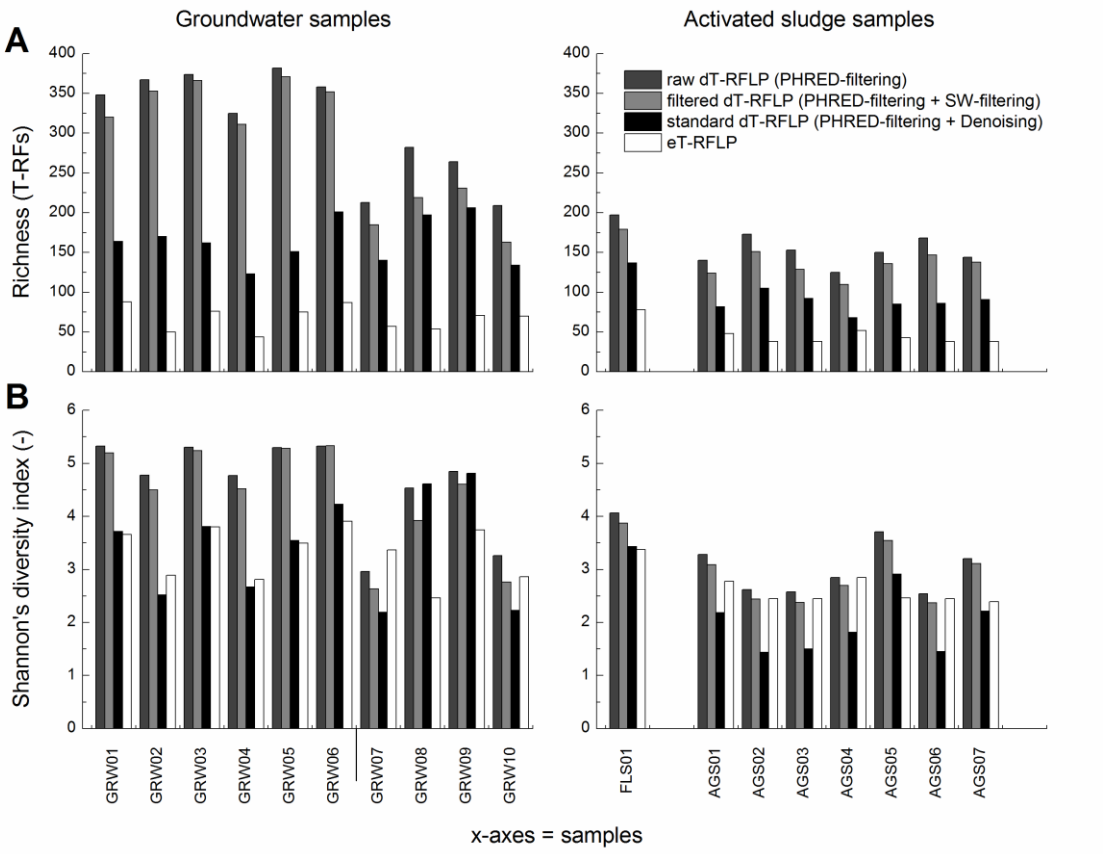


Figure 5. Comparison of the richness in number of T-RFs **(A)** and of the Shannon's H' diversity index **(B)** obtained for fingerprints related to groundwater samples pyrosequenced with the HighRA (GRW01-GRW06) and LowRA methods (GRW07-GRW10), to flocculent activated sludge (FLS01) and to aerobic granular sludge samples (AGS01-AGS07) pyrosequenced with the LowRA method.

3.3.8 Peak annotation and phylogenetic affiliation of T-RFs

Peak annotation provided comprehensive information for each T-RF, as exemplified by Table 2. Depending on the sample type, between 45 and 60% of all dT-RFs contained in dT-RFLP profiles were affiliated each with a single bacterial population (Figure 6). The other dT-RFs were composed of several contributing phylotypes. However in such cases one phylotype was displaying clear predominance between 50 and 99% of all phylotypes contributing to the T-RF. For some T-RFs no clear dominant phylotype emerged (e.g. eT-RF 216 in granular sludge samples, Table 2) and different reference sequences sometimes affiliated to several T-RFs (Table 3). For instance, in samples of aerobic granular sludge performing dephosphatation such as AGS01, six different dT-RFs (34, 194, 213, 214, 220, 247 bp) were affiliated to the same reference sequence of *Rhodocyclus tenuis* (accession number AB200295), with T-RF 214 being predominant with 769 affiliated pyrosequencing reads. The *Dehalococcoides* sp. affiliation in the chloroethene-contaminated groundwater sample GRW05 was related to

eight different T-RFs, with T-RF 163 being clearly dominant (143 affiliated reads). To elucidate the reasons of this phenomenon, the reads mapping with the same reference and resulting in different T-RF sizes were retrieved from the pyrosequencing database and aligned in ClustalX (*Additional file 5*). The analysis showed that the discrepancies observed in the T-RF sizes were due to insertions or deletions of nucleotides before the restriction site.

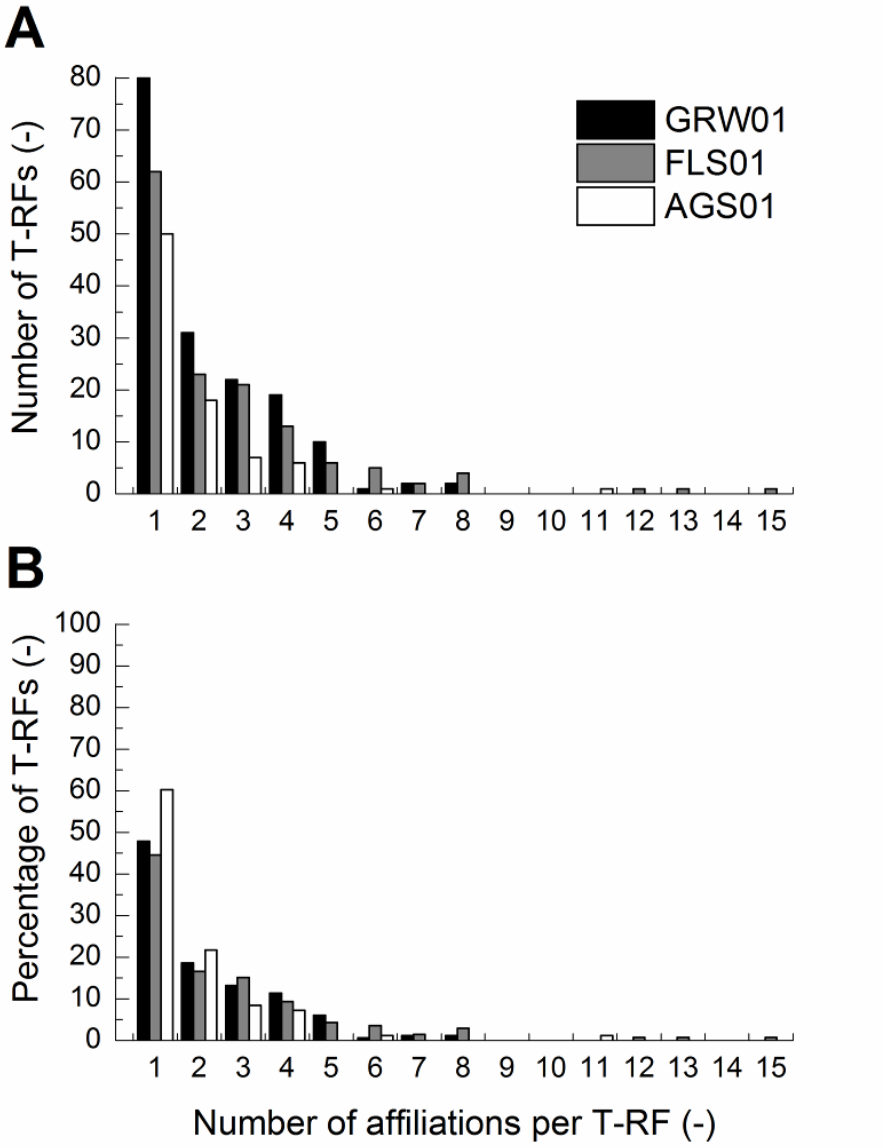


Figure 6. Amount of bacterial affiliations contributing to T-RFs. Number of T-RFs **(A)** and percentage of all T-RFs **(B)** composed of one to several bacterial contributions. Groundwater sample GRW01 (black bars), flocculent activated sludge sample FLS01 (gray bars) and aerobic granular sludge sample AGS01 (white bars).

3.3.9 Screening of restriction enzymes

Over the six endonucleases tested, the average richness of the groundwater dT-RFLP profiles (129 dT-RFs) was on average 1.2- and 2.1-times higher than the richness of flocculent (104 dT-RFs) and granular (61 dT-RFs) sludge dT-RFLP profiles, respectively (Figure 7). The decrease in richness from groundwater to flocculent and granular sludge samples was independent of the endonuclease (Figures 7 and 8).

*Hae*III provided dT-RFLP profiles with the highest richness from the tested groundwater sample. dT-RFLP profiles exhibiting highest diversities were obtained with *Msp*I. The results of the dT-RFLP profiles constructed with the different enzymes diverged between the flocculent and granular sludges. For the flocculent sludge, the highest richness was obtained with *Alu*I and the highest diversity with *Hae*III. In contrast, the highest richness in granular sludge was obtained with *Hae*III and the highest diversity with *Rsa*I. For all sample types, *Taq*I provided both the lowest richness and diversity.

According to the density plots presented in Figure 8, each restriction enzyme generated characteristic dT-RFLP features independent of the sample type. With *Hae*III, dT-RFs were stacked between 200 and 300 bp. Digestion with *Msp*I resulted in more homogeneous distributions of dT-RFs than *Hae*III up to ca. 300 bp as reflected by the linear tendency of the curve of cumulated numbers of dT-RFs (Figures 3 and 8). The second half of the profile did not contain numerous dT-RFs, and the cumulative curve thus became flat. This was also the case for the other restriction nucleases, except *Hha*I. The curve of the cumulative numbers of dT-RFs obtained with this endonuclease increased step-wise. *Rsa*I provided a homogeneous distribution of dT-RFs for the groundwater sample (Figure 8a). The cumulated curve was almost linear, indicating that the dT-RFs appeared at regular intervals. However, the profile obtained with this enzyme displayed a lower diversity than *Hae*III, *Alu*I, *Msp*I, and *Hha*I.

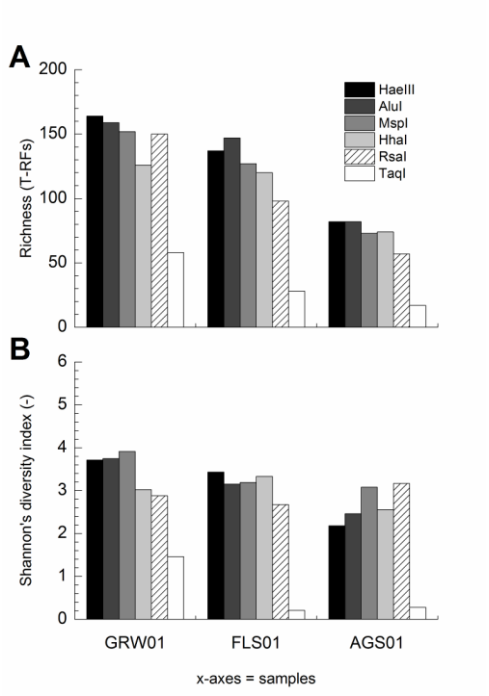


Figure 7. Richness and diversity indices of dT-RFLP profiles generated with different restriction enzymes. Graphs showing the richness in number of T-RFs **(A)** and Shannon's H' diversity index **(B)** calculated on digital T-RFLP profiles obtained with the standard PyroTRF-ID procedure involving denoising, for groundwater (GRW01), flocculent activated sludge (FLS01) and aerobic granular sludge (AGS01) with different restriction endonucleases (*Hae*III, *Alu*I, *Msp*I, *Hha*I, *Rsa*I and *Taq*I).

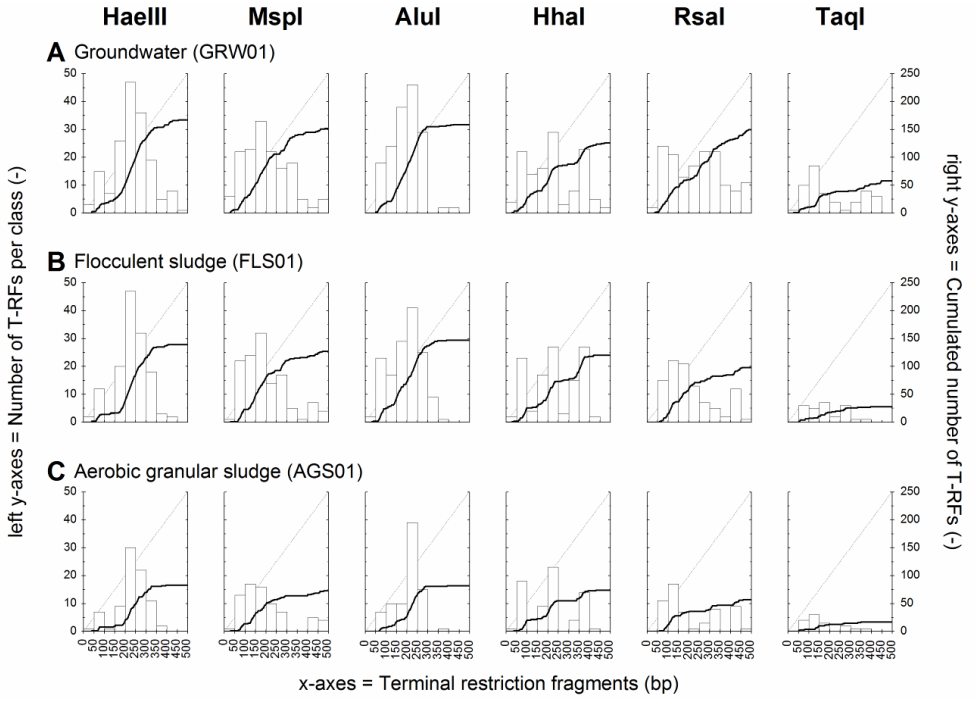


Figure 8. Density plots displaying the repartition of T-RFs along the 0-500 bp domain obtained with different endonucleases. The effect of the different restriction endonucleases *Hae*III, *Alu*I, *Msp*I, *Hha*I, *Rsa*I and *Taq*I was tested on pyrosequencing datasets from groundwater sample GRW01 **(A)**, flocculent activated sludge sample FLS01 **(B)** and aerobic granular sludge sample AGS01 **(C)**. Histograms represent the number of T-RFs produced per class of 50 bp (to read on the left y-axes). Thick black lines represent the cumulated number of T-RFs over the 500-bp fingerprints (to read on the right y-axes).

3.4 Discussion

The PyroTRF-ID software was developed as a high-througput bioinformatics procedure for the anticipation of T-RFLP profiles from microbiome pyrosequencing datasets and for the identification of eT-RFs. It combines the high resolution power of pyrosequencing and the convenience of T-RFLP for routine analysis. This chapter aimed at describing the development, optimization, performance and applicability of PyroTRF-ID for the analysis of microbial communities from various low- to high-complexity environments.

3.4.1 Setup of the standard PyroTRF-ID procedure

454 pyrosequencing errors, such as inaccurate resolving of homopolymers and single base insertions (*49*) can lead to overestimates of the diversity (*50, 51*). These errors were expected to impact the quality of the dT-RFLP profiles generated from the pyrosequencing datasets, leading to erroneous interpretation. Quality filters were therefore included into the PyroTRF-ID software. The PHRED quality score threshold of 20 was used for preliminary removal of low quality sequences, but had only limited impact on the noise present in the datasets by removing only less than 1% of the reads. Two additional procedures were implemented to remove the remaining noise. First, a SW-cutoff arbitrarily set at 150 was introduced to remove sequences with a low SW mapping score. This filtering only slightly reduced both richness and diversity of the dT-RFLP profiles. Furthermore, filtering the reads using the mapping score implies that the researcher assumes that sequences with a low mapping score are likely to be of poor quality (*49, 52*). This approach does not specifically remove reads based on their intrinsic quality, but rather on similarities with existing sequences from the database, hence reducing the complexity of the studied bacterial community to what is already known. As exemplified by *Additional file 2*, the reference mapping database can in addition have an impact on the final identification of T-RFs. Between 35 and 45% of the reads were unassigned during mapping in MG-RAST with the Greengenes database, while only 3-5% were unassigned with RDP with identical mapping criteria. This aspect stresses the need of standardized databases and microbiome dataset processing approaches within the microbial ecology discipline.

Since the proposed method intended to preserve the complexity of the bacterial community, we opted for a denoising approach based on the analysis of rank-abundance distributions (*43*), and implemented QIIME denoising libraries into PyroTRF-ID. As expected, denoising helped to strongly reduce the overestimated number of pyrosequencing reads, to filter virtual T-RFs out of the dT-RFLP profiles, and to generate dT-RFLP profiles approaching the structure of eT-RFLP profiles. The observation that for some samples the denoised dT-RFLP profiles exhibited higher richness equivalents but lower diversity than the eT-RFLP profiles is explained by the integration of both richness and evenness – i.e. distribution of abundances among T-RFs (*53*) – indicators in the Shannon's H' diversity index. The observed feature indicated that these dT-RFLP profiles contained more T-RFs (high richness equivalents) but with only few highly abundant ones and the others in very low proportion (low evenness). Even though the denoising process modified the original pyrosequencing dataset, in that only

centroids of each cluster of sequences were selected and reproduced artificially, this resulted in improved cross-correlations between denoised dT-RFLP and eT-RFLP profiles, reaching correlation coefficients of up to 0.91 for granular sludge samples.

The impacts of the data processing steps were different for the samples sequenced with the HighRA (>10'000 reads) or LowRA (>3'000 reads) methods. A total of 55±7% of the dT-RFs were removed on average with the standard procedure applied on the larger pyrosequencing datasets, whereas 37±7% were removed from the smaller ones. This difference was related to the size of the initial pyrosequencing datasets, and to the proportion of small-size reads, which were more abundant in HighRA datasets. However, the pyrosequencing method and the initial amount of reads did favorably not impact the final standard PyroTRF-ID output. Similar ratios between dT-RFLP and eT-RFLP profile characteristics were obtained with the HighRA and LowRA methods. Only the level of complexity of the studied ecosystems, the natural variations in the bacterial communities in the sampled environments and the common intrinsic variations in pyrosequencing methods could have explained the differences in richness among samples.

From the six endonucleases tested, *Hae*III, *Alu*I and *Msp*I were good candidates for the generation of rich and diverse dT-RFLP profiles for all types of low- and high-complexity environments. In comparison to previous studies involving *in silico* restriction of artificial microbial communities synthesized with selected reference sequences from public or cloning-sequencing databases (*31*, *30*, *32*), PyroTRF-ID has the advantage to work on full pyrosequencing datasets. The optimization of the T-RFLP fingerprints can therefore be performed by integrating the whole microbial complexity of the studied environments. An additional module analog to REPK (*30*) is planned to be implemented in a further version of PyroTRF-ID for fast determination of optimal restriction enzymes and combinations.

3.4.2 Aspects of the PyroTRF-ID standard procedure

Over all samples, it was possible to affiliate 63±18% of all eT-RFs with corresponding dT-RFs. The phylogenetic affiliation of T-RFs with PyroTRF-ID was based on the correspondence of eT-RFLP and dT-RFLP profiles which was achieved by calculating cross-correlation coefficients for different shifts and by choosing the shift with highest coefficient. For the samples investigated, the cross-correlation between these profiles amounted on average between 0.60±0.13 for complex groundwater bacterial communities and 0.82±0.10 for low-complexity granular sludge communities. Maximum cross-correlation of 0.91 was obtained for some flocculent and granular sludge samples. These values attest the high level of correspondence between T-RFLP profiles aligned with PyroTRF-ID. Discrepancies between the size of dT-RFs and eT-RFs have previously been described by different authors. In a review, Kitts (*54*) has reported that the predicted T-RFs using sequence databases were regularly 0-7 bp longer than eT-RFs. Junier et al. (*48*) have obtained predicted T-RFs 1-6% longer than the measured T-RFs. In the present study, similar shifts were observed between *in silico* and

experimental digestion of a set of 150 clones from a clone library. Average lags of 4-6 bp between the dT-RFLP and the eT-RFLP profiles were determined for all samples.

To overcome the bias induced by the shift, we introduced the calculation of a cross-correlation between dT-RFLP and eT-RFLP profiles. The entire dT-RFLP profiles were shifted by the number of base pairs enabling optimal alignment of both profiles. Nevertheless, such discrepancy is not constant across the dT-RFs since it positively correlates with the true T-RF length, negatively correlates with the purine content of T-RFs, and depends on the secondary structure of each T-RF as well as on the dye label (*54*, *46*, *47*). Thus, mirror plots sometimes displayed major T-RFs that were not perfectly aligned. Often, a 1-bp difference was observed between the eT-RFs and dT-RFs. It was therefore crucial for the user to analyze the mirror plots prior to assigning phylotypes to eT-RFs. The approach adopted here consisted in selecting T-RFs to identify prior to checking their alignment with dT-RFs. User interpretation can introduce a subjective step in the phylogenetic assignment process, especially when analyzing a large number of peaks and samples (*54*). The most precise way to overcome manual inspection could consist in computing a lag for each single dT-RF in function of its sequence composition and theoretical secondary structure following the regression function described by Bukovska et al. (*46*). However, the standard deviation associated with this method is still higher than 1 bp, which was the most often encountered discrepancy in the mirror plots, and shifting each single dT-RF based on this function was therefore not expected to improve the alignment accuracy. According to Kitts (*54*), final manual processing of T-RFLP profiles can remain the only way to resolve T-RF alignment problems.

Correspondence between dT-RFs and eT-RFs was relatively obvious for T-RFs which were abundant in both profiles but less obvious for low abundance T-RFs. dT-RFLP profiles contained numerous low abundance peaks which were absent in eT-RFLP profiles. Conversely, eT-RFs were sometimes lacking a corresponding dT-RF. This mainly occurred in profiles generated using pyrosequencing datasets with an initially low amount of reads exceeding 400 bp, and arose from the decreasing probability of finding a restriction site in the final portion of the sequences while the number of long reads decreased. For eT-RFs near 500 bp, incomplete enzymatic restriction could explain that undigested amplicons were detected in the electrophoresis run (*55*, *56*). These features however do not explain all missing dT-RFs, which sometimes occurred in the initial portion of the T-RFLP profile. Egert and Friedrich (*57*) have attributed the presence of ‘pseudo T-RFs’ to undigested single stranded DNA amplicons. They have demonstrated that cleaving amplicons with single-strand-specific mung bean nuclease was enabling getting rid of all pseudo T-RFs.

The majority of dT-RFs were affiliated to several phylotypes, revealing the phylogenetic complexity underlying each T-RF, which was in agreement with Kitts (*54*). PyroTRF-ID enabled assessing the relative contributions of each phylotype, and determining the most abundant ones. In most cases, one phylotype displayed clearly the highest number of reads for one dT-RF. However, for some dT-RFs several phylotypes contributed almost equally to the total number of reads. Although problematic while aiming at identifying T-RFs, this information is of primary importance if PyroTRF-ID is intended to be used for designing the most adapted T-RFLP procedure

for the study of a particular bacterial community. The resolution between T-RF identities could potentially be improved with other restriction endonucleases, or optimally with a multiple enzyme restriction (*54*, *58*). This would lead to improved identification of T-RFs and routine eT-RFLP analysis of spatial and temporal shifts of target microbial populations.

3.5 Conclusion

This study resulted in the successful development of the PyroTRF-ID software for high-throughput generation of digital T-RFLP profiles from pyrosequencing datasets and for identification of microbial contributions to eT-RFs. To the best of our knowledge, it is the first time that eT-RFs were massively affiliated to sequences obtained from the same samples. In addition, the following can be concluded:

- The combination of pyrosequencing and eT-RFLP data directly obtained from the same samples was a powerful characteristic of PyroTRF-ID, enabling generation of dT-RFLP profiles that integrate the whole complexity of microbiomes of interest.
- The type of 454 pyrosequencing method and the amount of expected sequences did not impact on the final results of the PyroTRF-ID procedure.
- Denoising was a crucial step in the 454 pyrosequencing dataset processing pipeline in order to generate representative digital fingerprints.
- PyroTRF-ID could be applied to the screening of restriction enzymes for the optimization of favorably distributed eT-RFLP profiles by considering the whole underlying microbial communities. *Hae*III, *Msp*I and *Alu*I were good candidates for T-RFLP profiling with high richness and diversity indices.
- PyroTRF-ID was validated with different samples from low- and high-complexity environments, and could therefore be implemented in a broad spectrum of environmental to medical applications with optimized laboratory and computational costs.

Acknowledgements

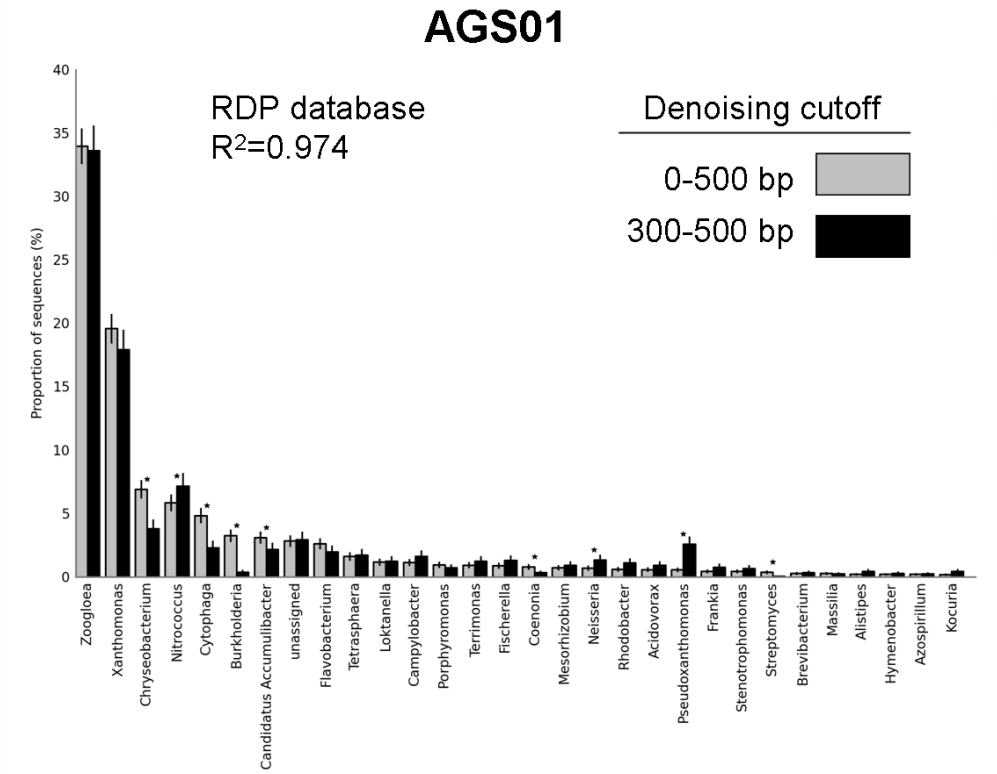
We recognize the excellent assistance of Yoan Rappaz in molecular biology analyses. We acknowledge Scot E. Dowd, Yan Sun, Lars Koenig and the Research and Testing Laboratory (Lubbock, Texas, USA), Timothy M. Vogel, Sébastien Cecillon and Tom O. Delmont and the Environmental Microbial Genomics Group at Ecole Centrale de Lyon (France), and GATC Biotech (Konstanz, Germany) for pyrosequencing analyses and advice. We are grateful to Ioannis Xenarios for support and access to the Vital-IT high performance computing cluster of the Swiss Institute for Bioinformatics (Lausanne, Switzerland).

Additional files

Additional file 1. Central procedure of the PyroTRF-ID software encoded in python (This procedure was encoded inside the module: `__init__.py`).

```
#-------------------------------------------------------------------------#

def digest(self, enzyme, primer_length, sw_threshold):
        """Virtually digest all the reads in the bam file.
        The documentation for the ``AlignedRead`` object is here:
        http://www.cgat.org/~andreas/documentation/pysam/api.html

        :param enzyme: The enzyme name.
        :type  enzyme: str
        :param primer_length: The amount of base pairs to cut off from every read.
        :type  primer_length: int
        :param sw_threshold: Ignore reads that mapped with a Smith-Waterman score below
this.
        :type  sw_threshold: int
        """

# Get the enzyme from the biopython library #
        try: self.enzyme = getattr(Bio.Restriction, enzyme)
        except AttributeError: raise Exception("Enzyme defined as '%s' is not present in the
biopython library." % enzyme)

        # Prepare the peak objects #
        self.peaks = dict([(l,Peak(l)) for l in range(self.max_frag_size+1)])

        # Open the bam file with the pysam tool #
        short_reads = pysam.Samfile(self.bam_path, "rb")

        # Do something for every read #
        for read in short_reads:
            # Skip the ones that didn't map #
            unmapped = True if read.flag != 0 and read.flag != 16 else False
            if unmapped: continue
            # Skip the ones that have a SW score too low #
            sw_score = read.tags[0][1]
            if sw_score < sw_threshold: continue
            # Cut the short read sequence in pieces #
            fragments = self.enzyme.search(Bio.Seq.Seq(read.seq))
            # Skip the ones that don't have a restriction site #
            if not fragments: continue
            # Only the first restriction site is considered #
            fragment_cut_location = fragments[0]
            # Cutting at position 6 will give a fragement of size 5 #
            fragment_length = fragment_cut_location - 1
            # Adjust the size to include the primer that was removed earlier #
            fragment_length += primer_length
            # Ignore those that are too long #
            if fragment_length > self.max_frag_size: continue
            # Retrieve the taxonomy #
            bacteria = Bacteria(short_reads.getrname(read.tid))
            # Create the fragment object #
            fragment = Fragment(fragment_length, bacteria, read.qname, read.rlen, sw_score)
            # Add the fragment to the right peak #
            self.peaks[fragment_length].add_fragment(fragment)

#-------------------------------------------------------------------------#
```

Additional file 2. Assessment of mapping performances with pyrosequencing datasets denoised without (0-500 bp) and with (300-500 bp) minimal read length cutoff. Examples are given for the groundwater sample GRW01, the flocculent activated sludge sample FLS01 and the aerobic granular sludge sample AGS01. After denoising with the one or the other method, each dataset was mapped against a reference database with MG-RAST (Meyer et al. 2008). No cutoff was set for e-value, minimum identity and minimum alignment length. After having observed that between 35-45% of the sequences were unassigned with Greengenes, RDP – the Ribosomal Database Project (*59*) (Cole et al. 2009) was used as reference database for this assessment (only 4% unassigned sequences). Correlations between bacterial community profiles obtained with both denoising methods and both reference databases were analyzed with STAMP (*60*) (Parks and Beiko 2010).



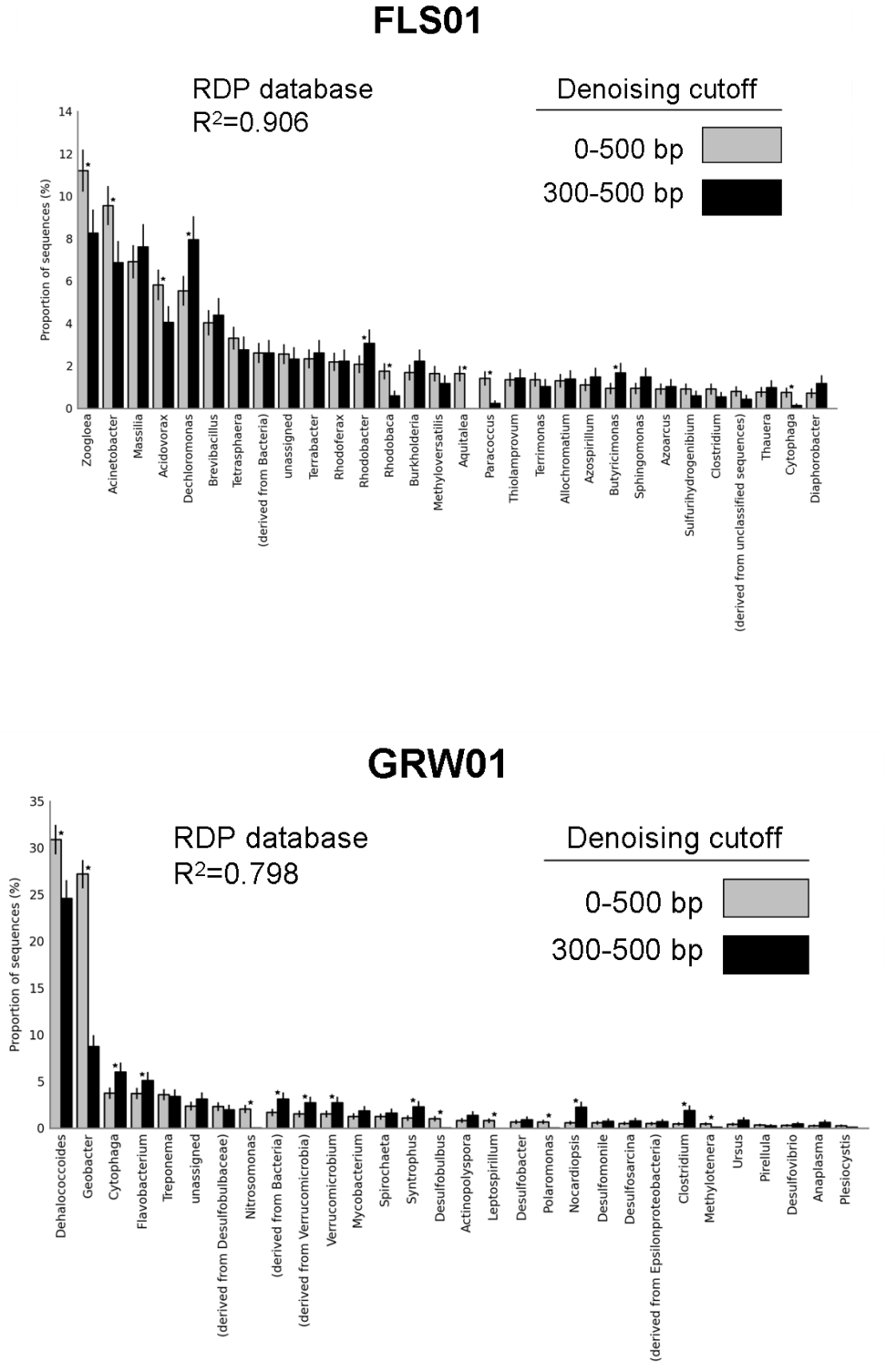
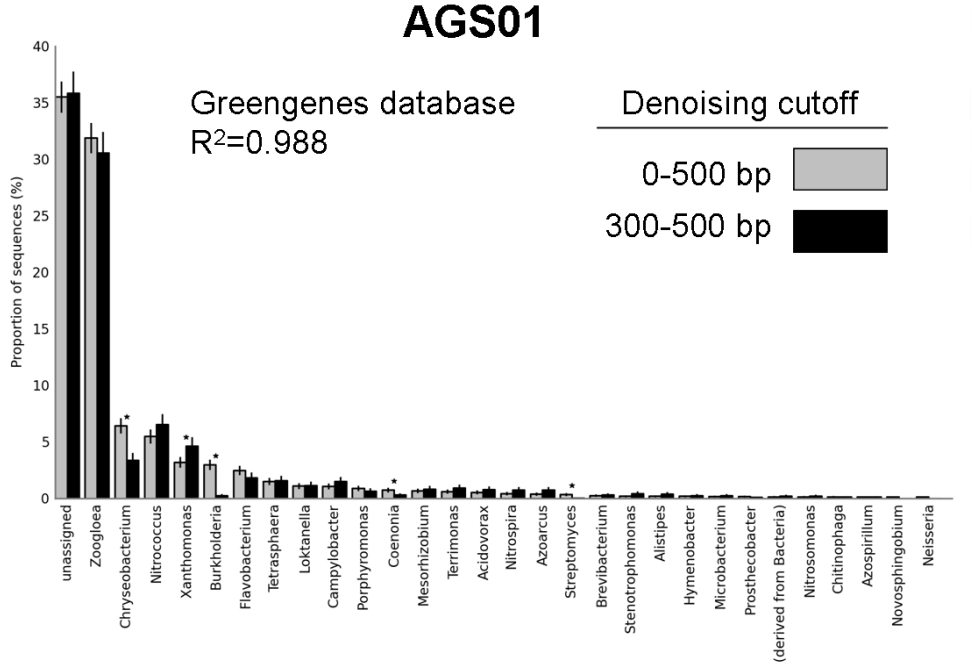
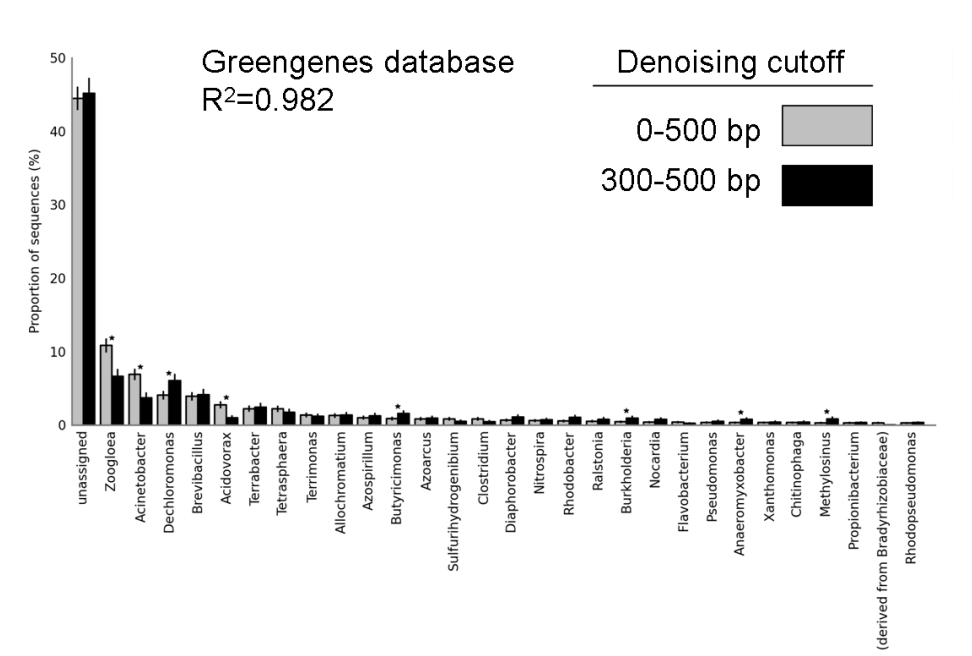


Figure AF2.1. Comparison of bacterial community profiles obtained with the two denoising processes, after mapping with RDP. The first 30 populations are shown.


AGS01
Greengenes database
R2=0.988
Denoising cutoff
0-500 bp
300-500 bp
Proportion of sequences (%)
unassigned
Zoogloea
Chryseobacterium
Nitrococcus
Xanthomonas
Burkholderia
Flavobacterium
Tetrasphaera
Loktanella
Campylobacter
Porphyromonas
Coenonia
Mesorhizobium
Terrimonas
Acidovorax
Nitrospira
Azoarcus
Streptomyces
Brevibacterium
Stenotrophomonas
Alistipes
Hymenobacter
Microbacterium
Prosthecobacter
(derived from Bacteria)
Nitrosomonas
Chitinophaga
Azospirillum
Novosphingobium
Neisseria


FLS01
Greengenes database
R2=0.982
Denoising cutoff
0-500 bp
300-500 bp
Proportion of sequences (%)
unassigned
Zoogloea
Acinetobacter
Dechloromonas
Brevibacillus
Acidovorax
Terrabacter
Tetrasphaera
Terrimonas
Allochromatium
Azospirillum
Butyricimonas
Azoarcus
Sulfurihydrogenibium
Clostridium
Diaphorobacter
Nitrospira
Rhodobacter
Ralstonia
Burkholderia
Nocardia
Flavobacterium
Pseudomonas
Anaeromyxobacter
Xanthomonas
Chitinophaga
Methylosinus
Propionibacterium
(derived from Bradyrhizobiaceae)
Rhodopseudomonas

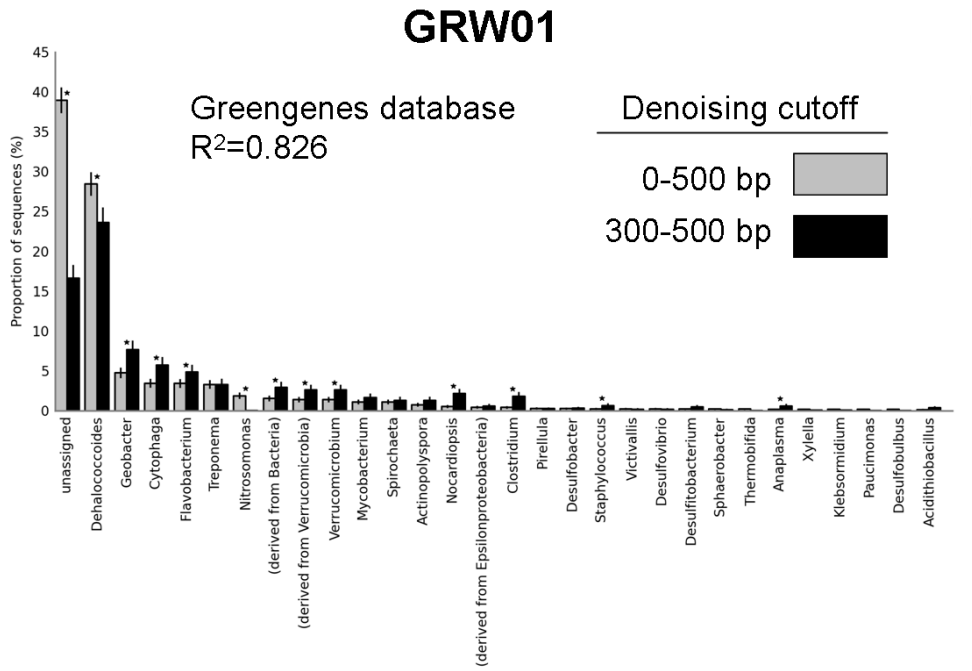
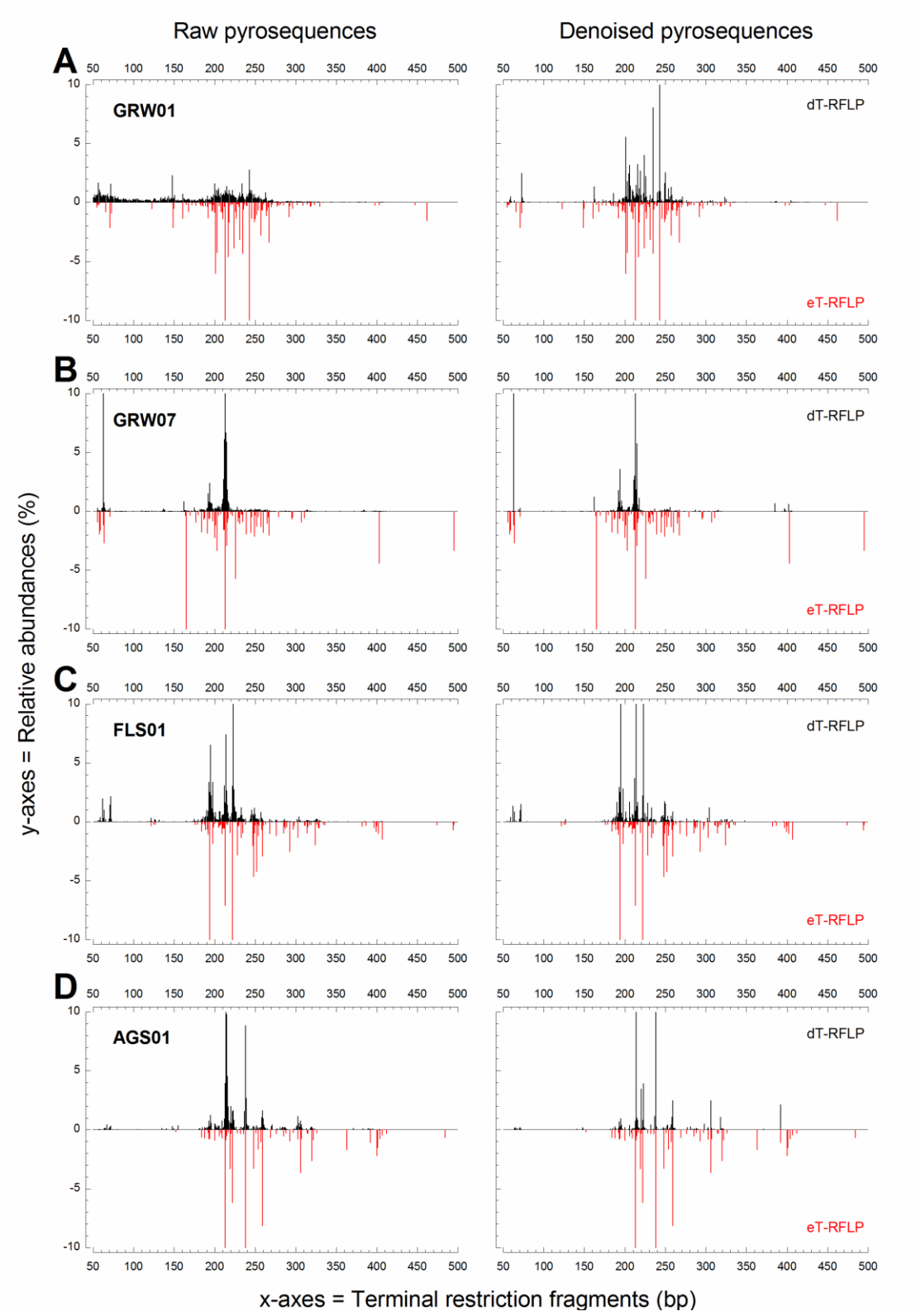
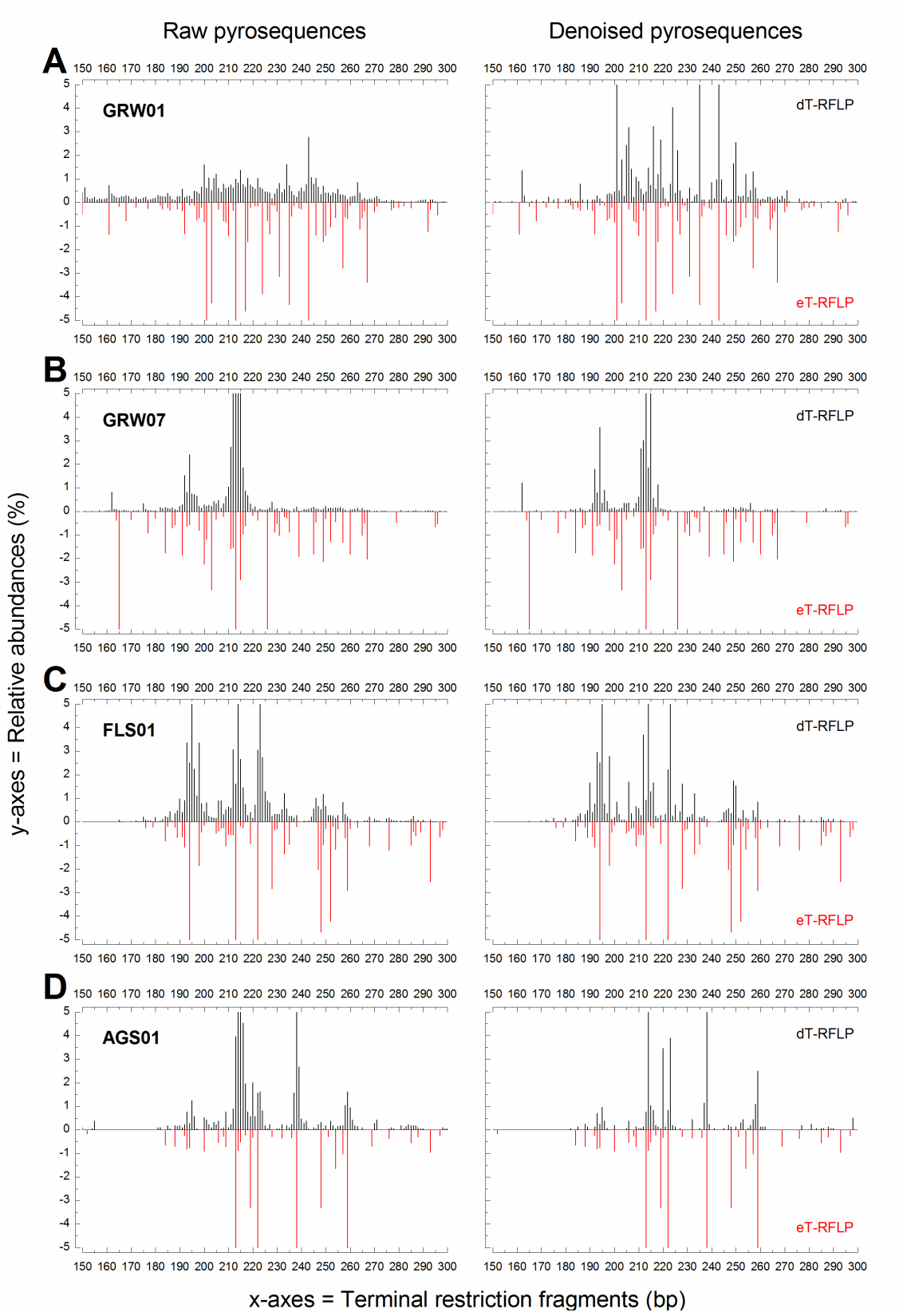


Figure AF2.2. Comparison of bacterial community profiles obtained with the two denoising processes, after mapping with Greengenes. The first 30 populations are shown.

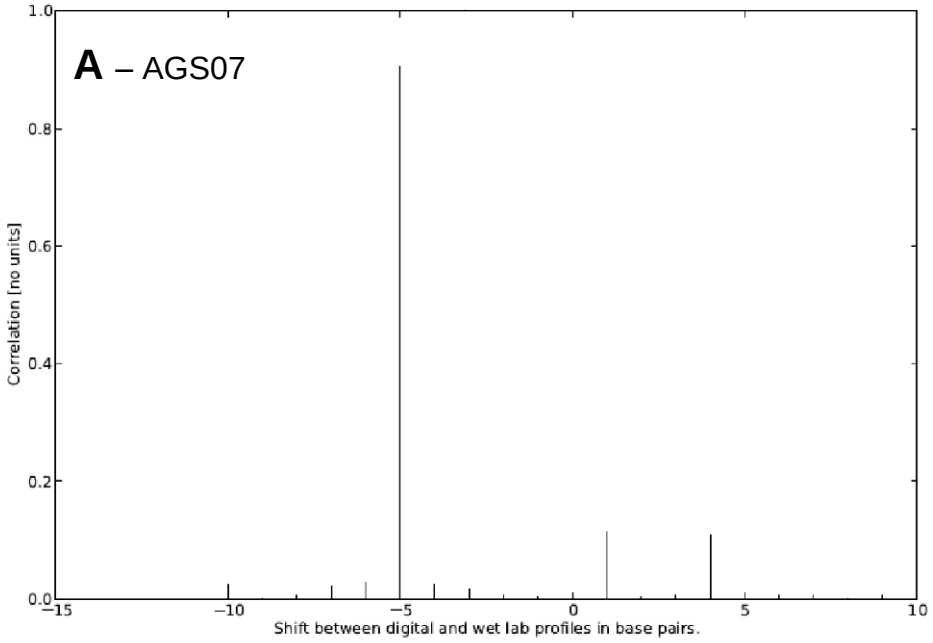
Additional file 3. Comparison of mirror plots obtained on raw and on denoised pyrosequencing datasets. Groundwater sample (GRW01) from a first sampling site pyrosequenced with the HighRA method **(A)**. Groundwater sample (GRW07) from a second sampling site **(B)**, flocculent activated sludge **(C)** and aerobic granular sludge **(D)** samples pyrosequenced with the LowRA method.

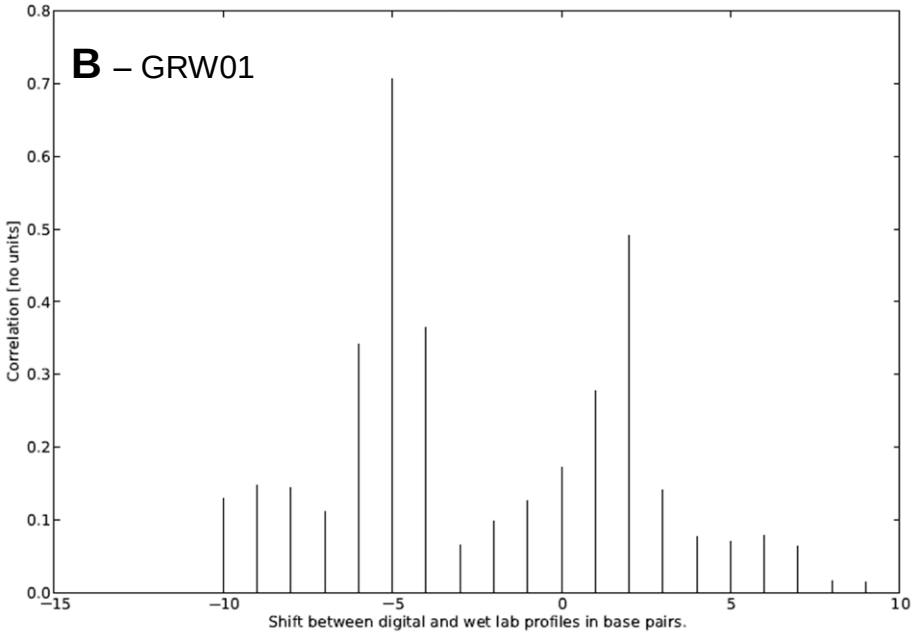


Additional file 3 (zoom)


Raw pyrosequences
Denoised pyrosequences
A
GRW01
B
GRW07
C
FLS01
D
AGS01
dT-RFLP
eT-RFLP
y-axes = Relative abundances (%)
x-axes = Terminal restriction fragments (bp)

Additional file 4. Assessment of cross-correlation and optimal lag between denoised dT-RFLP and eT-RFLP profiles. The denoised digital T-RFLP profiles of the aerobic granular sludge sample AGS07 **(A)** and of the groundwater sample GRW04 **(B)** were both shifted with optimal lags of -5 bp to match with the related experimental fingerprints. At these optimal lags, the maximum cross-correlation coefficients amounted to 0.91 (AGS07) and to 0.71 (GRW04).





Additional file 5. Alignment of sequences mapping with the same Greengenes reference sequence with identical accession number and resulting in different digital T-RFs. Examples are given for the *Rhodocyclus tenuis* affiliates (accession number AB200295) of sample AGS01 and for *Dehalococcoides* relatives (accession number EF059529) of sample GRW05.

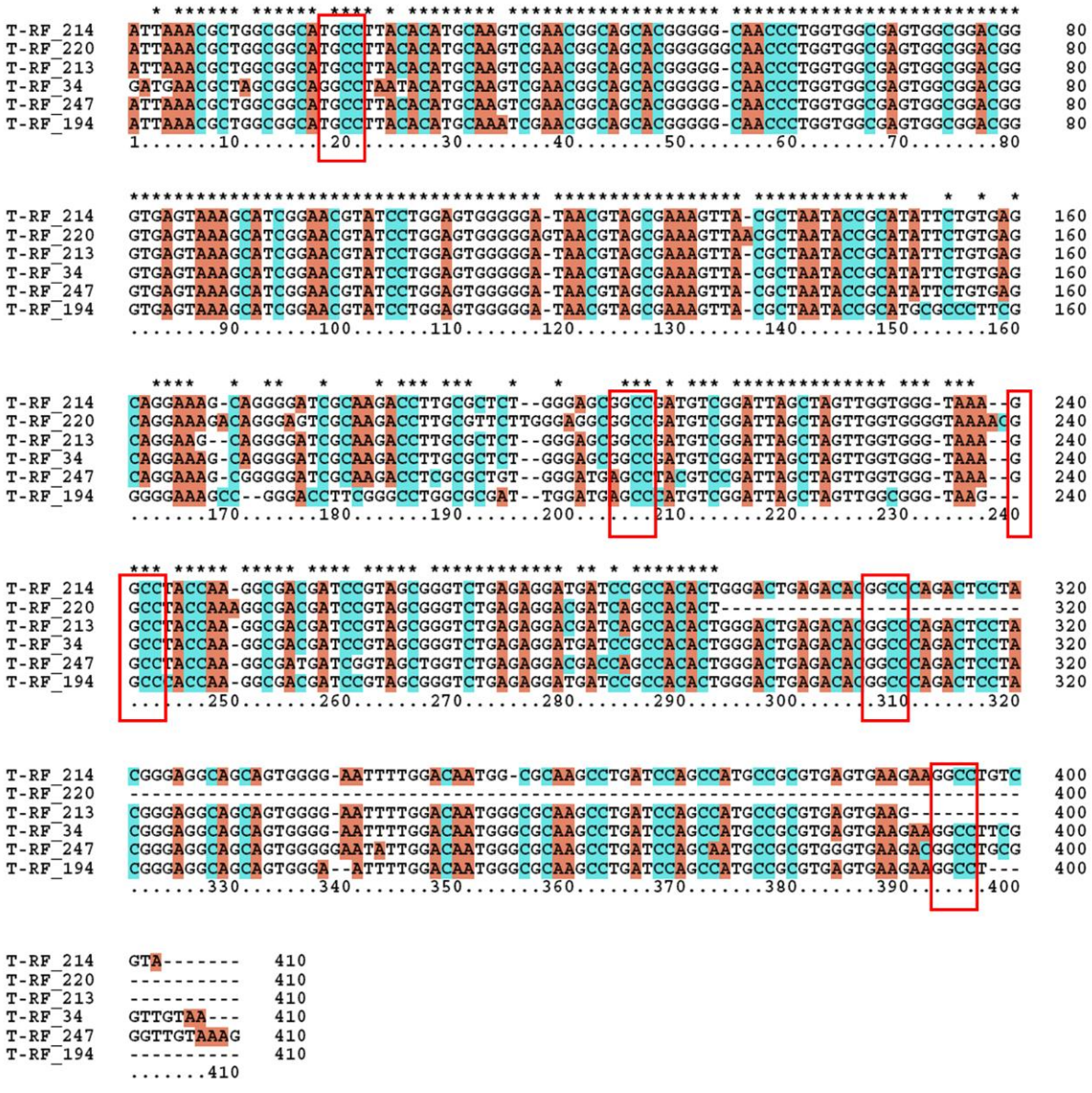


Figure AF5.1. Sequences affiliating with *Rhodocyclus tenuis* (accession number AB200295). The 'GGCC' restriction sites for the diverse T-RF lengths are delimited by the red rectangles.

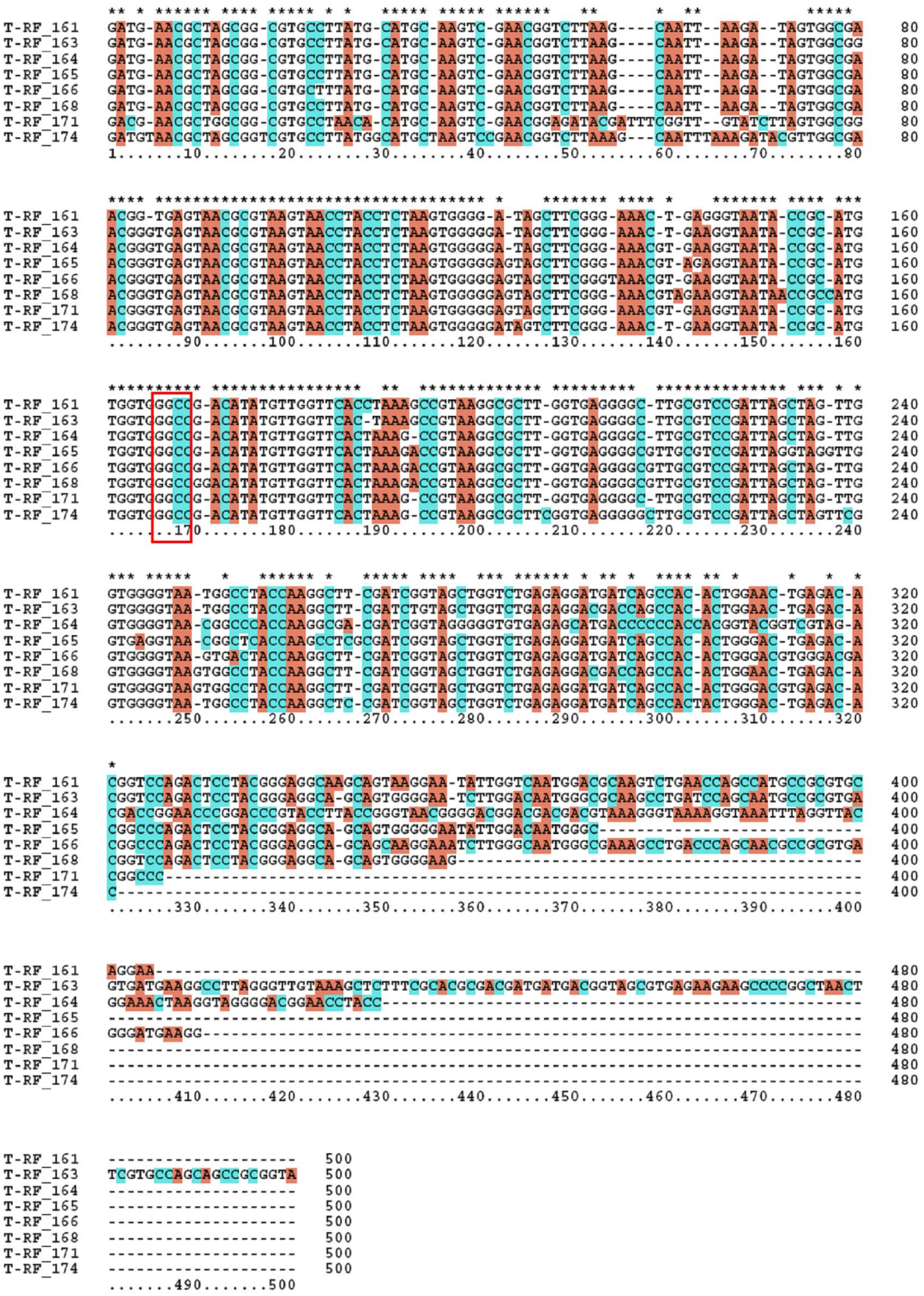


Figure AF5.2. Sequences affiliating with *Dehalococcoides* sp. (accession number EF059529). The common 'GGCC' restriction site of all presented T-RFs is delimited by the red rectangle.

References

1. **Mazzola, M.** (2004) Assessment and management of soil microbial community structure for disease suppression, *Annual Review of Phytopathology* 42 (1):35–59.
2. **Kent, A. D., Yannarell, A. C., Rusak, J. A., Triplett, E. W., and McMahon, K. D.** (2007) Synchrony in aquatic microbial community dynamics, *The ISME Journal* 1 (1):38–47.
3. **Gu, A. Z., Nerenberg, R., Sturm, B. M., Chul, P., and Goel, R.** (2011) Molecular methods in biological systems, *Water Environment Research* 82 (10):908–930.
4. **Marsh, T. L.** (1999) Terminal restriction fragment length polymorphism (T-RFLP): An emerging method for characterizing diversity among homologous populations of amplification products, *Current Opinion in Microbiology* 2 (3):323–327.
5. **Schutte, U. M. E., Abdo, Z., Bent, S. J., Shyu, C., Williams, C. J., Pierson, J. D., and Forney, L. J.** (2008) Advances in the use of terminal restriction fragment length polymorphism (T-RFLP) analysis of 16S rRNA genes to characterize microbial communities, *Applied Microbiology and Biotechnology* 80 (3):365–380.
6. **Militsopoulou, M., Lamari, F. N., Hjerpe, A., and Karamanos, N. K.** (2002) Adaption of a fragment analysis technique to an automated high-throughput multicapillary electrophoresis device for the precise qualitative and quantitative characterization of microbial communities, *Electrophoresis* 23 (7-8):1070–1079.
7. **Thies, J. E.** (2007) Soil microbial community analysis using terminal restriction fragment length polymorphisms, *Soil Science Society of America Journal* 71 (2):579–591.
8. **Mengoni, A., Grassi, E., and Bazzicalupo, M.** (2002) Cloning method for taxonomic interpretation of T-RFLP patterns, *BioTechniques* 33 (5):990–992.
9. **Grant, A., and Ogilvie, L. A.** (2004) Name that microbe: rapid identification of taxa responsible for individual fragments in fingerprints of microbial community structure, *Molecular Ecology Notes* 4 (1):133–136.
10. **Mao, Y., Yannarell, A. C., and Mackie, R. I.** (2011) Changes in N-transforming archaea and bacteria in soil during the establishment of bioenergy crops, *PLoS ONE* 6(9).
11. **Ronaghi, M.** (2001) Pyrosequencing sheds light on DNA sequencing, *Genome Research* 11 (1):3–11.
12. **Roesch, L. F. W., Fulthorpe, R. R., Riva, A., Casella, G., Hadwin, A. K. M., Kent, A. D., Daroub, S. H., Camargo, F. A. O., Farmerie, W. G., and Triplett, E. W.** (2007) Pyrosequencing enumerates and contrasts soil microbial diversity, *The ISME Journal* 1 (4):283–290.
13. **Wommack, K. E., Bhavsar, J., and Ravel, J.** (2008) Metagenomics: Read length matters, *Applied and Environmental Microbiology* 74 (5):1453–1463.
14. **Petrosino, J. F., Highlander, S., Luna, R. A., Gibbs, R. A., and Versalovic, J.** (2009) Metagenomic pyrosequencing and microbial identification, *Clinical Chemistry* 55 (5):856–866.
15. **Sun, Y., Wolcott, R. D., and Dowd, S. E.** (2011) Tag-encoded FLX amplicon pyrosequencing for the elucidation of microbial and functional gene diversity in any environment, *Methods in molecular biology (Clifton, N.J.)* 733: 129–141.
16. **Trombetti, G. A., Bonnal, R. J. P., Rizzi, E., De Bellis, G., and Milanesi, L.** (2007) Data handling strategies for high throughput pyrosequencers, *BMC Bioinformatics* 8 (suppl. 1).
17. **Edwards, R. A.** (2008) The smallest cells pose the biggest problems: High-performance computing and the analysis of metagenome sequence data, *Journal of Physics: Conference Series 125*.

18. **Glenn, T. C.** (2011) Field guide to next-generation DNA sequencers, *Molecular Ecology Resources* 11 (5):759–769.

19. **Desai, N., Antonopoulos, D., Gilbert, J. A., Glass, E. M., and Meyer, F.** (2012) From genomics to metagenomics, *Current Opinion in Biotechnology* 23 (1):72–76.

20. **Rodriguez-Ezpeleta, N., Hackenberg, M., and Aransay, A. M.** (2012) Bioinformatics for High Throughput Sequencing, Springer, New York.

21. **Field, D., Tiwari, B., Booth, T., Houten, S., Swan, D., Bertrand, N., and Thurston, M.** (2006) Open software for biologists: from famine to feast, *Nature Biotechnology* 24 (7):801–803.

22. **Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., and Edwards, R. A.** (2008) The metagenomics RAST server - A public resource for the automatic phylogenetic and functional analysis of metagenomes, *BMC Bioinformatics* 9.

23. **Quince, C., Lanzen, A., Curtis, T. P., Davenport, R. J., Hall, N., Head, I. M., Read, L. F., and Sloan, W. T.** (2009) Accurate determination of microbial diversity from 454 pyrosequencing data, *Nature Methods* 6 (9):639–641.

24. **Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., and Weber, C. F.** (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities, *Applied and Environmental Microbiology* 75 (23):7537–7541.

25. **Huse, S. M., Welch, D. M., Morrison, H. G., and Sogin, M. L.** (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering, *Environmental Microbiology* 12 (7), 1889–1898.

26. **Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Pena, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J., and Knight, R.** (2010) QIIME allows analysis of high-throughput community sequencing data, *Nature Methods* 7 (5):335–6.

27. **Hume, M. E., Barbosa, N. A., Dowd, S. E., Sakomura, N. K., Nalian, A. G., Martynova-Van Kley, A., and Oviedo-Rondon, E. O.** (2011) Use of pyrosequencing and denaturing gradient gel electrophoresis to examine the effects of probiotics and essential oil blends on digestive microflora in broilers under mixed Eimeria infection, *Foodborne Pathogens and Disease* 8 (11):1159–1167.

28. **Jakobsson, H. E., Jernberg, C., Andersson, A. F., Sjolund-Karlsson, M., Jansson, J. K., and Engstrand, L.** (2010) Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome, *PLoS ONE* 5 (3).

29. **Camarinha-Silva, A., Wos-Oxley, M. L., Jauregui, R., Becker, K., and Pieper, D. H.** (2012) Validating T-RFLP as a sensitive and high-throughput approach to assess bacterial diversity patterns in human anterior nares, *FEMS Microbiology Ecology* 79 (1):98–108.

30. **Collins, R. E., and Rocap, G.** (2007) REPK: an analytical web server to select restriction endonucleases for terminal restriction fragment length polymorphism analysis, *Nucleic Acids Research* 35 (suppl. 2):W58–W62.

31. **Marsh, T. L., Saxman, P., Cole, J., and Tiedje, J.** (2000) Terminal restriction fragment length polymorphism analysis program, a web-based research tool for microbial community analysis, *Applied and Environmental Microbiology* 66 (8):3616–3620.

32. **Fernandez-Guerra, A., Buchan, A., Mou, X., Casamayor, E. O., and Gonzalez, J. M.** (2010) T-RFPred: a nucleotide sequence size prediction tool for microbial community description based on terminal-restriction fragment length polymorphism chromatograms, *BMC Microbiology* 10:262.

33. **Aeppli, C., Hofstetter, T. B., Amaral, H. I. F., Kipfer, R., Schwarzenbach, R. P., and Berg, M.** (2010) Quantifying in situ transformation rates of chlorinated ethenes by combining compound-specific stable isotope analysis, groundwater dating, and carbon isotope mass balances, *Environmental Science & Technology* 44 (10):3705–3711.

34. **Ebrahimi, S., Gabus, S., Rohrbach-Brandt, E., Hosseini, M., Rossi, P., Maillard, J., and Holliger, C.** (2010) Performance and microbial community composition dynamics of aerobic granular sludge from sequencing batch bubble column reactors operated at 20 °C, 30 °C, and 35 °C, *Applied Microbiology and Biotechnology* 87 (4):1555–1568.

35. **Rossi, P., Gillet, F., Rohrbach, E., Diaby, N., and Holliger, C.** (2009) Statistical assessment of variability of terminal restriction fragment length polymorphism analysis applied to complex microbial communities, *Applied and Environmental Microbiology* 75 (22):7268–7270.

36. **Lane, D. J.** (1991) 16S/23S rRNA sequencing. In *Nucleic acid techniques in bacterial systematics*, pp 115–175, John Wiley & Sons, New York.

37. **Muyzer, G., Dewaal, E. C., and Uitterlinden, A. G.** (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction amplified genes coding for 16S ribosomal RNA, *Applied and Environmental Microbiology* 59 (3):695–700.

38. **R Development Core Team**. (2008) R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.

39. **Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G. L., Solymos, P., Stevens, M. H. H., and Wagner, H.** (2009) vegan: Community Ecology Package. R package version 1.15-4. *http://CRAN.R-project.org/package=vegan*.

40. **Hall, T. A.** (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, pp 95–98.

41. **Huber, T., Faulkner, G., and Hugenholtz, P.** (2004) Bellerophon: a program to detect chimeric sequences in multiple sequence alignments, *Bioinformatics* 20 (14):2317–9.

42. **Ewing, B., and Green, P.** (1998) Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities, *Genome Research* 8 (3):186–194.

43. **Reeder, J., and Knight, R.** (2010) Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions, *Nature Methods* 7 (9):668–669.

44. **Li, H., and Durbin, R.** (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform, *Bioinformatics* 26 (5):589–95.

45. **McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., Desantis, T. Z., Probst, A., Andersen, G. L., Knight, R., and Hugenholtz, P.** (2011) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea, *The ISME Journal* 6 (3):610-618.

46. **Bukovska, P., Jelinkova, M., Hrselova, H., Sykorova, Z., and Gryndler, M.** (2010) Terminal restriction fragment length measurement errors are affected mainly by fragment length, G plus C nucleotide content and secondary structure melting point, *Journal of Microbiological Methods* 82 (3):223–228.

47. **Kaplan, C. W., and Kitts, C. L.** (2003) Variation between observed and true Terminal Restriction Fragment length is dependent on true TRF length and purine content, *Journal of Microbiological Methods* 54 (1):121–125.

48. **Junier, P., Junier, T., and Witzel, K. P.** (2008) TRiFLe, a program for in silico terminal restriction fragment length polymorphism analysis with user-defined sequence sets, *Applied and Environmental Microbiology* 74 (20):6452–6456.

49. **Huse, S. M., Huber, J. A., Morrison, H. G., Sogin, M. L., and Welch, D. M.** (2007) Accuracy and quality of massively parallel DNA pyrosequencing, *Genome Biology* 8 (7):R143.

50. **Kunin, V., Engelbrektson, A., Ochman, H., and Hugenholtz, P.** (2010) Wrinkles in the rare biosphere: Pyrosequencing errors can lead to artificial inflation of diversity estimates, *Environmental Microbiology* 12 (1), 118–123.

51. **Niu, B., Fu, L., Sun, S., and Li, W.** (2010) Artificial and natural duplicates in pyrosequencing reads of metagenomic data, *BMC Bioinformatics* 11 (1):187.

52. **Gilbert, M. T. P., Binladen, J., Miller, W., Wiuf, C., Willerslev, E., Poinar, H., Carlson, J. E., Leebens-Mack, J. H., and Schuster, S. C.** (2007) Recharacterization of ancient DNA miscoding lesions: Insights in the era of sequencing-by-synthesis, *Nucleic Acids Research* 35 (1):1–10.

53. **Legendre, P., and Legendre, L.** (1998) Numerical ecology: second english edition, Elsevier, Amsterdam.

54. **Kitts, C. L.** (2001) Terminal restriction fragment patterns: A tool for comparing microbial communities and assessing community dynamics, *Current Issues in Intestinal Microbiology* 2 (1):17–25.

55. **Clement, B. G., Kehl, L. E., DeBord, K. L., and Kitts, C. L.** (1998) Terminal restriction fragment patterns (TRFPs), a rapid, PCR-based method for the comparison of complex bacterial communities, *Journal of Microbiological Methods* 31 (3):135–142.

56. **Osborn, A. M., Moore, E. R. B., and Timmis, K. N.** (2000) An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics, *Environmental Microbiology* 2 (1):39–50.

57. **Egert, M., and Friedrich, M. W.** (2003) Formation of pseudo-terminal restriction fragments, a PCR-related bias affecting terminal restriction fragment length polymorphism analysis of microbial community structure, *Applied and Environmental Microbiology* 69 (5):2555–2562.

58. **Osborne, C. A., Rees, G. N., Bernstein, Y., and Janssen, P. H.** (2006) New threshold and confidence estimates for terminal restriction fragment length polymorphism analysis of complex bacterial communities, *Applied and Environmental Microbiology* 72 (2):1270–1278.

59. **Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., Kulam-Syed-Mohideen, A. S., McGarrell, D. M., Marsh, T., and Garrity, G. M.** (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis, *Nucleic Acids Research* 37: D141-D145.

60. **Parks, D. H., and Beiko, R. G.** (2010) Identifying biologically relevant differences between metagenomic communities, *Bioinformatics* 26 (6):715-721.

Accumulation of lower chlorinated ethenes in a contaminated aquifer due to competition between iron- and organohalide-respiring bacterial guilds

4 Accumulation of lower chlorinated ethenes in a contaminated aquifer due to competition between iron- and organohalide-respiring bacterial guilds

Abstract

Natural attenuation of tetrachloroethene (PCE)-contaminated aquifers often results in the accumulation of the toxic intermediates cis-dichloroethene and vinyl chloride (VC) which are even more toxic than the parent compound. Reasons for this accumulation were investigated in a PCE-contaminated aquifer using multivariate statistics applied to bacterial community structures determined by terminal-restriction fragment length polymorphism (T-RFLP) and environmental variables. Multifactorial analysis showed that both VC and iron reductions (IR) were key terminal electron accepting processes, suggesting a competition between iron- and VC-reducing bacteria. T-RFs showing significant correlations with VC reduction were identified using the PyroTRF-ID bioinformatics tool and found to be closely affiliated to the genus "*Dehalococcoides"* and to uncultured bacteria of the "Lahn Cluster" (LC) within the class *Dehalococcoidetes*. A major T-RF negatively correlated with the LC T-RF was affiliated to the genus *Rhodoferax* that contains members identified as iron-reducing bacteria (IRB). Finally, the level of activity of sulfate-reducing bacteria probably played an important role in the present environment by decreasing Fe(III) contents through a mutualistic interaction with IRB.

4.1 Introduction

Monitored natural attenuation (MNA) is a cost-effective strategy for bioremediation of sites contaminated with chlorinated ethenes (CEs) such as tetrachloroethene (PCE) or trichloroethene (TCE). The main drawback of this technique is the frequent undesired accumulation of the toxic intermediates cis-dichloroethene (cDCE) and vinyl chloride (VC) (*1–4*). This can be the result either of an incomplete sequence of dechlorination or of a decreased efficiency of lower CEs dechlorination (*5*). Incomplete or impeded reductive dechlorination has been formerly explained by (i) the absence of bacterial species able to dechlorinate down to ethene (*6–8*), (ii) the depletion of the organic substrates used as electron donors by fermenting guilds (*9*, *10*), and (iii) the competition for the electron donors with other TEAPs (*11–13*). The reasons for lower CEs accumulation have been investigated in microcosm studies or enrichment cultures (*14–16*). For example, Hoelen & Reinhard (2004) showed that sulfate reduction (SR) affected the rate of lower CEs dechlorination in microcosm experiments (*15*). Addition of toluene stimulated H_2 production probably via fermentation, and increased the rate of dechlorination. Similar conclusions were obtained by Heimann et al. (2005) in mixed anaerobic cultures (*14*). Wei et al. (2011) showed that iron reduction (IR) was not inhibiting organohalide respiration (OHR) (*16*).

Reductive dechlorination has been previously investigated in the Zuchwil aquifer. Aeppli et al. (2010) have used carbon stable isotope analysis to analyze the extent and the rates of dechlorination *in situ* (*17*). In this study, qualitative evidence of PCE reductive dechlorination was provided by the detection of significant concentrations of daughter compounds such as TCE, cDCE, and VC, as well as by the observation of a gradual dechlorination pattern along a transect from the source zone to the fringe of the contaminant plume. Along this transect, proportions of cDCE and VC were increasing with increasing distance to the source. Similar conclusion was obtained from the carbon isotopic signature. The authors have shown that PCE was efficiently transformed at most sampling points, with degrees of dechlorination of more than 60%. Isotopic mass balance based on the concentration-weighted ^{13}C isotopic signatures of CEs has enabled assessing the potential of VC accumulation on the site. The authors have concluded that VC was overall accumulating on the site, although complete OHR to ethene was probably occurring locally. The reasons for this accumulation have not been investigated at this stage, although the hypothesis was formulated that sulfate-reducing bacteria (SRB) were outcompeting VC-reducing bacteria.

Based on the findings that VC was accumulating, the objective of this study was to identify the mechanisms leading to lower CEs accumulation *in situ* in the CE-contaminated Zuchwil aquifer. We adopted an integrative approach to first assess the aquifer ecosystem as a whole. The investigation of the microbial processes running in the Zuchwil aquifer was carried out using: (i) PCR detection of both reductive dehalogenase functional genes and OHR bacterial guild members, (ii) bacterial community structures (BCS) analysis using terminal-restriction fragment length polymorphism (T-RFLP), (iii) exploration of the relationships between BCS and measured environmental variables using multivariate statistics in order to

identify the mechanisms leading to the observed incomplete degradation of CEs, and (iv) finally, identification of T-RFs significantly correlated with TEAP variables both by Sanger sequencing of cloned 16S rRNA genes and by pyrosequencing of the 16S rRNA gene pool analyzed using the PyroTRF-ID software (see CHAPTER 3).

4.2 Materials and Methods

4.2.1 Test site

The test site and the natural attenuation process have been previously described in details (*17*). Briefly, the aquifer is composed of 4-12 m of silty sand and gravel of glacio-fluvial deposition, covered by a 2-5 m thick confining layer of silty clay. It is characterized by very low Darcy groundwater velocities of 1-8 m/yr (hydraulic conductivities: 1-8 x 10^{-5} m/s; average hydraulic gradient: 0.3%). According to a historical investigation, the contamination with PCE most probably occurred in 1969. Today, the area is covered with buildings, parking areas and roads, and is equipped with monitoring wells to follow-up the natural attenuation process. A total of 17 monitoring wells were chosen for the present study, covering several contamination states, from the highly contaminated source zone (RB1/99) to the plume fringe (KB6, P20) (Figure 1).

4.2.2 Groundwater sampling

Thirty-one groundwater samples were collected during two sampling campaigns in April and October 2008. Sampling wells were purged using a 12 V submersible pump (Whale submersible 921, Whale Water Systems, Bangor, Northern Ireland) until temperature, electric conductivity (EC), dissolved oxygen (DO), pH, and redox potential values (Eh) reached stable readings, which corresponded to at least three well volumes. Samples for both chemical and microbiological analyses were collected at a flow rate of about 100 mL/min using a peristaltic pump (Type P2.52, Eijkelkamp, Giesbeek, The Netherlands) through PTFE tubing. Water samples for the analyses of CE concentrations were stored in 120 mL glass bottles (Infochroma AG, Switzerland), completely filled, sealed using PTFE-lined screw caps, and stored in the dark at 4°C until analysis that was performed within four weeks from time of sampling. Samples for the analysis of major anions and total organic carbon (TOC) were stored in 125 mL Nalgene® polypropylene bottles (Thermo Fisher Scientific, USA) at 4°C. Aliquots for analysis of cations, dissolved Fe^{2+}, Mn^{2+} and Sr^{2+} were filtered on-site through 0.45 µm nitrocellulose membranes (Millipore, USA), acidified to pH < 2 with ultrapure nitric acid (Sigma-Aldrich, USA) and stored at 4°C. TOC was analyzed within one day, whereas anions, cations and dissolved metals were analyzed within one week. Groundwater samples for the microbial analyzes were stored in sterile 1-L polypropylene Nalgene® bottles (Thermo Fisher Scientific, USA) in the dark and at 4°C and were processed within 2 days.

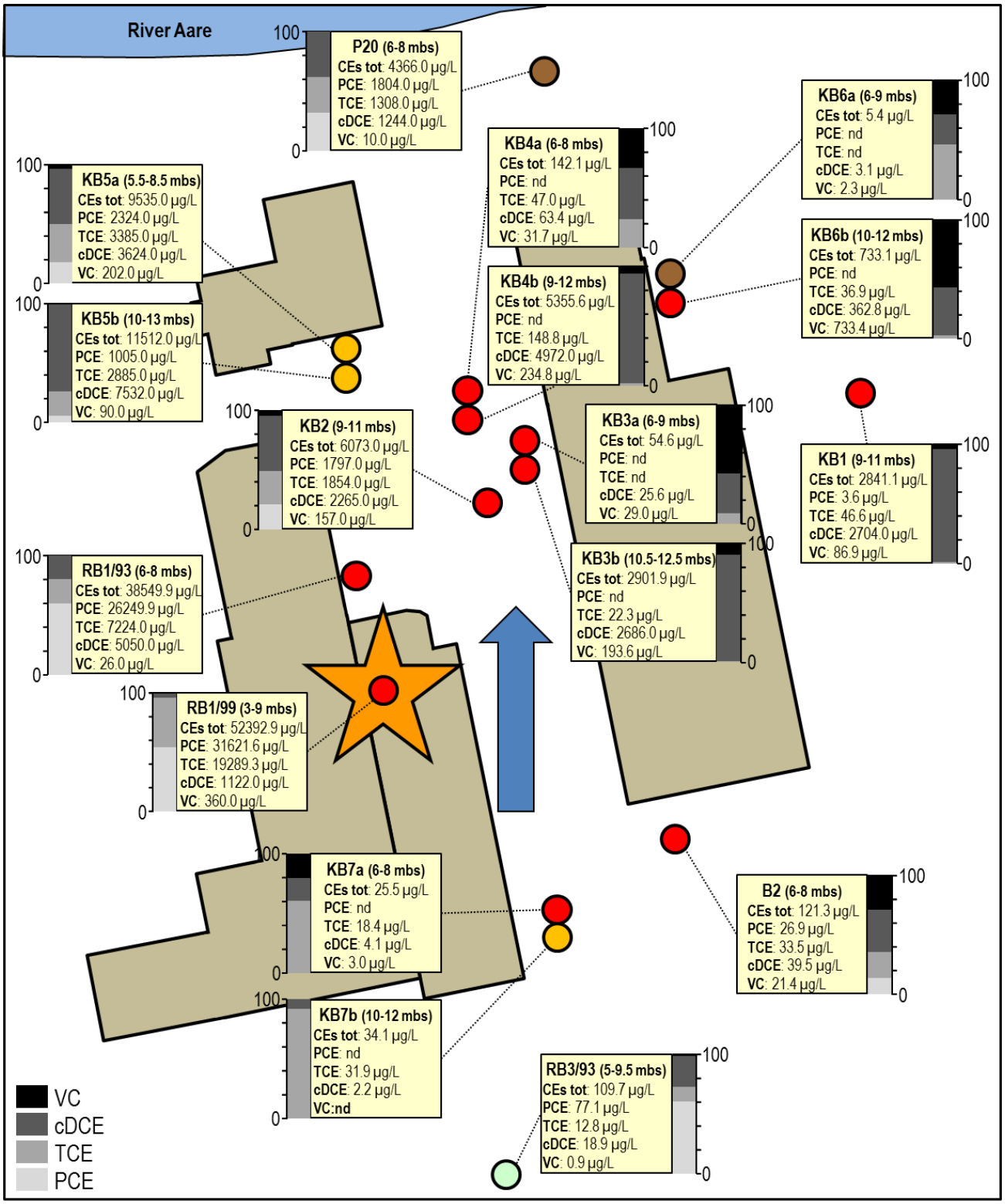


Figure 1. Map of the contaminated test site indicating the monitoring wells with colored circles, buildings present during the sampling campaign of April 2008 (brownish square shapes), main groundwater flow direction (blue arrow) and source of PCE contamination (orange star). The colors of the circles indicate the oxidation-reduction state of the aquifer at the monitoring location (green: oxic / nitrate-reducing; orange: nitrate- to Fe(III)-reducing; red: Fe(III)-reducing; brown: methanogenic). Names of the monitoring wells are indicated in yellow squares together with the chlorinated ethenes (CEs) concentrations measured in April 2008 and the percentage of each CE (the grey scale is given on the bottom left of the figure).

4.2.3 Chemical analyzes

Dissolved oxygen (DO), NO_3^-, SO_4^{2-}, Fe^{2+} and Mn^{2+} concentrations were assessed as indicators of the electron-accepting processes occurring in the aquifer. Methane and ethene, quantified only in the second sampling event, also provided information about the redox conditions of the aquifer. Major anions (Cl^-, NO_3^-, SO_4^{2-}) and cations (Na^+, K^+, Ca^{2+}, Mg^{2+}, NH_4^+) were analyzed by ion chromatography respectively in a DX-500 with an AS12A column and an ICS-3000A with an IonPacCS16 column (Dionex Corp., USA). Fe^{2+} and Mn^{2+} were quantified with an inductively coupled plasma optical emission spectrometer (ICP-OES, Perkin-Elmer Optima 3300, Perkin-Elmer, USA) and TOC was analyzed in a Shimadzu ASI-5000 TOC-analyzer (Shimadzu Corp., Japan). CE concentrations were quantified using direct aqueous injection-GC/MS (*18*) with typical accuracy of 92-108%. Ethene and methane were measured by GC-FID in samples collected in November 2008, providing qualitative information only. The aquifer oxidation-reduction status was defined using an Excel spreadsheet program designed to assist in the assignment of redox conditions based on water chemistry data (*19*, *20*).

To explore the relationships between the environmental variables, pairwise Pearson product-moment correlation coefficients were calculated with R (*21*) as well as their p-values using a permutation test with 1000 permutations.

4.2.4 DNA extraction

Water samples were filtered through 0.2 µm autoclaved polycarbonate membranes (Isopore™ Membrane Filters, Millipore) with a mobile filtration system (Filter Funnel Manifolds, Pall Corporation). DNA was extracted using the PowerSoil™ DNA Extraction Kit (Mo Bio Laboratories, Inc.) following the manufacturer instructions, except that the samples were processed in a bead-beater (Fastprep FP120, Bio101) at 4.5 m/s for 30 s after the addition of solution C1. Extracted DNA was quantified with a ND-1000 Nanodrop® spectrophotometer (Thermo Fisher Scientific, USA).

4.2.5 Specific detections of OHRB and reductive dehalogenase genes

Presence of OHR-related genes (16S rRNA and reductive dehalogenase (*rdhA*) genes) was assessed by PCR amplification, and the primers are listed in Table 1. Unless stated otherwise, all PCR reagents were provided by PeqLab Biotechnologie GmbH (Germany). Amplification reactions were carried out in a Thermal Cycler T3 Biometra (Biolabo Scientific Instruments, Switzerland). The VC reductase genes *vcr*A and *bvc*A (GenBank YP003330719 and AY563562) were amplified in 10 µl reaction mixtures containing 1 µl 10x buffer S, 0.3 µl of 10 mM dNTPs, 1 µl of each primer at 10 µM, 0.3 µl DMSO (Sigma-Aldrich, USA), 0.5 U of PeqGOLD Taq polymerase. 247 and 139 base pairs DNA fragments were amplified as follows: 4 min 30 s preheating at 94°C, 40 cycles of 30 s denaturation at 94°C, 45 s of primer annealing at 55°C, 30 s of

elongation at 72°C, with a 10 min final elongation step at 72°C. The 16S rRNA genes of OHRB were amplified in a nested PCR procedure. In the first step, 16S rRNA gene pools were amplified using Eubacterial primers Eub-27f / Univ-518r as follows: 10 µl reaction mixtures contained 1 µl 10x buffer Y, 0.8 µl of 10 mM dNTPs, 0.5 µl of each primer at 10 µM, 0.3 µl DMSO (Sigma-Aldrich, USA), 0.5 U of PeqGOLD Taq polymerase. Amplification was carried out with the following program: 4 min 30 s preheating at 94°C, 30 cycles of 30 s denaturation at 94°C, 45 s of primer annealing at 56°C, 1 min of elongation at 72°C. A final elongation step of 10 min at 72°C was included. The PCR products were visually inspected on 1.5% agarose gels stained with GelRed (Stratec Molecular, Germany) and diluted prior to the second amplification step. In the second amplification step, taxon-specific primers were used to amplify the 16S rRNA gene of known OHRB as follows: 10 µl reaction mixtures contained 1 µl 10x buffer Y, 0.8 µl of 10 mM dNTPs, 0.5 µl of each primer at 10 µM, 0.3 µl DMSO (Sigma-Aldrich, USA), 0.5 U of PeqGOLD Taq polymerase. Amplifications were carried out with the following program: 4 min 30 s preheating at 94°C, 30 cycles of 30 s denaturation at 94°C, 45 s of primer annealing at 55°C, 45 s of elongation at 72°C. A final elongation step of 10 min at 72°C was included.

Table 1. List of primers used to amplify specific organohalide-respiring bacterial genera (16S rRNA gene) or reductive dehalogenase-encoding genes.

Target organism / gene	**Name of primers**	**Sequence (5'-3')**	**Reference**
Eubacteria 16S rRNA gene	Eub27f	AGAGTTTGATCMTGGCTCAG	(*22*)
	Univ518r	ATTACCGCGGCTGCTGG	(*23*)
Dehalococcoides sp.	DHC587f	GGACTAGAGTACAGCAGGAGAAAAC	(*8*)
	DHC1212r	GGATTAGCTCCAGTTCACACTG	(*8*)
Dehalobacter sp.	Dre441f	GTTAGGGAAGAACGGCATCTGT	(*24*)
	Dre645r	CCTCTCCTGTCCTCAAGCCATA	(*24*)
Sulfurospirillum sp.	DHSPM576f	GCTCTCGAAACTGGTTACCTA	(*25*)
	DHSPM1210r	GTATCGCGTCTCTTTGTCCTA	(*25*)
Desulfuromonas sp.	Desulfo494f	AGGAAGCACCGGCTAACTCC	(*26*)
	Desulfo1050r	CGATCCAGCCGAACTGACC	(*26*)
Desulfitobacterium sp.	Dsb406f	GTACGACGAAGGCCTTCGGGT	(*24*)
	De2	CCTAGGTTTTCACACCAGACTT	(*27*)
*vcr*A	vcrA880f	CCCTCCAGATGCTCCCTTTA	(*28*)
	vcrA1018r	ATCCCCTCTCCCGTGTAACC	(*28*)
*bvc*A	bvcA227f	TGGGGACCTGTACCTGAAAA	(*28*)
	bvcA523r	CAAGACGCATTGTGGACATC	(*28*)

4.2.6 Terminal-restriction fragment length polymorphism (T-RFLP)

The analysis of the bacterial communities using terminal-restriction fragment length polymorphism (T-RFLP) as well as their numerical treatment were based on the protocol described in (*29*) with the following modifications: (i) 30 µl PCR reactions were composed of 3 µL 10X Y buffer, 2.4 µL 10 mM dNTPs, 1.5 µL of each 10 µM primer, 6 µL 5X enhancer P, 1.5 U PeqGold Taq polymerase and 0.2 ng/µl template DNA (final concentration), completed with autoclaved and UV-treated Milli-Q water (Millipore, USA); (ii) for each DNA extract, PCR amplification was carried out in triplicates; (iii) the PCR products were purified with the purification kit MSB® Spin PCRapace (Stratec Molecular GmbH, Germany) according to the manufacturer's instructions.

4.2.7 Multivariate statistical analyses based on the T-RFLP profiles

All multivariate statistical analyses in this section were carried out with R (*21*) and the additional package Vegan (*30*). In a preliminary step, the bacterial community dataset was constructed as follows: Ruzicka dissimilarities were calculated between replicate T-RFLP profiles; the profile at the centroid (i.e. displaying the lowest dissimilarity with its replicates) was selected for each sample. The bacterial community dataset was transformed using a Hellinger transformation (*31*), whereas the environmental dataset was standardized (*32*).

Multifactorial analysis (MFA) (*33*, *34*) was chosen to analyze the data. This symmetrical canonical ordination technique based on principal component analysis (PCA) allowed the simultaneous assessment of 4 datasets: (i) the bacterial community dataset, (ii) the proportion of each CE among the total amount of all CEs ("%CEs"), indicating the reduction state of the contaminant plume and a potential accumulation of CEs, (iii) environmental variables related to TEAPs ("TEA"), and (iv) all remaining variables ("OTHERS") (see Tables S1 to S4 for details). RV multivariate correlation coefficients and p-values (*35*) were calculated in order to get numerical correlations between the environmental dataset and the bacterial communities. They were computed between the bacterial community dataset and (i) the whole environmental dataset, (ii) each defined group of environmental variables (%CEs, TEA and OTHERS), and (iii) single environmental variables. Spearman pairwise correlations were computed between single T-RFs and environmental variables so as to reveal significantly positive or negative correlations with ongoing biogeochemical processes.

Finally, T-RFs co-occurrence was assessed according to their relative abundance among samples with an agglomerative hierarchical clustering using the Ward's minimum variance algorithm. The Mantel-optimal number of clusters was chosen to create the clusters (*36*). T-RFs belonging to the same cluster were considered as co-occurring.

4.2.8 Identification of T-RFs

Cloning and sequencing

DNA extracts of samples containing T-RFs showing significant correlations with selected environmental variables were used to construct clone libraries. For this purpose, the total 16S rRNA gene pool was amplified by PCR as described above, except that the 8f primer was not labeled. The PCR products were purified as described above, ligated into pGEM®-T Easy Vector (Promega, USA) and transformed into *E.coli* XL1-Blue competent cells (Agilent Technologies, USA). The clone libraries were first screened by applying the T-RFLP procedure on each clone separately. Clones containing T-RFs with interesting correlations with the environmental variables were isolated. The 16S rRNA gene was extracted from the plasmid using a PCR amplification with primers T7 and SP6 (Promega, USA) and purified as described above. A sequencing reaction was carried out on each purified PCR product as follows: 10 µl reaction mixtures contained 2 µl of Terminator mix V3.1 (Applied Biosystems Inc., USA), 2 µl of 5X BigDye V3.1 Sequencing buffer (Applied Biosystems Inc., USA), 1.6 µl of primer T7 at 1 µM, 200 ng DNA template and completed with autoclaved and UV-treated Milli-Q water (Millipore, USA). The sequencing reaction was carried out with the following program: 30 cycles of 10 s denaturation at 94°C, 5 s of primer annealing at 50°C, and 4 min of elongation at 60°C. The products were sequenced in an ABI 3130xl DNA capillary sequencer (Applied Biosystems Inc., USA) equipped with 50 cm long capillaries (80 µm inner diameter) and POP 7 electrophoresis matrix (dye set G5, Applied Biosystems Inc., USA). Sequences were tested for chimeras using Bellerophon (*37*) and aligned with ClustalX 2.1 (*38*). They were digested *in silico* with HaeIII, and compared to the experimental length measured by T-RFLP. Sequences showing a discrepancy higher than 5bp between the in silico and the experimental T-RF were discarded, as well as short sequences. The remaining sequences were compared with those available in GenBank® (*39*) by BLAST searches (*40*) to find the nearest relatives and percentage of gene sequence identity. Phylogenetic trees were built with Mega5 (*41*).

Pyrosequencing of selected aquifer samples and Pyrosequencing T-RFs Identification software (PyroTRF-ID)

The selection of six aquifer samples was carried out according to the results obtained by the clustering of all T-RFLP profiles (n = 30) using the Hellinger distance and the Ward algorithm. One sample per cluster was selected for further pyrosequencing analysis (GATC, Germany) using a 454/Roche GS-FLX Titanium Instrument (Roche, NJ, USA). Amplicons of the V3 region of the 16S rRNA genes were produced and purified in the exact same way as shown above. PyroTRF-ID software (see CHAPTER 3) running on the Vital-IT plateform (Vital-IT, Swiss Institute of Bioinformatics, Switzerland) was used to handle pyrosequencing data and identify T-RFs. Briefly, quality assessment, de-multiplexing (with forward primer removal) and de-noising of the sequences was carried out with QIIME libraries (*42*) implemented in the PyroTRF-ID software. Mapping of the sequences was performed using BWA (*43*) and the Greengenes database (*44*). Digital T-RFLP profiles were produced on the basis of 5234 reads per sample on average,

with an average length of 402 bp. Reads were digested *in silico* with the *Hae*III endonuclease (5'-GG^CC-3' enzyme's recognition site). Most probable drift between experimental and digital T-RFLP profiles (eT-RFLP and dT-RFLP, respectively) (*45*, *46*) was assessed by computing cross-correlations, and the dT-RFLP profile was shifted accordingly. A consistent match was then obtained between eT-RFs obtained experimentally and dT-RFs obtained *in silico*. The putative corresponding sequence affiliation of eT-RFs previously linked with significant statistical correlations with environmental variables was thus obtained.

4.3 Results

Groundwater samples from two sampling campaigns in 2008 have been analyzed to investigate the possible reasons for VC accumulation observed in the Zuchwil aquifer. Environmental variables and bacterial communities have been assessed and jointly analyzed using multivariate statistical tools.

4.3.1 Environmental variables

Groundwater level was 0.1 to 0.3 m higher in October than in April, indicating that the aquifer was recharged by heavy rain events preceding shortly the second sampling campaign. Accordingly, the mean groundwater temperature was 1.8 °C higher in October than in April. The pH values ranged from 6.2 to 7.0, with a mean pH 6.9 measured for both sampling events. TOC concentrations measured in all groundwater samples averaged 3.2 mg/L. The reference well RB3/93 displayed the lowest concentrations (1.2 and 1.4 mg/L in April and October, respectively). Maximal TOC concentrations were measured in well RB1/99, with 6.2 and 4.9 mg/L in April and October, respectively. Organic carbon content in the sediments was estimated from geochemical analyzes performed previously on core samples obtained during the installation of monitoring wells (data not shown). It was comprised between 6 and 9% (w/w), based on the approximation that approximately 50% of the mass of combustible organic matter is composed of organic carbon (*47*). Very low values of dissolved oxygen concentrations (DO) were measured during the first sampling event (0.1 to 0.3 mg/L), with the exception of well KB3b (1.0 mg/L). Nitrate concentrations were low as well (0.12 to 0.74 mg/L), except in well KB7b in October (1.29 mg/L). The highest values were found in well RB3/93 (5.51 and 1.77 mg/L in April and October, respectively) located upstream of the contaminated zone. Sulfate concentrations varied considerably, with lowest concentrations found in wells RB1/99, P20, KB4b, and KB6a (0.2 to 1.9 mg/L). Higher sulfate concentrations were measured in wells B2, KB2, KB3a, KB4a, KB5a, and KB7a, ranging from 2.2 to 4.6 mg/L, and in wells RB1/93, KB1, KB3b, KB5b, and KB7b, ranging from 9.2 to 15.8 mg/L. The highest sulfate concentrations were measured in well RB3/93, with a mean value of 24.9 mg/L. Sulfide was not measured in this study, since previous investigation (*17*) indicated that it was always at or below the detection limit of 0.6 mg/L. Fe^{2+} was below detection limit in well RB3/93 (< 0.1 mg/L). Higher Fe^{2+} concentrations (1.3 to 4.4 mg/L) were measured in wells RB1/93, RB1/99, KB3a,

KB5a, KB5b, KB6a and KB7b, indicating that IR was taking place (*48*, *49*). Particularly high Fe^{2+} concentrations were detected in wells B2, P20, KB1, KB2, KB3b, KB4a, KB4b, KB6b, and KB7a, ranging from 7.5 to 18.7 mg/L. Mn^{2+} concentrations were comprised between 0.13 and 1.90 mg/L, except in well RB3/93 in which it was barely detected (0.05 mg/L in April and below detection limit in October). The redox states calculated based on ion concentrations showed that most wells were in Fe(III)-reducing conditions. Only P20 and KB6a, located at the plume fringe, were in a methanogenic redox state, whereas wells KB5 (a and b), KB7b and RB3/93 were in nitrate-reducing state. Well KB3b was slightly oxic.

The following section describing CE concentrations is based on measurements carried out in April 2008. Though the concentrations in October were slightly different from those measured in April, the overall pattern of concentrations along the site was similar. Total CE concentrations varied throughout the site, reaching about 50 mg/L in monitoring well RB1/99 (Figure 1). PCE and TCE concentrations were highest in the vicinity of the source zone with cumulated concentrations of 95.1 and 80.3% in monitoring wells RB1/99 and RB1/93, respectively. Both concentrations and proportions decreased together with increasing distance to the source zone. cDCE concentrations and proportions already increased along the short distance separating monitoring wells RB1/99 and RB1/93, passing from 1122 to 5050 µg/L. Highest concentrations of this compound were measured in well KB5b (7532 µg/L), and its highest proportion was observed in well KB1 (94%). VC concentrations were highest in RB1/99 in the vicinity of the source zone, but the proportion of VC among all CEs was low in this well (1.6%) and increased together with the distance to the source zone. The highest proportion of VC was reached in well KB3a. Total CE concentrations in this well amounted 54.6 µg/L only, and only cDCE and VC could be quantified therein. At the plume fringe, concentration of VC was high in well KB6b, amounting 57% of all CEs. CEs (110 µg/L cumulated concentrations) were detected in well RB3/93 during the first sampling campaign only. All details about environmental variables are given in Tables S1 to S4.

4.3.2 Pairwise correlations between environmental variables

Pairwise correlations were calculated to explore the relationships between environmental variables (Figure 2). "NO_3^-", indicating the energetically most favorable TEAP after aerobic respiration, as well as "SO_4^{2-}", was anti-correlated with all TEA variables. Both NO_3^- and SO_4^{2-} are parent compounds that were present in relatively high concentration in oxic water and drastically lower concentrations in most monitoring wells of the Zuchwil aquifer. "SO_4^{2-}" was strongly anti-correlated with "TOC", showing that SR was more active in fractions of the aquifer with high organic carbon content. "Fe^{2+}", the product of Fe(III) reduction, was positively correlated with bivalent cations such as "Sr^{2+}", "Ca^{2+}", and "Mg^{2+}". It was positively correlated with "%VC" as well, suggesting that VC accumulation was higher where Fe(III) reduction occurred. Overall, "Fe^{2+}" and "%VC" were following an almost similar pattern of correlations. The correlation between "Fe^{2+}" and "%cDCE" was not significant. The latter was strongly negatively correlated with "%TCE", indicating that the processes underlying the relative accumulation of these compounds were dependent of each other.

Conversely, "%VC" was not correlated with "cDCE", indicating that these processes were independent of each other. "PCE" was positively correlated with "Eh", reflecting the fact that the highest PCE concentrations were found in less reducing sections of the aquifer. All CEs were correlated with each other, the strongest correlations being displayed by adjacent compounds in the sequence of dechlorination.

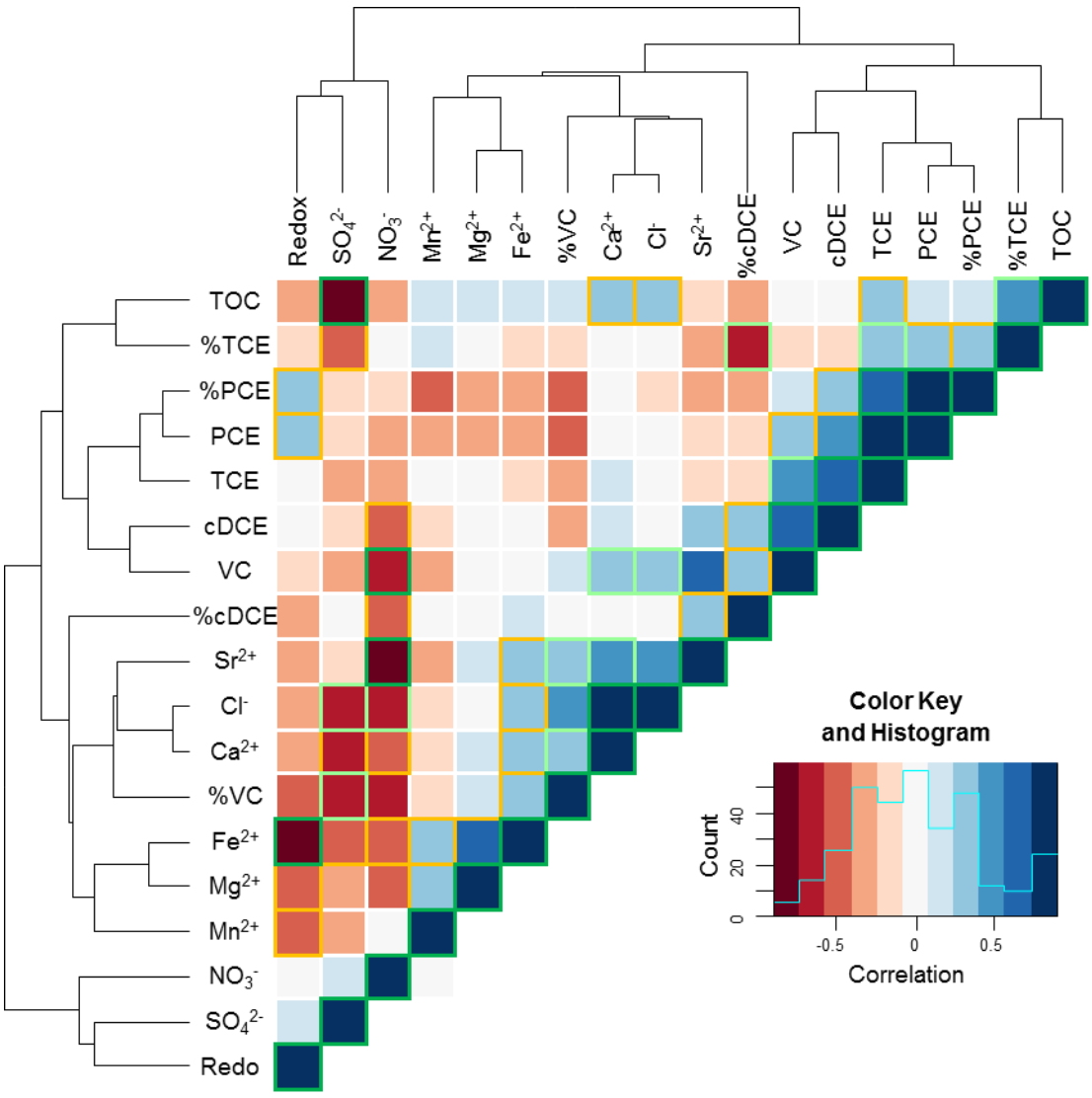


Figure 2. Heatmap of pairwise correlations between environmental variables. The color key of the correlations is shown on the left. Statistically significant correlations are indicated with a color surrounding the squares. Dark green: $p < 0.001$; light green: $p < 0.01$; orange: $p < 0.05$.

4.3.3 Bacterial communities

DNA extracted from October sample KB7b could not be amplified and was consequently removed from further analyses. In total, 30 samples provided bacterial community profiles (17 in April and 14 in October). The total number of T-RFs per sample was on average greater in October (80 T-RFs) than in April (62 T-RFs). Mean Shannon's H' diversity index (*50*) calculated based on the T-RFLP profiles was a little higher in

October (H=3.68) than in April (H=3.27) as well. Endpoint PCR detection of known OHRB led to positive results in most of the investigated wells (Table 2). *Desulfitobacterium* sp. was detected in all monitoring wells whereas *Dehalococcoides* sp. was detected in all but two wells (RB1/93 and KB7b). *Dehalobacter* was not detected in 6 monitoring wells (RB3/93, B2, KB4a, KB5b, KB6a, and KB6b), *Desulfuromonas* sp. in 5 (RB3/93, KB1, KB4b, KB6a, and KB7a), and *Sulfurospirillum* sp. was undetected in 3 (RB3/93, KB5a, and KB5b). Finally, the two genes *vcr*A and *bvc*A encoding VC reductive dehalogenases also provided positive results in most of the wells.

Furthermore, only the OHRB *Dehalococcoides* sp. and *Desulfitobacterium* sp. were detected in pyrosequencing datasets of 6 samples. The former was present in all but one sample (AB2) and the latter in two samples (BRB1/99 and BKB5b).

Table 2. PCR detections of known OHRB and genes encoding the two VC reductive dehalogenases known to date.

	***Dehalococcoides* sp.**	***Dehalobacter* sp.**	***Desulfitobacterium* sp.**	***Desulfuromonas* sp.**	***Sulfurospirillum multivorans***	***vcrA***	***bvcA***
RB1/93	+	+	+	+	+	+	+
RB1/99	+	+	+	+	+	+	+
RB3/93	-	-	+	-	-	-	-
B2	+	-	+	+/-	+	+	+
P20	+	+	+	+	+	+	+
KB1	+	+	+	-	+	+	+
KB2	+	+	+	+	+	+	+
KB3a	+	+	+	+	+	+	-
KB3b	+	+	+	+	+	+	+
KB4a	+	-	+	+	+	+	+
KB4b	+	+	+	-	+	+	+
KB5a	+	+	+	+	-	+	+
KB5b	+	-	+	+	-	+	+
KB6a	+	-	+	-	+	+	+
KB6b	+	-	+	+	+	+	+
KB7a	+	+	+	-	+	-	-
KB7b	-	+	+	+	+	-	+

4.3.4 Relationships between environmental variables and bacterial community structures

RV correlation coefficients and p-values calculated between the bacterial community dataset and environmental variables datasets are summarized in Table 3. An MFA ordination plot with environmental variables displayed in three groups, namely "%CEs", "TEA" and "OTHERS", is depicted in Figure 3. All three groups of variables were significantly correlated with BCS, with "TEA" showing the highest correlation.

Table 3. RV correlation coefficients calculated between the bacterial community dataset and (i) the whole environmental dataset (ENV), (ii) each defined group of environmental variables (%CEs, TEA, OTHERS), and (iii) each environmental variable. The p-values were calculated using a permutation test with 1000 permutations.

Variable or group of variables	RV coefficient	p-value	Significance[a]
ENV	0.62	<0.001	***
%CEs	0.34	0.002	**
TEA	0.54	<0.001	***
OTHERS	0.49	<0.001	***
PCE [nM]	0.24	0.004	**
PCE [%]	0.19	0.087	ns
TCE [nM]	0.29	<0.001	***
TCE [%]	0.18	0.108	ns
cDCE [nM]	0.22	0.017	*
cDCE [%]	0.19	0.063	ns
VC [nM]	0.34	<0.001	***
VC [%]	0.25	0.003	**
NO_3^-	0.20	0.026	*
Fe^{2+}	0.32	<0.001	***
Mn^{2+}	0.18	0.105	ns
SO_4^{2-}	0.26	0.002	**
TOC	0.24	0.005	**
Mg^{2+}	0.25	0.004	**
Ca^{2+}	0.28	<0.001	***
Sr^{2+}	0.28	<0.001	***
Na^+	0.17	0.147	ns
NH_4^+	0.18	0.098	ns
K^+	0.18	0.098	ns
Cl^-	0.23	0.010	*
Eh[b]	0.23	0.006	**
Temperature	0.26	0.002	**
EC[c]	0.21	0.025	*
pH	0.16	0.239	ns

[a] Degree of significance based on the calculated p-value: ***: highly significant; **: significant; *: moderately significant; ns: non-significant.
[b] Oxidation-reduction potential.
[c] Electrical conductivity

The two TEA variables with highest RV coefficients of 0.34 and 0.32, respectively, were "VC" and "Fe^{2+}". However, no significant correlation was obtained between these two variables (Spearman's rank correlation $\rho = 0.19$, $p = 0.31$; Figure 2). "TCE" was the third most correlated TEA variable (RV = 0.29, $p < 0.001$) followed by "SO_4^{2-}" and "PCE", displaying weaker correlations (RV=0.26 and 0.24, respectively) and higher p-values ($p = 0.004$ and 0.002, respectively). The correlations displayed by "NO_3^-" and "cDCE" with BCS were poorly significant, with p-values of 0.026 and 0.017, respectively. "Mn^{2+}" was not significantly correlated with BCS ($p = 0.11$).

Relative accumulation of individual CEs in a sample is reflected by their proportion among the cumulated concentrations of all CEs in this sample. Among the proportions of the different CEs, only "%VC" was significantly correlated with BCS (RV = 0.25, $p = 0.003$). Furthermore and in contrast to "VC", "%VC" was significantly correlated with "Fe^{2+}" (Spearman's rank correlation $\rho = 0.45$, $p = 0.012$, Figure 2). This result, together with the strong RV correlation coefficient displayed by "VC" and "Fe^{2+}" pointed out a possible competition between IR and VC reduction.

Among other ions, bivalent cations ("Mg^{2+}", "Ca^{2+}", and "Sr^{2+}") and "Cl^-" were significantly correlated with BCS. "Ca^{2+}" and "Sr^{2+}" even displayed a highly significant correlation (RV = 0.28, $p < 0.001$ for both variables). Temperature, redox potential and TOC were also significantly correlated, whereas pH displayed no significant correlation with BCS.

4.3.5 Correlations between T-RFs and environmental variables

T-RFs displaying significant correlations with axes MFA1 and MFA2 (Figure 3) were selected for further statistical treatment. Figure 4 shows the correlations calculated between each selected T-RF and selected environmental variables. These correlations were used for deducing potential implications of the T-RFs in the microbiological processes going on in the aquifer. Four T-RFs showed strong significant positive correlations with higher CEs (T-RFs 55, 192, 323, and 462). Six T-RFs displayed significant correlations with VC (T-RFs 165, 198, 210, 218, 264, and 270). A single T-RF (T-RF 243) displayed strong significant correlations with all CEs. Note that all the T-RFs present in these three categories showed no significant correlations with "Fe^{2+}". Two T-RFs showed a particularly strong negative correlation with "VC", namely T-RFs 214 and 223. A single T-RF (T-RF 219) was strongly positively correlated with "Fe^{2+}". Seven T-RFs (T-RFs 56, 182, 221, 256, 266, 289, and 292) displayed particularly strong significant negative correlations with "Fe^{2+}".

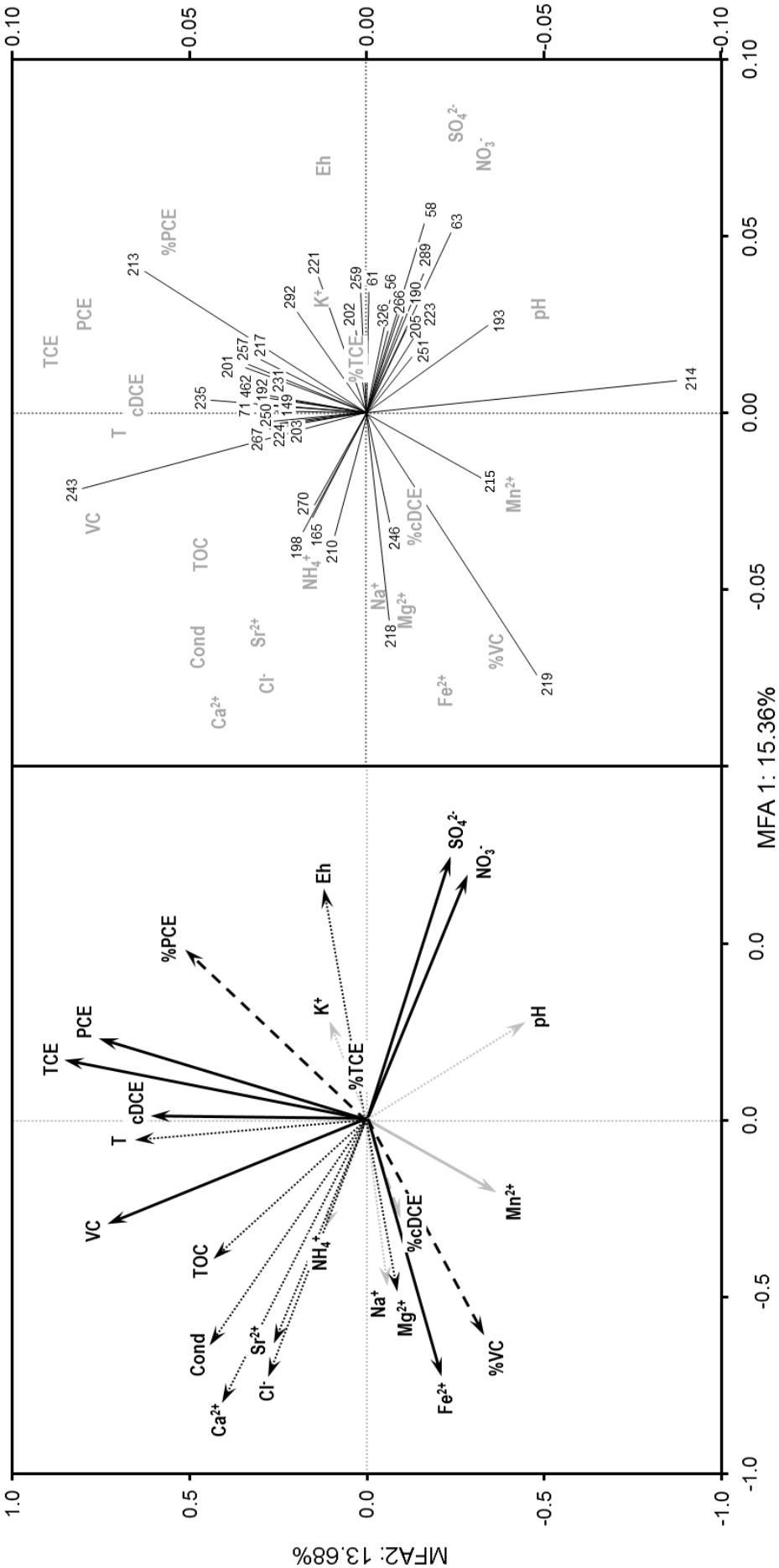


Figure 3. Multifactorial Analysis (MFA) carried out on environmental variables and bacterial communities showing the first two principal components (MFA1 and MFA2). The percentages indicated on the axes are the variance expressed by the principal component. Note that scales of the plots are adapted to facilitate readability. **A**: Significant and non-significant environmental variables are shown with black and grey arrows, respectively. Three groups of environmental variables are displayed: terminal electron-accepting processes (TEAPs, full arrows), relative proportions of chlorinated ethenes (%CEs, dashed arrows), and other environmental variables (OTHERS, dotted arrows). **B**: T-RFs displaying significant correlation with MFA axes 1 and 2 at the p = 0.05 level.

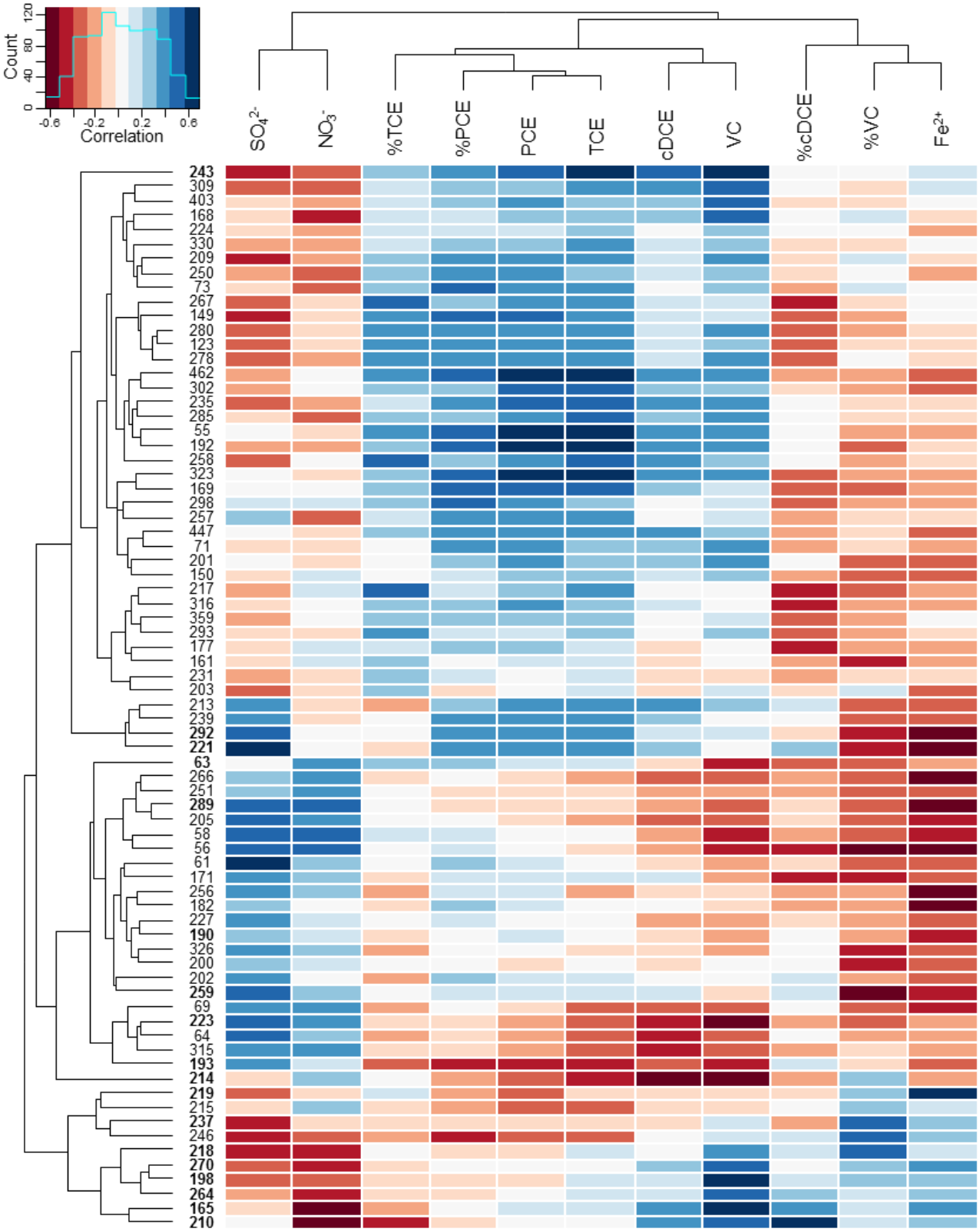


Figure 4. Heatmap of the correlations between selected environmental variables and T-RFs selected based on their correlations to the MFA first two principal components. The color key of the correlations is shown on the top left. T-RFs identified with PyroTRF-ID are displayed with bold characters.

4.3.6 Identification of T-RFs

Selection of T-RFs

Among the set of T-RFs selected above, a subset of 24 T-RFs were correlated at the p = 0.01 level with either "VC", "%VC", or "Fe^{2+}". These T-RFs were selected for identification (Table 4), as the correlations with these variables indicated that they were possibly involved in the competition between IR- and VC reductions.

Pyrosequencing T-RFs Identification (PyroTRF-ID)

Six samples (AB2, AKB2, AKB6b, BKB1, BKB5b, and RB1/99) were selected for in-depth analysis of the 16S rRNA gene pools by pyrosequencing. The selection among all samples was carried out based on the clustering of their respective experimental T-RFLP (eT-RFLP) data sets, and one representative sample per cluster was selected (data not shown). As expected, the results showed that digital T-RFs (dT-RFs) were built predominantly with the contributions of several 16S rRNA gene sequences, showing different taxonomic levels (Table 4). Use of the PyroTRF-ID software enabled identifying the phylotypes contributing to some of the T-RFs that were significantly correlated with selected environmental variables. Seven T-RFs with low relative abundance in the selected samples could not be identified, since they found no corresponding dT-RFs.

Among the T-RFs identified with PyroTRF-ID that were positively correlated with VC, 3 were unambiguously affiliated to the *Dehalococcoidaceae*. The T-RF 165 was clearly affiliated to *Dehalococcoides* sp., with 100% of all contributing reads. The affiliation of T-RFs 210 and 243 stopped at the family level. Their contributing reads were processed in BLAST against the NCBI database to gain additional information. The results indicated that T-RFs 210 and 243 were affiliated to sequences of the "Lahn Cluster" (LC). The LC is in fact composed of two subclusters, namely LC1 and LC2 (*51*). The T-RFs 243 and 210 displayed 97 and 99% sequence identity with LC1 and LC2, respectively. The T-RF 198 was unambiguously affiliated to *Desulfovibrio* sp., a SRB. Among the T-RFs that were negatively correlated with VC, T-RF 63 was affiliated to the *Methylococcaceae*, T-RF 214 to the *Comamonadaceae*, and T-RF 216 to the *Gallionellaceae*, composed of IRB. T-RFs 218 and 237, both displaying a positive correlation with %VC, and T-RF 219, displaying a positive correlation with Fe^{2+}, were all affiliated to the *Gallionellaceae* as well. Among the T-RFs that were negatively correlated with %VC, unambiguous identification was obtained for T-RF 289, affiliated to the family *Sphingomonadaceae*, T-RF 292, affiliated to the *Rhodospirillaceae*, and T-RF 326, affiliated to *Nitrospira* sp. The T-RF 190, negatively correlated with Fe^{2+}, was affiliated to the *Desulfobulbaceae*. For more details, please refer to Table 4.

Table 4. Identities of selected T-RFs obtained with PyroTRF-ID. Only those T-RFs that were successfully identified are shown.

eT-RF[a] (bp)	Correlation[b] with: "VC"	"%VC"	"Fe^{2+}"	Sample	Relative abundance of eT-RF in sample[c]	Corresponding dT-RF[d] (bp)	Dominant phylotype[e]	Normalized SW score[f]	Clone identification[g]
165	+	ns	+	BKB1	1.5%	163	*Dehalococcoides* sp. (100%; 143/143)	0.71-0.95	*Dehalococcoides* sp. (2)
198	+	ns	ns	AKB6b	4.4%	196	*Desulfovibrio* sp. (98.6%; 140/142)	0.55-1.0	*Desulfovibrio* sp. (2)
210	+	ns	ns	BKB1	3.4%	209	*Dehalococcoidaceae* sp. (98.3%; 233/237)	0.71	*Dehaloccoidetes* Lahn Cluster (2)
243	+	ns	ns	BRB1-99	10.0%	243	*Dehalococcoidaceae* (99.7%; 389/390)	0.45-0.63	*Dehaloccoidetes* Lahn Cluster (3)
270	+	ns	+	BRB1-99	0.4%	271	*Synergistaceae* (77.8%; 7/9)	0.73	na
63	-	ns	ns	AB2	3.3%	64	*Methylococcales* (98.2%; 107/111)	0.69-0.92	na
							Methylococcaceae (85.3%; 93/111)	0.68-0.91	na
214	-	ns	ns	AKB6b	3.2%	214	*Comamonadaceae* (81.5%; 75/92)	0.74-0.97	*Rhodoferax* sp. (4)
223	-	-	ns	BKB5b	1.3%	223	*Actinomycetales* (41.8%; 41/122)	0.51-0.92	na
218	ns	+	ns	AB2	15.7%	217	*Gallionellaceae* (97.3%; 248/255)	0.59-0.86	na
237	ns	+	ns	AKB6b	2.3%	238	*Gallionellaceae* (98.6%; 610/619)	0.49-0.89	na
221	ns	-	-	BKB5b	2.4%	221	*Comamonadaceae* (21.4%; 3/14)	0.60-0.81	na
259	ns	-	-	BKB5b	3.3%	259	*Solibacterales* (97.4%; 38/39)	0.49-0.93	na
289	ns	-	-	BKB5b	3.1%	287	*Sphingomonadaceae* (100%; 123/123)	0.77-0.98	na
							Kaistobacter sp. (33.8%106/314)	0.77-0.98	na
292	ns	-	-	BKB1	6.8%	290	*Rhodospirillaceae* (100%; 314/314)	0.46-0.79	na
219	ns	ns	+	AB2	22.6%	217	*Gallionellaceae* (97.3%; 248/255)	0.67-0.88	na
190	ns	ns	-	BKB5b	3.3%	191	*Desulfobulbaceae* (54.6%; 12/22)	0.95	na

[a] Experimental Terminal-Restriction Fragment (eT-RF).
[b] Spearman's rank correlation of the eT-RF with 3 variables informative of the competition between Fe(III)- and VC reduction. +: positive correlation; -: negative correlation; ns: non-significant correlation.
[c] Relative abundance (%) of the eT-RF in the sample used for PyroTRF-ID computation.
[d] Digital Terminal-Restriction Fragment corresponding to the eT-RF.
[e] Main phylotype contributing to the dT-RF. In brackets: relative contribution of the phylotype to the dT-RF; number of reads showing this affiliation among the total number of reads contributing to the dT-RF.
[f] Range of Smith-Watermann mapping scores of reads contributing to the dT-RF normalized by the sequence length.
[g] T-RF identity obtained by cloning-sequencing. The number of clones with this identity is indicated in brackets.

Cloning and sequencing

In total, 105 clones were sequenced. Among the retrieved sequences, 67 displayed the required features for further analyzes, i.e. the forward primer could be detected, they were not identified as chimeras, the first restriction site of the enzyme *Hae*III was present, and the resulting dT-RF corresponded to the eT-RF within a maximal 10-bp drift. Among these, 16 belonged to the list of selected T-RFs. All identified T-RFs were positioned in phylogenetic trees of bacterial 16S rRNA sequences (Figures S1 to S4). More than one phylogenetic affiliation was found for 13 T-RFs (T-RFs 63, 183, 193, 194, 197, 198, 210, 213, 214, 216, 218, 221, and 255).

Chloroflexi

Sixteen sequences, related to 9 different T-RFs were affiliated to the *Chloroflexi* phylum (Figure S1). Two sequences were obtained for T-RF 165, which displayed the strongest correlation with VC, and no significant correlation with higher CEs (Figure 4). These sequences were both affiliated to *Dehalococcoides* sp. (with 100% sequence identity), confirming the identity obtained with PyroTRF-ID. Two T-RFs were affiliated to a deep-branching cluster within the *Dehalococcoidetes* class. They shared respectively 100% (T-RF 210, clone z31), 99% (T-RF 243, clones z60 to z62), and 97% (T-RF 210, clone z32) sequence identity with sequences from the LC. Confirming the results of PyroTRF-ID, T-RF 210 was affiliated to LC2, and T-RF 243 to LC1.

Two other T-RFs (T-RFs 213 and 255a) were affiliated to the *Dehalococcoidetes* class, and three to the *Anaerolineae* class (T-RFs 221, 255, 256). T-RF 220 shared 91% sequence identity with an unidentified *Chloroflexi* clone from a CE-contaminated aquifer (GenBank accession number DQ223083).

β-Proteobacteria

Twenty-six sequences were affiliated to the *β-Proteobacteria* (Figure S2) and were related to 11 different T-RFs. Sixteen sequences were affiliated to the *Comamonadaceae* and were related to a total of 9 different T-RFs. Among the latter, T-RF 214 and T-RF 194 were affiliated to *Rhodoferax* sp. T-RFs 218 (3 different sequences) and 219 were affiliated to members of the *Gallionellaceae*, 2 T-RFs (T-RFs 197 and 216) were affiliated to the *Methylophilaceae*, and 2 (T-RFs 193 and 194) to the *Rhodocyclaceae*. Finally, one sequence related to T-RF 213 was affiliated to a group of unclassified *β-Proteobacteria*.

δ-Proteobacteria

Two almost identical sequences were related to T-RF 198 and were affiliated to the *Desulfovibrio* genus (Figure S3). T-RF 198 shared 100% sequence identity with clone DCE36 (GenBank accession number AJ249261), isolated from a CE-dechlorinating enrichment culture, and 99% identity with a clone isolated from an anaerobic methane-producing enrichment culture (*52*).

α-, δ-, and γ-Proteobacteria, and Firmicutes

T-RF 63 and 64 were classified in the methanotrophic *Methylococcaceae* within the *γ-Proteobacteria* (Figure S3). T-RF 495 shared 98% sequence identity with a clone obtained from an arsenite-oxidizing biofilm (GenBank accession number AY168736) and 93% sequence identity with *Bdellovibrio bacteriovorus* str. H100 (GenBank accession number BX842648). The phylum *Firmicutes* and the class of the *α-Proteobacteria* were each represented by a single clone.

Others / unclassified

The phylum *Nitrospirae* and the candidate divisions OP11, TM7, and NC10 were each represented by a single sequence (Figure S4). Thirteen sequences related to 11 different T-RFs could not be classified in any known bacterial taxon.

4.4 Discussion

4.4.1 Incomplete CE dechlorination cannot be explained by the lack of electron donors and the absence of biochemical potential

As mentioned above, a major drawback of MNA of CE-contaminated aquifers is the frequent accumulation of undesired intermediates as a result of an incomplete dechlorination, or of a decreased efficiency of lower CEs dechlorination (*5*). The explicit and unambiguous explanation of the reasons for which this natural mechanism is not continuing until the release of the final products (i.e. ethene or ethane) is therefore mandatory. Incomplete or impeded reductive dechlorination is often the result of a depletion of organic substrates metabolized by the fermenting guild, which in turn provides OHRB with suitable electron donors (*53, 54*). Based on geochemical analysis of core samples collected in a previous campaign (data not shown), the Zuchwil aquifer was shown to be rich in organic carbon (estimated between 6 - 9% (w/w)). Therefore, lack of carbon and electron sources fuelling the OHR bacterial guild was therefore excluded as a possible explanation for incomplete CE dechlorination. The presence of OHRB such as *Sulfurospirillum sp.*, *Desulfitobacterium* sp., *Dehalobacter* sp., *Desulfuromonas* sp., and *Dehalococcoides* sp. was confirmed in most wells of the Zuchwil aquifer. End-point PCR carried out using specific primers also revealed the presence of *vcr*A and *bvc*A, two genes encoding VC reductases (*55*). Based on pyrosequencing results, only 2 OHRB were detected in the Zuchwil aquifer: *Dehalococcoides* sp. was present in all analyzed samples but AB2, whereas *Desulfitobacterium* sp. was only present in BRB1-99 and BKB5b. The discrepancy between end-point PCR and pyrosequencing probably arose from the greater sensitivity of the nested PCR approach. Nevertheless, these observations, together with the presence of small amounts of ethene, clearly indicated that the biochemical potential for CEs dechlorination was present in the Zuchwil aquifer and that bacteria responsible for lower CEs reduction were present and active locally, but were unable to fully exploit their catabolic potential throughout the whole aquifer.

4.4.2 TEAPs shape the whole bacterial community present

The multivariate statistical analysis of the BCS data and the environmental variables showed that the group composed of terminal electron acceptor variables ("TEA") displayed the highest correlations with the BCS. In general, the energetically most favorable electron-accepting process predominates in the habitat as long as the concentration of the required TEA is sufficient. When the TEA is depleted or its concentration too low, the next most favorable TEAP takes place. This phenomenon gives rise to the ecological succession of TEAPs (*56*). Microbial populations involved in anaerobic TEAPs make often use of electron sources provided by upstream processes such as fermentation and syntrophy (*11*). These populations can be referred to as keystone populations (*57*). Their abundance may be limited, but their respiratory function influences the structure and the functioning of the microbial community as a whole, through the transmission of the electron fluxes across the successive bacterial guilds. This influence of TEAPs over the entire bacterial community is reflected here by the strong correlation between TEAs and the BCS.

"%CE", an indicator of the reduction state of the contaminants, was the second most correlated set of variables. T-RFLP profiles thus showed an important adequacy with the different stages of reduction of CEs. This significant correlation demonstrated that BCS were not only conditioned by the presence of chlorinated compounds, but also that they presented an adaptive response to their relative proportions, and more particularly the proportion of VC which was the most strongly correlated with BCS. The discrepancies observed in the T-RFLP profiles were not induced necessarily by major changes in the OHR guild itself, as the diversity thereof was quite limited to a few populations only, as shown by metagenomic analysis. Classically, OHR of cDCE and VC occurs under more reducing conditions than required for OHR of PCE and TCE (*58*). In anoxic environments, OHR of cDCE and VC is therefore more susceptible of being hampered. The degradation of both higher and lower CEs is known to involve the activity of different bacterial populations. Maillard et al. (2011) showed that *Sulfurospirillum* sp. was involved in the reduction of PCE and TCE to cDCE, whereas *Dehalococcoides* sp. was responsible for the dechlorination of cDCE and VC to Eth (*59*). In contrast, Flynn et al. (2000) showed that T-RFs appearing or disappearing after the switch from higher to lower CE dechlorination were not belonging to known dehalogenating bacteria, but to a nitrogen-fixing *β-Proteobacterium* and to members of the *Clostridiaceae*, respectively, which were present in the consortia as accompanying populations (*60*). When the flux of electrons switches from higher CEs (PCE and TCE) to the lower ones (cDCE and VC), the thermodynamics and kinetics of the overall reaction, i.e. from the organic substrate fermentation to the reduction of the TEA, is modified (*61*). This change undoubtedly selects for the guilds that are best adapted to the new conditions.

CEs displayed contrasting RV correlation coefficients. "PCE" and "TCE" were significantly correlated with BCS. The reduction process of these two compounds was expected to be favored under the IR conditions prevailing in the Zuchwil aquifer, allowing for the development of a dedicated OHRB population. These TEAPs are indeed thermodynamically more favorable than IR, and their H_2 threshold concentrations requirements are in the same range (*62, 63*). "cDCE" was slightly correlated with BCS, whereas "VC"

displayed the highest significant correlation with BCS among all measured environmental variables. "VC" can be an indicator either of cDCE or VC reduction. Whatever the case, the variations in VC concentrations were directly related to the reduction of lower CEs and to bacterial communities associated with one or the other processes it may indicate.

4.4.3 CEs were dechlorinated by several *Chloroflexi*

Interestingly, and at present knowledge, cDCE and VC can be reduced by some strains of *Dehalococcoides* only, involving deeper redox conditions. Five T-RFs (165, 198, 210, 264 and 270) displayed a strong correlation with lower CEs as well as no significant correlation with higher ones. PyroTRF-ID and the sequences of two clones confirmed that T-RF 165 was closely affiliated to *Dehalococcoides* sp. This confirmed the implication of this obligate OHRB in the reduction of lower CEs, which was further demonstrated by the strong positive correlation shown between T-RF 165 and "VC". T-RF 210 was affiliated to the sequences belonging to the LC (*51*) based on the results of PyroTRF-ID and confirmed by two clone sequences. Uncultured bacteria belonging to the LC are members of the *Chloroflexi* but they are only distantly related to *Dehalococcoides* sp., with only 92 to 94% sequence identity. They were detected in pristine river sediment microcosms incubated at 15°C with PCE as sole electron acceptor and ^{13}C-labeled acetate as electron donor. Based on stable isotope probing (SIP) analysis, they were shown to reductively dechlorinate PCE to cDCE. In the present study, they were detected for the first time in groundwater DNA extracts. According to T-RFLP profiles, this population was dominant in some sampling wells (T-RF 243 represented more than 10% of the bacterial community in well RB1/99), suggesting that bacteria of the LC are putative major protagonists in particular PCE-contaminated groundwater systems. The significant correlation between T-RF 210 and lower CEs and the absence of correlation with higher CEs indicated that it was directly involved in the reduction of lower CEs. Should this finding be confirmed, it would be the first observation of a member of the LC able to dechlorinate lower CEs. T-RF 243 was unambiguously related to the LC as well, as confirmed by both PyroTRF-ID and cloning-sequencing approaches. In contrast with T-RF 210, it was significantly strongly correlated with all CEs, indicating that it was involved in the reductive dechlorination of both higher and lower CEs. The large proportion represented by this T-RF in the bacterial communities of several samples (up to 14.8%) indicated that it was a major protagonist in the reduction of CEs in the Zuchwil aquifer. Two sequences among the four obtained for T-RF 198 were related to a SRB *Desulfovibrio* sp. clone detected in drinking water. *Desulfovibrio* sp. has been reported as a facultative SRB and former studies have also demonstrated the syntrophic role of members of this genus (*64–66*). Men et al. (2011) showed that the dechlorination rate was enhanced in a co-culture containing *Dehalococcoides ethenogenes* str. 195 and a strain of *Desulfovibrio*, as compared to the culture of the OHRB alone (*65*). In the Zuchwil aquifer, *Desulfovibrio* sp. was not correlated with "SO_4^{2-}", indicating that it played a non-significant role in the reduction of sulfate. However, this T-RF was present in all samples and was co-occurring (data not shown) with T-RF 165 (*Dehalococcoides* sp.), supporting the hypothesis of a syntrophic activity. Three clone sequences corresponding to the T-RF 218 were closely affiliated to members of the *Gallionellaceae*.

PyroTRF-ID analysis showed equally that 97.3% of the pyrosequencing sequences obtained from sample AB2 and showing a T-RF of this size were associated to this T-RF. However, the reasons for which members of this family were significantly correlated with "%VC" could not be explained at this stage. Finally, no phylogenetic affiliation could be attributed to T-RF 270, as pyrosequencing provided no dT-RF related to this size.

4.4.4 Iron reduction is another main TEAP in the Zuchwil aquifer

The value of the RV correlation coefficient computed between "Fe^{2+}" and BCS was almost identical as the one computed between "VC" and BCS. Relatively important concentrations of Fe^{2+} in the Zuchwil aquifer unambiguously indicated that IR was a major TEAP in the Zuchwil aquifer (*56*). This was supported by the significant positive correlations computed between "Fe^{2+}" and bivalent cations such as Mg^{2+} and Sr^{2+}, as it has been shown previously that metals and bivalent ions that have been co-precipitated with Fe(III) are being solubilized together with Fe^{2+} during IR (*67*, *68*). IR was also consistent with results of geochemical analyzes carried out on core samples collected in a previous campaign that revealed overall high Fe(III) contents (between 1.6 and 4.4 % w/w) in the aquifer matrix (data not shown). Among all T-RFs, T-RF 214 showed a distinctive pattern of correlations, displaying significant negative correlations with all CEs as well as with "SO_4^{2-}" (Figure 4). Furthermore, T-RF 214 showed a positive correlation with "%VC". At the same time, this T-RF was strongly negatively correlated with T-RF 243, affiliated to the LC. The results of these correlations indicated that the sequences composing this T-RF (or at least a fraction of it) were related to a bacterial taxon implicated in an activity which was incompatible with the reduction of lower CEs. PyroTRF-ID showed that this T-RF was predominantly composed of sequences affiliated to the *Comamonadaceae* (81.5% relative contribution to the dT-RF) at the family level. However, nothing could be concluded at the genus level with this technique, as too many different phylotypes contributed to the dT-RF. However, the cloning-sequencing approach provided 4 clones among five which were closely affiliated to *Rhodoferax* sp., a genus containing known IRB (*69*, *70*). T-RF 219 was the most strongly correlated T-RF with "Fe^{2+}", indicating a link with the iron metabolism. Its identification both by cloning-sequencing and by PyroTRF-ID unambiguously revealed an affiliation to members of the iron-oxidizing *Gallionellaceae*. Aerobic iron oxidation was not a major metabolism in the Zuchwil aquifer as this metabolism is known to occur in microaerophilic conditions only (*71*), and since the Zuchwil aquifer was mainly under IR conditions. Nevertheless, several locations of the aquifer were undeniably under microaerophilic conditions. This was the case for the upstream zone of the plume, where oxic water was penetrating the area. Also, water present in wells was in contact with atmospheric oxygen. A reddish-brownish organic precipitate was observed in some wells, characteristic of iron precipitates formed by iron-oxidizing bacteria. Thus, iron-oxidizing bacteria observed in our samples may originate either from a limited microaerophilic zone around the well formed by O_2 diffusion from the well, or directly from water present in the well. In both cases, they would represent a sampling artifact in the sense that the microaerophilic conditions result from the presence of the monitoring well.

It has to be mentioned at this stage that although several T-RFs could be successfully identified based on cloning-sequencing and PyroTRF-ID in this study, some uncertainties still prevailed for some of them. Indeed, several phylotypes may contribute to a single T-RF (*45*, *72*, *73*). Conversely, several T-RFs may refer to a single phylotype. For instance, the genus *Desulfitobacterium* contains several different 16S rRNA genes (*74*), with restriction sites located at different locations and resulting in different T-RFs. In our case, 7 T-RFs could not be resolved at all because they found no corresponding dT-RF, mainly due to the low relative abundance of these fragments in the investigated T-RFLP profile. Furthermore, as explained above, some dT-RFs could not be elucidated at the genus level, since they were composed of too many phylotypes with no clear dominance.

4.4.5 Competition between VC reduction and iron reduction in the Zuchwil aquifer

Despite the statistical evidence that IR and the reduction of lower CEs were both prominent TEAPs, no correlation was observed between "Fe^{2+}" and "VC". A significant positive correlation was however detected between "Fe^{2+}" and "%VC", an indicator of the extent of CE reduction and of the relative accumulation of VC. The positive correlation emphasized the competitive nature of the interaction between the two TEAPs for electron donors. The competition between IRB- and VC-reducing bacteria is therefore the most likely hypothesis explaining the accumulation of VC in the Zuchwil aquifer. In anoxic groundwater, most TEAPs mainly rely on H_2 as electron donor for their respiratory metabolism. H_2 is produced by a series of processes involving fermenting bacteria and syntrophs. It is preferentially consumed by the thermodynamically more potent TEAPs. The bacteria mediating these processes lower H_2 concentration towards a physiologically-based H_2 concentration threshold, making it unavailable to less potent TEAPs (*11*). As a consequence, H_2 threshold concentration requirements of bacteria mediating TEAPs in groundwater, together with kinetic aspects, determine their competitiveness. To the best of our knowledge, few studies specifically investigated the competition of these two TEAPs or even the potential inhibition of the reduction of lower CEs by IRB. Wei et al. showed that IR was not necessarily inhibiting the complete reduction of CEs to ethene (*16*). Conversely, several studies showed that OHR requires H_2 threshold concentrations similar to, or higher than other TEAPs, including IR (reviewed in (*11*)). In particular, one study measured H_2 steady-state concentrations during cDCE and VC reduction in batch cultures and found that they were higher than for IR (*63*). The hypothesis that VC-reducing bacteria were outcompeted by IRB in the Zuchwil aquifer was therefore consistent with literature data and fitted with the interpretation of statistical results.

4.4.6 Sulfate reduction potentially helps VC to be dechlorinated

"SO_4^{2-}" displayed a significant correlation with BCS. Concentrations of sulfate were highest in well RB3/93, which was under nitrate-reducing conditions, as indicated by the nitrate concentrations and the calculated redox state. In the other locations of the aquifer, sulfate concentrations were lower, indicating that SR was an active process, although globally not influencing the BCS. Bethke et al. (2011) showed that IR and SR were not necessarily excluding each other, but that a mutualistic relationship could exist (*75*). SR and IR are inhibited by the accumulation of their respective reduction products, namely S^{2-} and Fe^{2+}. The precipitation of these products into FeS(s) was proposed as the mechanism leading to the observed mutualism (*61*, *76*). The relationship between SRB and OHRB in groundwater is still controversial. Several studies carried out on dechlorinating consortia reported the inhibition of OHRB by SRB (*14*, *77*, *78*). Contrarily, other studies provided some evidence that the reduction of lower CEs can compete with SRB (*15*, *79*, *80*). Measured threshold and steady-state H_2 concentrations (*11*) suggested that OHRB should be able to compete with SRB. In the Zuchwil aquifer, sulfate concentrations were the lowest in the zones where complete dechlorination was highest (*17*), as shown by the significant negative correlation between "SO_4^{2-}" and "VC". In other words, VC reduction occurred preferentially in micro-habitats where SR was enhanced, leading to a noticeable decrease of sulfate concentration. At the same time, "SO_4^{2-}" and "TOC" were strongly significantly anti-correlated, indicating that SR was enhanced in zones with high organic matter content. "SO_4^{2-}" and "Fe^{2+}" were equally negatively correlated. This correlation could be interpreted as the consequence of the precipitation of Fe^{2+} in the form of FeS, which in turn explains why Fe^{2+} was not accumulating in excess in these zones as compared with other sections of the aquifer. This hypothesis was supported by the undetected or very low concentrations of sulfide measured in the investigated wells (*17*), as S^{2-} produced by SR has probably systematically reacted with Fe^{2+} in solution (*75*). Assuming that the reduction of lower CEs was predominantly taking place in zones of low SO_4^{2-} concentrations, RV correlation coefficients, combined with the identification of T-RFs, provided indications on the bacterial taxa involved in this process. T-RF 243, affiliated to the LC, was the only T-RF identified as an OHRB which displayed a significant negative correlation with "SO_4^{2-}", indicating that it was particularly abundant in zones with low SO_4^{2-} concentrations, where VC was reduced to ethene. It was therefore considered as the main driver of the reduction of lower CEs in the Zuchwil aquifer. Surprisingly, T-RF 165 displayed no correlation with "SO_4^{2-}", indicating that *Dehalococcoides* sp., although undoubtedly implicated in the reduction of CEs throughout the Zuchwil aquifer, was not restricted to the zones with low SO_4^{2-} concentrations.

4.5 Conclusion

The integrative approach used in this study allowed the global assessment of the ecological mechanisms leading to the accumulation of the intermediate toxic compound VC. Statistical analyzes showed that IR and VC reduction were the main TEAPs involved in the Zuchwil aquifer and suggested that IRB and VC-reducing bacteria were competing for electron donors. The identification of T-RFs provided further details on the bacteria involved in this competition, indicating that the most significant IRB were composed of *Rhodoferax*-like bacteria or other species belonging to the *Comamonadaceae*. VC-reducing bacteria were unsurprisingly composed of *Dehalococcoides* sp. More surprising was the high abundance and strong correlation with VC of sequences related to the LC. To the best of our knowledge, this is the first report indicating that members of this cluster could be involved in the reduction of lower CEs, a result that will have to be confirmed using enrichment cultures. Finally, our results suggest that SRB could play a role in the attenuation of the contaminated site, by stimulating the depletion of Fe(III) through mutualism with IRB.

Acknowledgements

I warmly thank Christoph Aeppli (Woods Hole Oceanographic Institution, MA, USA) for the analysis of chlorinated and non-chlorinated ethenes, as well as for providing precious information on the contaminated site. I cheerfully thank Timothy M. Vogel, Sébastien Cécillon, and Tom O. Delmont (Environmental Microbial Genomics Group, Ecole Centrale de Lyon, France) for their kind collaboration in pyrosequencing analysis and data processing. I thank Emmanuelle Rohrbach, David Weissbrodt, Yoan Rappaz (Laboratory for Environmental Biotechnology, EPFL, Lausanne, Switzerland), Céline Ruegsegger (Institute of Cell Biology, University of Bern, Switzerland), and Claudia Pasche (CEL, EPFL, Switzerland) for their precious technical assistance in sampling and in laboratory analyses, Ioannis Xenarios for support and access to the Vital-IT high performance computing cluster of the Swiss Institute for Bioinformatics (Lausanne, Switzerland), Johannes Dollinger (SolGeo AG, Solothurn, Switzerland) for the access to the contaminated site, and François Gillet (Université de Franche-Comté, France) for his kind advices in multivariate statistics.

Supplementary material

Table S1. General water characteristics measured in samples of the Zuchwil aquifer in April 2008.

Monitoring well	Screen depth [mbs]	Water level [mbs]	T°C [°C]	pH	EC[a] [µS cm^{-1}]	Eh[b] [mV]	TOC[c] [mg l^{-1}]	Fe^{2+} [mg l^{-1}]	Mn^{2+} [mg l^{-1}]	Na^{+} [mg l^{-1}]	NH_4^{+} [mg l^{-1}]	K^{+} [mg l^{-1}]	Mg^{2+} [mg l^{-1}]	Ca^{2+} [mg l^{-1}]	Sr^{2+} [µg l^{-1}]	Cl^{-} [mg l^{-1}]	NO_3^{-} [mg l^{-1}]	SO_4^{2-} [mg l^{-1}]
RB1/93	6-8	1.8	13.5	6.9	619	66	3.2	4.2	0.3	17.1	0.5	1.0	11.4	140.1	395.4	42.0	0.2	9.2
RB1/99	3-9	2.3	15.8	6.8	831	-23	6.2	4.3	0.2	23.8	0.9	1.0	12.1	185.8	486.8	92.3	0.2	1.3
RB3/93	5-9.5	5.1	11.2	7.0	515	78	1.2	0.1	<0.02	13.4	<0.2	1.6	7.8	45.8	266.3	7.0	5.5	19.9
B2	6-8	2.7	12.9	6.7	890	-46	4.5	8.4	0.3	20.0	0.6	0.4	13.9	213.3	563.0	131.1	0.2	4.2
P20	6-8	1.0	12.3	6.8	551	-55	3.7	10.9	0.9	13.6	2.7	0.4	12.3	131.4	395.8	28.0	0.2	0.4
KB1	9-11	2.3	12.3	7.0	594	71	2.0	7.6	0.6	21.8	<0.2	1.7	15.9	130.1	589.1	40.7	0.1	9.8
KB2	9-11	2.8	12.6	6.9	770	-136	3.4	9.8	0.4	17.3	0.4	0.9	13.7	178.6	565.9	78.2	0.1	4.4
KB3a	6-9	2.7	12.5	6.8	727	-37	3.5	5.3	0.3	30.6	0.9	0.4	8.8	171.5	499.8	87.1	0.2	1.8
KB3b	10.5-12.5	2.7	13.7	7.0	742	-54	2.4	17.6	0.5	31.0	0.6	2.2	15.3	155.1	752.4	65.8	0.1	7.7
KB4a	6-8	2.5	12.3	6.9	641	-143	4.9	9.6	1.9	30.6	1.7	1.6	10.8	131.6	457.8	63.2	0.2	2.5
KB4b	9-12	2.4	13.0	6.8	801	-88	4.5	12.9	0.6	17.5	0.7	1.2	13.5	192.7	604.8	87.7	0.1	1.9
KB5a	5.5-8.5	1.3	14.0	6.8	658	-15	4.6	3.6	1.5	20.3	1.3	1.5	12.2	147.3	473.4	44.3	0.5	2.6
KB5b	10-13	1.2	13.4	6.7	652	27	3.0	0.8	0.6	23.2	1.0	2.2	8.6	150.6	468.0	48.9	0.7	11.0
KB6a	6-9	2.2	13.1	6.8	695	-51	3.7	3.7	0.4	15.8	1.0	0.5	12.3	139.1	459.5	23.1	0.2	0.2
KB6b	10-12	2.0	12.8	6.8	568	-42	3.2	11.4	0.4	16.0	0.3	1.3	15.7	169.0	585.7	76.8	0.2	0.8
KB7a	6-8	3.0	12.6	6.8	891	-56	3.3	18.7	1.6	30.1	1.7	1.1	14.8	199.9	646.2	148.0	0.3	3.2
KB7b	10-12	2.9	13.0	7.1	416	-66	2.3	4.2	1.0	13.6	<0.2	0.7	10.5	91.4	388.7	14.4	0.7	8.8

[a] Electrical conductivity.
[b] Oxidation-reduction potential.
[c] Total organic carbon.

Table S2. Chlorinated ethenes measured in the aquifer in April 2008 and their relative proportions.

Monitoring well	PCE [µg l^{-1}]	TCE [µg l^{-1}]	cDCE [µg l^{-1}]	VC [µg l^{-1}]	Total CEs [µg l^{-1}]	%PCE	%TCE	%cDCE	%VC
RB1/93	26249.9	7224.0	5050.0	26.0	38549.9	59.6	20.7	19.6	0.2
RB1/99	31621.6	19289.3	1122.0	360.0	52392.9	53.7	41.4	3.3	1.6
RB3/93	77.1	12.8	18.9	0.9	109.7	60.2	12.6	25.3	1.9
B2	26.9	33.5	39.5	21.4	121.3	13.9	21.8	34.9	29.4
P20	1804.0	1308.0	1244.0	10.0	4366.0	32.2	29.4	37.9	0.5
KB1	3.6	46.6	2704.0	86.9	2841.1	0.1	1.2	94	4.7
KB2	1797.0	1854.0	2265.0	157.0	6073.0	21.3	27.8	46	4.9
KB3a	<0.5	<10	25.6	29.0	54.6	0	8.8	33.1	58.1
KB3b	<0.5	22.3	2686.0	193.6	2901.9	0	0.5	89.5	10
KB4a	<0.5	47.0	63.4	31.7	142.1	0	23.6	43	33.4
KB4b	<0.5	148.8	4972.0	234.8	5355.6	0	2	91.3	6.7
KB5a	2324.0	3385.0	3624.0	202.0	9535.0	17.4	32	46.5	4
KB5b	1005.0	2885.0	7532.0	90.0	11512.0	5.7	20.5	72.5	1.3
KB6a	<0.5	<10	3.1	2.3	5.4	0	45.8	25.3	28.8
KB6b	<0.5	36.9	362.8	333.4	733.1	0	3	40	57
KB7a	<0.5	18.4	4.1	3.0	25.5	0	60.6	18.5	20.9
KB7b	<0.5	31.9	2.2	<0.5	34.1	0	91.6	8.4	20.9

Table S3. General water characteristics measured in the aquifer in October 2008.

Monitoring well	Screen depth [mbs]	Water level [mbs]	Temp. [°C]	pH	EC[a] [µS cm^{-1}]	Eh[b] [mV]	TOC[c] [mg l^{-1}]	Fe^{2+} [mg l^{-1}]	Mn^{2+} [mg l^{-1}]	Na^{+} [mg l^{-1}]	NH_4^{+} [mg l^{-1}]	K^{+} [mg l^{-1}]	Mg^{2+} [mg l^{-1}]	Ca^{2+} [mg l^{-1}]	Sr^{2+} [µg l^{-1}]	Cl^{-} [mg l^{-1}]	NO_3^{-} [mg l^{-1}]	SO_4^{2-} [mg l^{-1}]
RB1/93	6-8	1.9	14.4	6.4	723.0	-22	2.8	2.51	0.15	16.7	1.0	1.8	8.3	145.9	565.48	53.5	0.2	9.2
RB1/99	3-9	2.5	16.3	6.5	970.0	13	4.9	2.58	0.13	22.7	1.4	1.7	9.4	194.8	717.56	122.8	0.1	1.7
RB3/93	5-9.5	6.0	14.3	6.9	523.0	176	1.4	0.06	0.05	17.5	0.3	3.3	8.7	70.2	552.81	16.5	1.8	29.9
B2	6-8	2.9	15.0	6.6	1084.0	-24	3.0	7.59	0.23	18.3	0.6	0.6	11.2	199.4	792.60	139.8	0.1	4.2
KB1	9-11	2.5	15.6	6.6	773.0	-76	2.2	7.37	0.23	17.6	0.2	2.1	12.2	135.8	721.66	47.5	0.1	10.1
KB2	9-11	3.0	15.8	6.6	906.0	-22	3.0	6.17	0.19	15.3	0.5	1.0	10.1	180.5	672.28	88.4	0.1	4.7
KB3a	6-9	2.8	14.8	6.6	893.0	-20	3.1	3.58	0.22	35.5	1.4	0.6	7.0	167.4	597.05	104.5	0.1	2.5
KB3b	10.5-12.5	2.9	13.6	6.6	850.0	-70	1.4	10.55	0.19	24.0	1.2	2.6	12.8	155.4	965.52	69.6	0.2	8.8
KB4a	6-8	2.7	13.8	6.2	783.0	-51	3.9	6.71	0.89	28.6	2.2	2.6	7.9	140.0	610.65	75.2	0.2	2.0
KB4b	9-12	2.7	14.5	6.4	957.0	-32	3.0	7.02	0.27	15.6	1.0	1.4	9.5	199.0	730.42	101.6	0.2	1.6
KB5a	5.5-8.5	1.4	15.4	6.2	798.0	-18	3.3	3.68	0.47	20.7	1.8	1.3	10.5	155.2	621.27	57.4	0.1	2.8
KB5b	10-13	1.4	14.0	6.5	783.0	-15	2.8	1.82	0.34	18.0	0.6	2.1	7.5	161.4	571.56	58.1	0.2	8.9
KB7b	10-12	3.1	15.0	7.0	465.0	-49	1.6	1.2	0.31	13.5	0.4	1.3	7.4	87.2	426.28	14.8	1.3	22.7

[a] Electrical conductivity.
[b] Oxidation-reduction potential.
[c] Total organic carbon.

Table S4. Chlorinated ethenes, methane, and ethene concentrations measured in the aquifer in October 2008 and their relative proportions.

Monitoring well	PCE [µg l^{-1}]	TCE [µg l^{-1}]	cDCE [µg l^{-1}]	VC [µg l^{-1}]	Total CEs [µg l^{-1}]	%PCE	%TCE	%cDCE	%VC	Methane [µg l^{-1}]	Ethene [µg l^{-1}]
RB1/93	12100.0	6210.0	5250.0	300.0	23860.0	40.7	26.4	30.2	2.7	222.3	56.2
RB1/99	13300.0	11800.0	7600.0	900.0	33600.0	30.5	34.2	29.8	5.5	466.2	101.5
RB3/93	<0.5	<10	<1	<0.5	0.0	0	0	0.0	0	<5	<5
B2	8.0	29.4	107.1	39.1	183.6	2.4	11.2	55.2	31.3	254.2	6.2
KB1	138.9	153.1	1310.0	445.0	2047.0	3.7	5.1	59.7	31.5	120.4	42.1
KB2	548.9	1380.0	1480.0	389.5	3798.4	9.4	29.7	43.2	17.6	492.0	118.9
KB3a	12.0	<10	172.1	147.1	331.2	1.7	0	42.3	56	151.2	37.2
KB3b	13.1	26.0	5530.0	546.6	6115.7	0.1	0.3	86.6	13.3	105.6	61.4
KB4a	503.4	355.5	280.8	93.9	1233.6	29.9	26.7	28.6	14.8	428.9	26.2
KB4b	171.0	238.4	4030.0	470.2	4909.6	2	3.5	80	14.5	337.6	159.4
KB5a	620.0	1520.0	1920.0	380.0	4440.0	9.1	28.1	48.1	14.8	523.5	159.9
KB5b	150.0	680.0	4170.0	120.0	5120.0	1.8	10.1	84.3	3.8	267.1	45.3
KB7b	<0.5	<10	1.1	<0.5	1.1	0	0	100.0	0	<5	<5

Figure S1. Phylogenetic placement of T-RFs identified by cloning-sequencing within the *Chloroflexi*. Clones are indicated with a "z" followed by a number and the T-RFs are shown in brackets. Bootstraps below 500 are not shown.

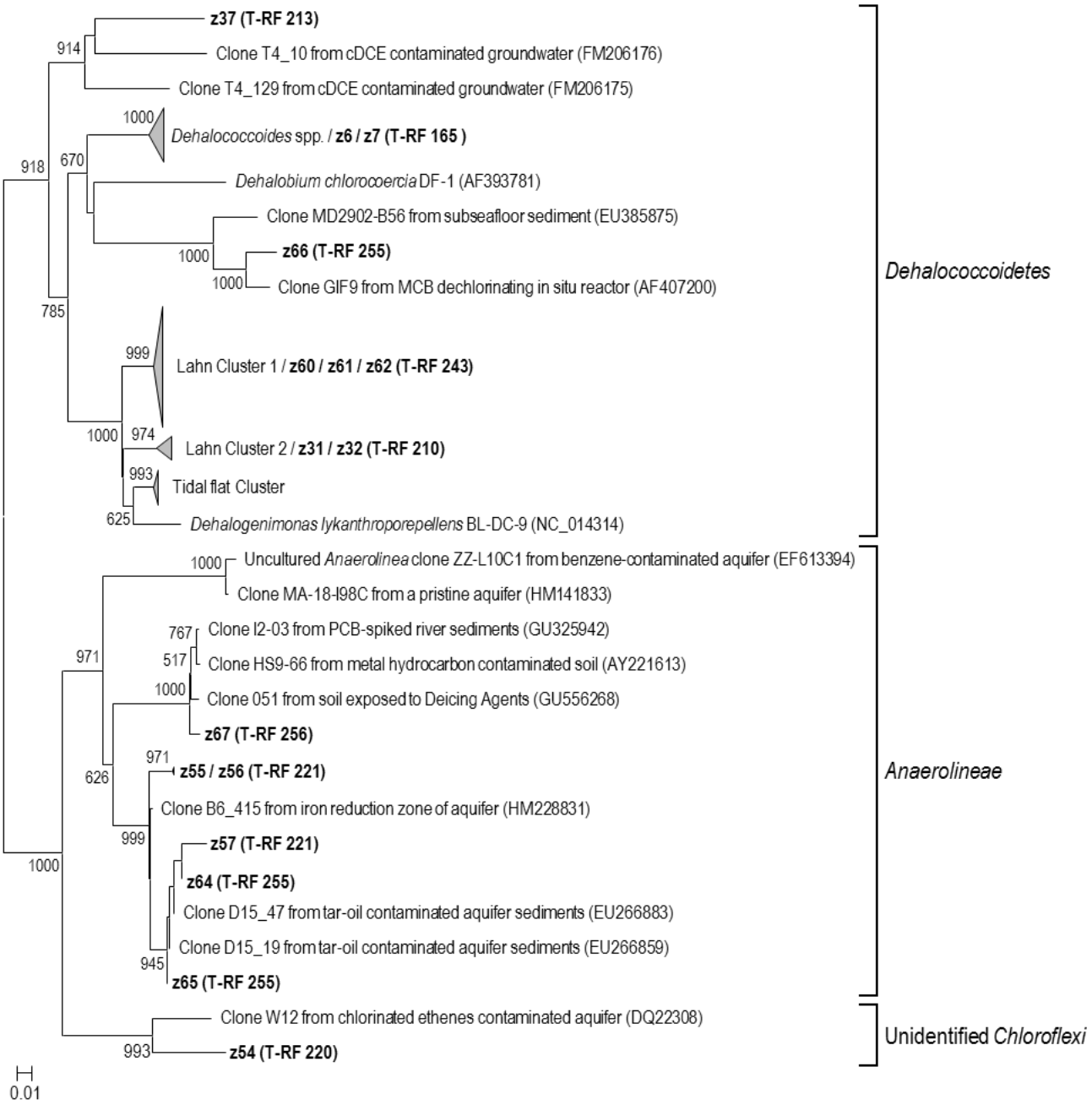


Figure S2. Phylogenetic placement of T-RFs identified by cloning-sequencing within the *β-Proteobacteria.* Clones are indicated with a "z" followed by a number and the T-RFs are shown in brackets. Bootstraps below 500 are not shown.

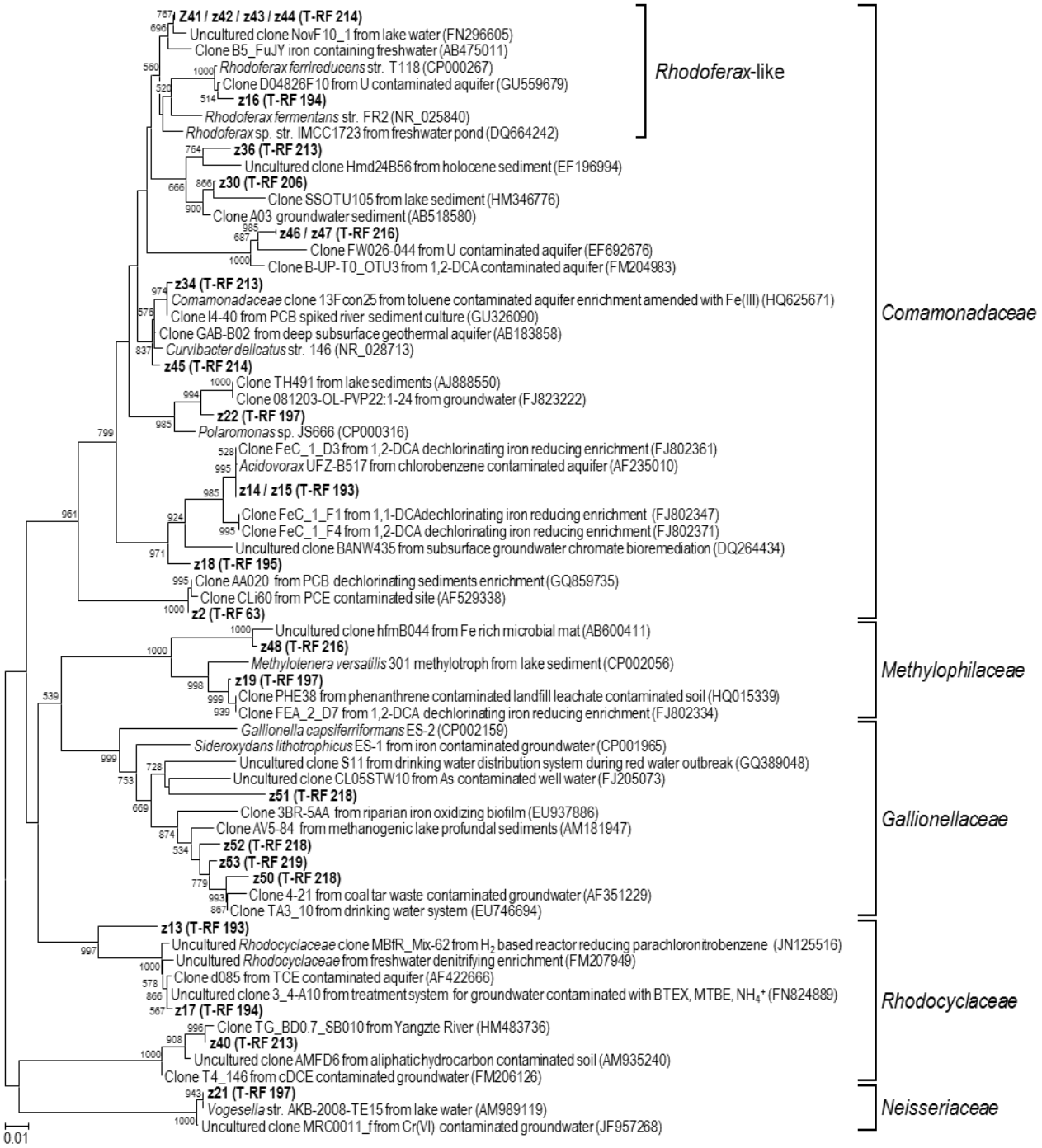


Figure S3 Phylogenetic placement of T-RFs identified by cloning-sequencing within the *Firmicutes*, *α*-Proteobacteria, *δ-Proteobacteria*, and *γ-Proteobacteria*,. Clones are indicated with a "z" followed by a number and the T-RFs are shown in brackets. Bootstraps below 500 are not shown.

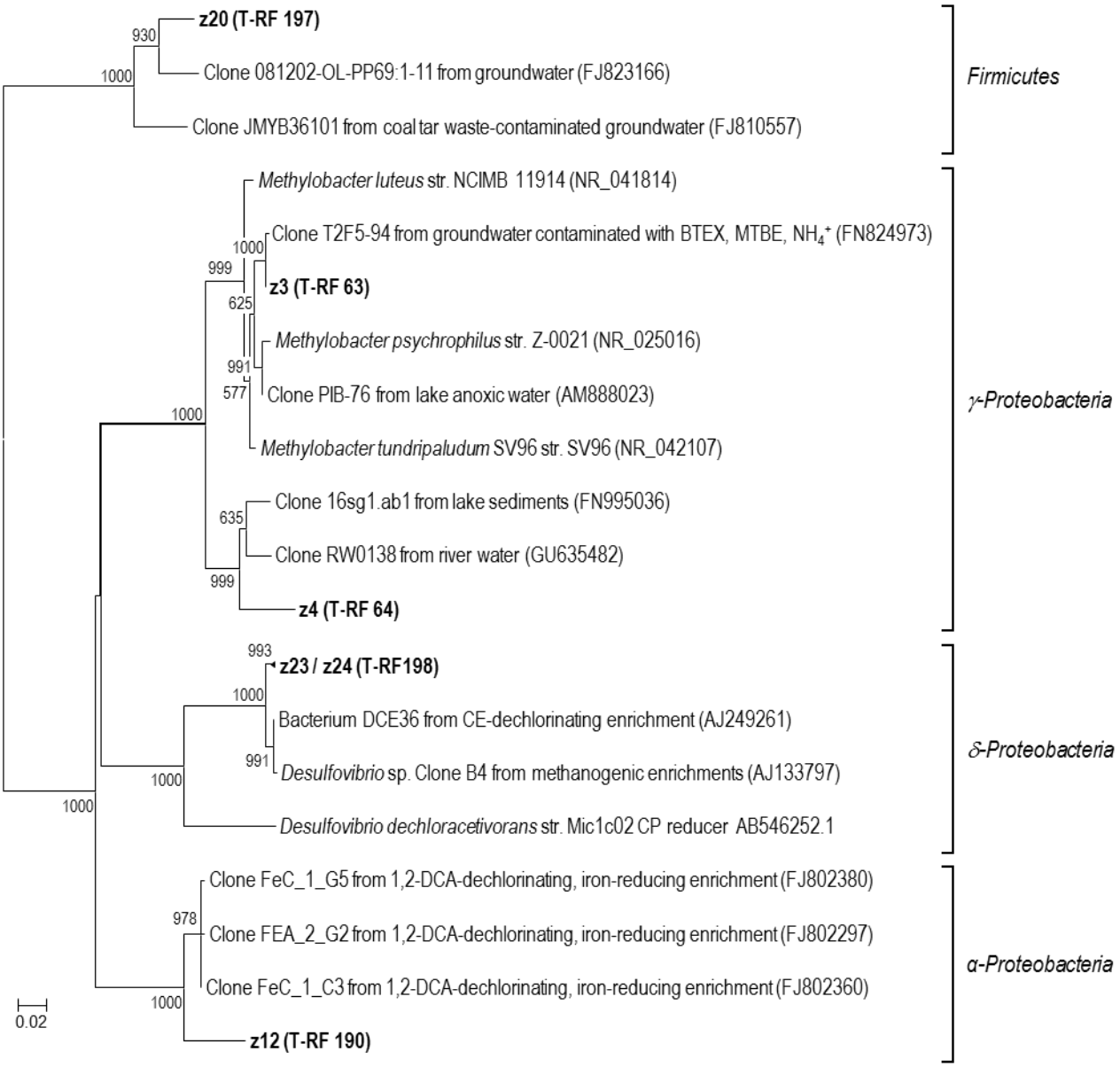
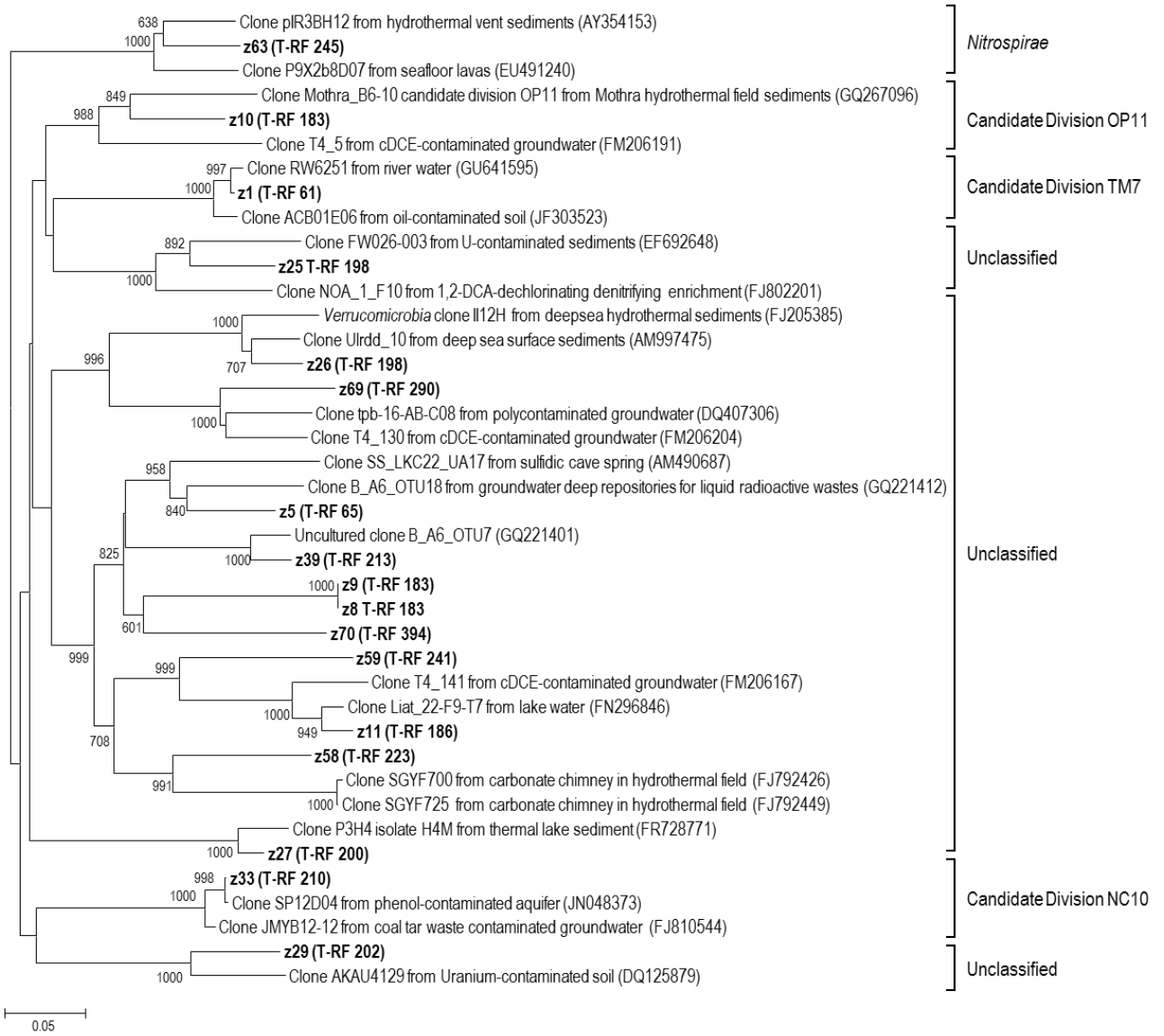


Figure S4. Phylogenetic placement of T-RFs identified by cloning-sequencing within the various phyla. Clones are indicated with a "z" followed by a number, and the T-RFs are shown in brackets. Bootstraps below 500 are not shown.



References

1. **Bradley, P. M.** (2000) Microbial degradation of chloroethenes in groundwater systems, *Hydrogeology Journal* 8 (1):104–111.
2. **Maymò-Gatell, X., Chien, Y. T., Gossett, J. M., and Zinder, S. H.** (1997) Isolation of a bacterium that reductively dechlorinates tetrachloroethene to ethene, *Science* 276 (5318):1568–1571.
3. **Middeldorp, P., Langenhoff, A., Gerritse, J., and Rijnaarts, H.** (2002) In situ biological soil remediation techniques. In Agathos, S., and Reineke, W. (Eds.) *Biotechnology for the Environment: Soil Remediation*, Springer, Dordrecht.
4. **Verce, M. F., Gunsch, C. K., Danko, A. S., and Freedman, D. L.** (2002) Cometabolism of cis-1,2-cichloroethene by aerobic cultures grown on vinyl chloride as the primary substrate, *Environmental Science & Technology* 36 (10):2171–2177.
5. **Bradley, P. M., and Chapelle, F. H.** (2010) Biodegradation of Chlorinated Ethenes, In Stroo, H.F., and Ward, C.H. (Eds), *In Situ Remediation of Chlorinated Solvent Plumes*, pp 39–67. Springer Science+Business Media, NY; USA.
6. **Cichocka, D., Nikolausz, M., Haest, P. J., and Nijenhuis, I.** (2010) Tetrachloroethene conversion to ethene by a *Dehalococcoides*-containing enrichment culture from Bitterfeld, *FEMS Microbiology Ecology* 72 (2):297–310.
7. **Fennell, D. E., Carroll, A. B., Gossett, J. M., and Zinder, S. H.** (2001) Assessment of indigenous reductive dechlorinating potential at a TCE-contaminated site using microcosms, polymerase chain reaction analysis, and site data, *Environmental Science & Technology* 35 (9):1830–1839.
8. **Hendrickson, E. R., Payne, J. A., Young, R. M., Starr, M. G., Perry, M. P., Fahnestock, S., Ellis, D. E., and Ebersole, R. C.** (2002) Molecular analysis of *Dehalococcoides* 16S ribosomal DNA from chloroethene-contaminated sites throughout north America and Europe, *Applied and Environmental Microbiology* 68 (2):485–495.
9. **McCarty, P. L., and Smith, D. P.** (1986) Anaerobic wastewater treatment, *Environmental Science & Technology* 20 (12):1200–1206.
10. **Rectanus, H. V., Widdowson, M. A., Chapelle, F. H., Kelly, C. A., and Novak, J. T.** (2007) Investigation of reductive dechlorination supported by natural organic carbon, *Ground Water Monitoring & Remediation* 27 (4):53–62.
11. **Heimann, A., Jakobsen, R., and Blodau, C.** (2010) Energetic constraints on H_2-dependent terminal electron accepting processes in anoxic environments: a review of observations and model approaches, *Environmental Science & Technology* 44 (1):24–33.
12. **Cord-Ruwisch, R., Seitz, H. J., and Conrad, R.** (1988) The capacity of hydrogenotrophic anaerobic bacteria to compete for traces of hydrogen depends on the redox potential of the terminal electron acceptor, *Archives of Microbiology* 149 (4):350–357.
13. **Lovley, D. R., and Goodwin, S.** (1988) Hydrogen concentrations as an indicator of the predominant terminal electron-accepting reactions in aquatic sediments, *Geochimica et Cosmochimica Acta* 52 (12):2993–3003.
14. **Heimann, A. C., Friis, A. K., and Jakobsen, R.** (2005) Effects of sulfate on anaerobic chloroethene degradation by an enriched culture under transient and steady-state hydrogen supply, *Water Research* 39 (15):3579–3586.
15. **Hoelen, T. P., and Reinhard, M.** (2004) Complete biological dehalogenation of chlorinated ethylenes in sulfate containing groundwater, *Biodegradation* 15 (6):395–403.
16. **Wei, N., and Finneran, K. T.** (2011) Influence of ferric iron on complete dechlorination of trichloroethylene (TCE) to ethene: Fe(III) reduction does not always inhibit complete dechlorination, *Environmental Science & Technology* 45 (17):7422–7430.

17. **Aeppli, C., Hofstetter, T. B., Amaral, H. I. F., Kipfer, R., Schwarzenbach, R. P., and Berg, M.** (2010) Quantifying in situ transformation rates of chlorinated ethenes by combining compound-specific stable isotope analysis, groundwater dating, and carbon isotope mass balances, *Environmental Science & Technology* 44 (10):3705–3711.

18. **Aeppli, C., Berg, M., Hofstetter, T. B., Kipfer, R., and Schwarzenbach, R. P.** (2008) Simultaneous quantification of polar and non-polar volatile organic compounds in water samples by direct aqueous injection-gas chromatography/mass spectrometry, *Journal of Chromatography* 1181 (1-2):116–124.

19. **Chapelle, F. H., Bradley, P. M., Thomas, M. A., and McMahon, P. B.** (2009) Distinguishing iron-reducing from sulfate-reducing conditions, *Ground Water* 47 (2):300–305.

20. **Jurgens, B. ., McMahon, P. B., Chapelle, F. H., and Eberts, S. M.** (2009) An Excel® workbook for identifying redox processes in ground water: U.S. Geological Survey Open-File Report 2009–1004 8p.

21. **R development Core Team** (2008), R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.

22. **Lane, D. J.** (1991) Nucleic acid techniques in bacterial systematics, John Wiley & Sons.

23. **Muyzer, G., Dewaal, E. C., and Uitterlinden, A. G.** (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction amplified genes coding for 16S ribosomal RNA, *Applied and Environmental Microbiology* 59 (3):695–700.

24. **Smits, T. H., Devenoges, C., Szynalski, K., Maillard, J., and Holliger, C.** (2004) Development of a real-time PCR method for quantification of the three genera *Dehalobacter*, *Dehalococcoides*, and *Desulfitobacterium* in microbial communities, *Journal of Microbiological Methods* 57 (3):369–378.

25. **Ebersole, R., and Hendrickson, E.** (2003) Nucleic acid fragments for the identification of dechloprinating bacteria. Patent WO/2003/064,695

26. **Bond, D. R., Holmes, D. E., Tender, L. M., and Lovley, D. R.** (2002) Electrode-reducing microorganisms that harvest energy from marine sediments, *Science* 295 (5554):483–485.

27. **Lanthier, M., Villemur, R., Lepine, F., Bisaillon, J. G., and Beaudet, R.** (2001) Geographic distribution of *Desulfitobacterium frappieri* PCP-1 and *Desulfitobacterium* spp. in soils from the province of Quebec, Canada, *FEMS Microbiology Ecology* 36 (2-3):185–191.

28. **Behrens, S., Azizian, M. F., McMurdie, P. J., Sabalowsky, A., Dolan, M. E., Semprini, L., and Spormann, A. M.** (2008) Monitoring abundance and expression of "*Dehalococcoides*" species chloroethene-reductive dehalogenases in a tetrachloroethene-dechlorinating flow column, *Applied and Environmental Microbiology* 74 (18):5695–5703.

29. **Rossi, P., Gillet, F., Rohrbach, E., Diaby, N., and Holliger, C.** (2009) Statistical assessment of variability of terminal restriction fragment length polymorphism analysis applied to complex microbial communities, *Applied and Environmental Microbiology* 75 (22):7268–7270.

30. **Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G. L., Solymos, P., Henry, M., Stevens, M. H. H., and Wagner, H.** (2009) Vegan: Community Ecology Package. R package version 1.16-22., *http://R-Forge.R-project.org/projects/vegan/*

31. **Legendre, P., and Gallagher, E.** (2001) Ecologically meaningful transformations for ordination of species data, *Oecologia* 129 (2):271–280.

32. **Legendre, P., and Legendre, L.** (1998) Numerical Ecology, Volume 24, Second Edition, Elsevier Science.

33. **Escofier, B., and Pages, J.** (1994) Multiple factor analysis (AFMULT package), *Computational Statistics & Data Analysis* 18 (1):121–140.

34. **Bécue-Bertaut, M., and Pagès, J.** (2008) Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data, *Computational Statistics & Data Analysis* 52 (6):3255–3268.

35. **Robert, P., and Escoufier, Y.** (1976) A unifying tool for linear multivariate statistical methods: The RV coefficient, *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 25 (3), 257–265.

36. **Borcard, D., Gillet, F., and Legendre, P.** (2011) Numerical Ecology with R, Springer.

37. **Huber, T., Faulkner, G., and Hugenholtz, P.** (2004) Bellerophon: a program to detect chimeric sequences in multiple sequence alignments, *Bioinformatics* 20 (14):2317–9.

38. **Thompson, J. D., Higgins, D. G., and Gibson, T. J.** (1994) Clustal W : Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Research* 22 (22):4673–4680.

39. **Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Sayers, E. W.** (2011) GenBank, *Nucleic Acids Res* 39:D32–7.

40. **Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., and Lipman, D. J.** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Research* 25 (17):3389–3402.

41. **Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S.** (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Molecular Biology and Evolution* 28 (10):2731–2739.

42. **Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J., and Knight, R.** (2010) QIIME allows analysis of high-throughput community sequencing data, *Nature Methods* 7 (5):335–336.

43. **Li, H., and Durbin, R.** (2010) Fast and accurate long-read alignment with Burrows–Wheeler transform, *Bioinformatics* 26 (5):589–595.

44. **McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., Desantis, T. Z., Probst, A., Andersen, G. L., Knight, R., and Hugenholtz, P.** (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea, *The ISME J*ournal 6:610–618.

45. **Liu, W. T., Marsh, T. L., Cheng, H., and Forney, L. J.** (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA., *Applied and Environmental Microbiology* 63 (11):4516–4522.

46. **Kaplan, C. W., and Kitts, C. L.** (2003) Variation between observed and true Terminal Restriction Fragment length is dependent on true TRF length and purine content, *Journal Microbiological Methods* 54 (1):121–125.

47. **Schwarzenbach, R. P., Gschwend, P. M., Imboden, D. M., and NetLibrary, I.** (2003) Environmental organic chemistry, Wiley Online Library.

48. **Lovley, D. R.** (1991) Dissimilatory Fe(III) and Mn(IV) reduction, *Microbiological Reviews* 55 (2):259–287.

49. **Lovley, D. R., Holmes, D. E., and Nevin, K. P.** (2004) Dissimilatory Fe(III) and Mn(IV) reduction, *Advances in Microbial Physiology* 49:219–286.

50. **Margalef, R.** (1958) Information theory in ecology, *General Systems* 3:36–71.

51. **Kittelmann, S., and Friedrich, M. W.** (2008) Identification of novel perchloroethene-respiring microorganisms in anoxic river sediment by RNA-based stable isotope probing, *Environmental Microbiology* 10 (1):31–46.

52. **Zengler, K., Richnow, H. H., Rosselló-Mora, R., Michaelis, W., and Widdel, F.** (1999) Methane formation from long-chain alkanes by anaerobic microorganisms, *Nature* 401 (6750):266–269.

53. **He, J. Z., Sung, Y., Dollhopf, M. E., Fathepure, B. Z., Tiedje, J. M., and Loffler, F. E.** (2002) Acetate versus hydrogen as direct electron donors to stimulate the microbial reductive dechlorination process at chloroethene-contaminated sites, *Environmental Science & Technology* 36 (18):3945–3952.

54. **Becker, J., Berardesco, G., Rittmann, B., and Stahl, D.** (2005) The role of syntrophic associations in sustaining anaerobic mineralization of chlorinated organic compounds, *Environmental Health Perspectives* 113 (3):310–316.

55. **Taş, N., van Eekert, M. H. A., de Vos, W. M., and Smidt, H.** (2010) The little bacteria that can - diversity, genomics and ecophysiology of "*Dehalococcoides*" spp. in contaminated environments, *Microbial Biotechnology* 3 (4):389–402.

56. **McMahon, P. B., and Chapelle, F. H.** (2008) Redox processes and water quality of selected principal aquifer systems, *Ground Water* 46 (2):259–271.

57. **Briones, A., and Raskin, L.** (2003) Diversity and dynamics of microbial communities in engineered environments and their implications for process stability, *Current Opinion in Biotechnology* 14 (3):270–276.

58. **Bradley, P. M.** (2003) History and ecology of chloroethene biodegradation: A review, *Bioremediation Journal* 7 (2):81–109.

59. **Maillard, J., Charnay, M.-P., Regeard, C., Rohrbach-Brandt, E., Rouzeau-Szynalski, K., Rossi, P., and Holliger, C.** (2011) Reductive dechlorination of tetrachloroethene by a stepwise catalysis of different organohalide respiring bacteria and reductive dehalogenases, *Biodegradation* 22 (5):949–960.

60. **Flynn, S. J., Loffler, F. E., and Tiedje, J. M.** (2000) Microbial community changes associated with a shift from reductive dechlorination of PCE to reductive dechlorination of cis-DCE and VC, *Environmental Science & Technology* 34 (6):1056–1061.

61. **Bethke, C. M., Ding, D., Jin, Q., and Sanford, R. A.** (2008) Origin of microbiological zoning in groundwater flows, *Geology* 36 (9):739–742.

62. **Heimann, A. C., Blodau, C., Postma, D., Larsen, F., Viet, P. H., Nhan, P. Q., Jessen, S., Duc, M. T., Hue, N. T. M., and Jakobsen, R.** (2007) Hydrogen thresholds and steady-state concentrations associated with microbial arsenate respiration, *Environmental Science & Technology* 41 (7):2311–2317.

63. **Lu, X. X., Tao, S., Bosma, T. N. P., and Gerritse, J.** (2001) Characteristic hydrogen concentrations for various redox processes in batch study, *Journal of Environmental Science and Health* 36 (9):1725–1734.

64. **Drzyzga, O., and Gottschal, J. C.** (2002) Tetrachloroethene dehalorespiration and growth of *Desulfitobactetium frappieri* TCE1 in strict dependence on the activity of *Desulfovibrio fructosivorans*, *Applied and Environmental Microbiology* 68 (2):642–649.

65. **Men, Y., Feil, H., Verberkmoes, N. C., Shah, M. B., Johnson, D. R., Lee, P. K., West, K. A., Zinder, S. H., Andersen, G. L., and Alvarez-Cohen, L.** (2012) Sustainable syntrophic growth of *Dehalococcoides ethenogenes* strain 195 with *Desulfovibrio vulgaris* Hildenborough and *Methanobacterium congolense*: global transcriptomic and proteomic analyses, *The ISME Journal* 6 (2):410-421.

66. **Stams, A. J. M., and Plugge, C. M.** (2009) Electron transfer in syntrophic communities of anaerobic bacteria and archaea, *Nature Reviews Microbiology* 7 (8):568–577.

67. **Crowe, S. A., O'Neill, A. H., Weisener, C. G., Kulczycki, E., Fowle, D. A., and Roberts, J. A.** (2007) Reductive dissolution of trace metals from sediments, *Geomicrobiology Journal* 24 (3-4):157–165.

68. **Ferris, F. G., Hallberg, R. O., Lyven, B., and Pedersen, K.** (2000) Retention of strontium, cesium, lead and uranium by bacterial iron oxides from a subterranean environment, *Applied Geochemistry* 15 (7):1035–1042.

69. **Zhuang, K., Izallalen, M., Mouser, P., Richter, H., Risso, C., Mahadevan, R., and Lovley, D. R.** (2011) Genome-scale dynamic modeling of the competition between *Rhodoferax* and *Geobacter* in anoxic subsurface environments, *The ISME Journal* 5 (2):305–316.

70. **Finneran, K. T., Johnsen, C. V., and Lovley, D. R.** (2003) *Rhodoferax ferrireducens* sp. nov., a psychrotolerant, facultatively anaerobic bacterium that oxidizes acetate with the reduction of Fe(III), *International Journal of Systematic and Evolutionary Microbiology* 53 (3):669–673.

71. **Emerson, D., and Weiss, J. V.** (2004) Bacterial iron oxidation in circumneutral freshwater habitats: findings from the field and the laboratory, *Geomicrobiology Journal* 21 (6):405–414.

72. **Marsh, T. L.** (2005) Culture-independent microbial community analysis with terminal restriction fragment length polymorphism. In *Environmental Microbiology*, Academic Press (Ed.), pp 308–329.

73. **Dunbar, J., Ticknor, L. O., and Kuske, C. R.** (2001) Phylogenetic specificity and reproducibility and new method for analysis of terminal restriction fragment profiles of 16S rRNA genes from bacterial communities, *Applied and Environmental Microbiology* 67 (1):190–197.

74. **Villemur, R., Constant, P., Gauthier, A., Shareck, M., and Beaudet, R.** (2007) Heterogeneity between 16S ribosomal RNA gene copies borne by one *Desulfitobacterium* strain is caused by different 100-200 bp insertions in the 5' region, *Canadian Journal of Microbiol*ogy 53 (1):116–28.

75. **Bethke, C. M., Sanford, R. A., Kirk, M. F., Jin, Q., and Flynn, T. M.** (2011) The thermodynamic ladder in geomicrobiology, *American Journal of Science* 311 (3):183–210.

76. **Neal, A. L., Techkarnjanaruk, S., Dohnalkova, A., McCready, D., Peyton, B. M., and Geesey, G. G.** (2001) Iron sulfides and sulfur species produced at hematite surfaces in the presence of sulfate-reducing bacteria, *Geochimica Et Cosmochimica Acta* 65 (2):223–235.

77. **Cabirol, N., Jacob, F., Perrier, J., Fouillet, B., and Chambon, P.** (1998) Interaction between methanogenic and sulfate-reducing microorganisms during dechlorination of a high concentration of tetrachloroethylene, *Journal of General and Applied Microbiology* 44 (4):297–301.

78. **Aulenta, F., Beccari, M., Majone, M., Papini, M. P., and Tandoi, V.** (2008) Competition for H_2 between sulfate reduction and dechlorination in butyrate-fed anaerobic cultures, *Process Biochemistry* 43 (2):161–168.

79. **Aulenta, F., Canosa, A., Leccese, M., Papini, M. P., Majone, M., and Viottit, P.** (2007) Field study of in situ anaerobic bioremediation of a chlorinated solvent source zone, *Industrial & Engineering Chemistry Research* 46 (21):6812–6819.

80. **Imfeld, G., Nijenhuis, I., Nikolausz, M., Zeiger, S., Paschke, H., Drangmeister, J., Grossmann, J., Richnow, H. H., and Weber, S.** (2008) Assessment of in situ degradation of chlorinated ethenes and bacterial community structure in a complex contaminated groundwater system, *Water Research* 42 (1), 871–882.

Incomplete reductive dechlorination of tetrachloroethene as a result of hydrogeological characteristics of the Lyss aquifer

5 Incomplete reductive dechlorination of tetrachloroethene as a result of hydrogeological characteristics of the Lyss aquifer

Abstract

Accumulation of toxic daughter molecules formed during incomplete organohalide respiration (OHR) of chloroethenes (CEs) represents a major obstacle for the remediation of contaminated sites using monitored natural attenuation. In this study, reasons for spatial discrepancies in the fate of OHR were investigated in a tetrachloroethene (PCE)-contaminated aquifer in which cDCE was accumulating locally. Two zones with *a priori* similar geological conditions were delimited on this site. The first one displayed complete OHR, whereas the second one displayed an accumulation of cDCE. Terminal-restriction fragment length polymorphism was used in combination with pyrosequencing to describe the bacterial communities in both zones. The results of these analyses were combined with hydrogeochemical data using multivariate statistical tools such as multifactorial analysis, enabling building-up a conceptual model of the aquifer functioning and hypothesizing the reasons for the accumulation of cDCE. Bacterial fingerprinting provided evidence of the presence of radically different bacterial communities in the two zones. Pyrosequencing datasets revealed that the bacterial populations in the zone of complete OHR were typical of highly reducing environments and were dominated by *Dehalococcoides* sp. Conversely, the zone where incomplete OHR was leading to the accumulation of cDCE was inhabited by bacterial populations typical of more oxidized environments, such as nitrate-reducing and aerobic bacteria. Finally, based on hydrogeological expertise, the observed OHR discrepancy could be attributed to the occasional upward recharge of the anoxic lithological layer with oxidized groundwater, coupled with local discrepancies in the aquifer conformation and organic carbon content. These processes were probably hindering the development of an efficient OHR process in Zone II, leading to the observed accumulation.

5.1 Introduction

Accumulation of cis-dichloroethene (cDCE) and vinyl chloride (VC) is frequently observed in groundwater contaminated with tetrachloroethene (PCE) or trichloroethene (TCE) as a result of the incomplete organohalide respiration (OHR) of chlorinated ethenes (CEs) (*1*, *2*). Differences in the extent of OHR was also shown to occur in limited sections of contaminated sites (*3*, *4*). Reasons for the uneven degradation of the contaminant has previously been attributed to the irregular distribution of organic substrates feeding the fermenting and syntrophic guilds, resulting in locally insufficient amounts of electron donors fuelling the OHR guild (*2*, *5*). Oxidation-reduction (redox) state has been shown to vary highly across contaminated sites, reflecting the activity of various terminal electron-accepting processes (TEAPs), which are sometimes thermodynamically more efficient than OHR (*6*, *7*). Furthermore, high amounts of selected electron acceptors coupled with low electron donor requirements may provide a competitive advantage to some TEAPs over OHR. In contrast with PCE and TCE, cDCE and VC are only efficiently dechlorinated between sulfate-reducing and methanogenic conditions (*3*, *8*). In addition, Abe et al. (2009) reported that the occurrence of the dedicated OHRB, "*Dehalococcoides*" sp., was positively correlated with the amount of soil organic content in streambed sediments (*3*). As a recall, only *Dehalococcoides* sp. has been shown to catalyze the reductive dechlorination of cDCE and VC to the final harmless product ethene (*9*). Although some strains have been shown to successfully dechlorinate PCE and TCE, the ecological niche of *Dehalococcoides* sp. has been shown to be mainly limited to the reductive dechlorination of cDCE and VC in consortia and in field conditions (*3*, *10*, *11*). According to the present state of knowledge, *Dehalococcoides* sp. is a highly specialized obligate OHRB (*12*), showing a metabolic versatility reduced to the use of halogenated compounds as electron acceptors, and H_2 as the sole electron donor (*13–15*). However, all *Dehalococcoides* sp. strains do not bear *vcr*A and *bvc*A, the genes known to encode the enzymes catalyzing the reductive dechlorination of cDCE and VC (*12*). Thus, the detection of specific strains of *Dehalococcoides* sp. using molecular tools such as quantitative PCR does not *per se* enable anticipating the complete degradation of CEs. Strains of *Dehalococcoides* sp. isolated to date show a high degree of homology (*12*) and no clear link can be drawn between their phylogeny and their catabolic capabilities. Moreover, the presence of a large catabolic potential is not a guarantee of the success of reductive dehalogenation if the required environmental conditions are not met. Investigation of the reasons for partial degradation of CEs can therefore not be reduced to simply testing the presence of *Dehalococcoides* sp. Such reductionist approaches can prove counter-productive in the study of complex microbial systems (*12*, *16*). Several other biological and hydrogeological characteristics must be investigated prior to the application of monitored natural attenuation (MNA) in a contaminated site (*5*, *17*). The features include the diversity of the OHR bacterial guild, the availability of electron donors, as well as the distribution of local redox conditions. The geological structure and groundwater fluxes should be characterized as well (*5*, *17*). Often neglected, the temporal behavior of the aquifer may be of primary importance in the success of natural attenuation. For instance, the aquifer recharge may considerably influence the redox

conditions, by bringing water showing a typical surface water signature to the groundwater system and accordingly causing a shift in the dominant TEAPs (*18*). Physical extent of the impacting of the aquifer recharge on OHR and its possible implication in the local accumulation of lower CEs is not well known. McGuire et al. (*19*) reported that shifts in TEAPs due to aquifer recharge were restricted to small-scale interfaces between already present groundwater and "fresh" groundwater. Yet, temporal fluctuations induced by hydrostatic pressures are likely to negatively impact the formation of a habitat having a redox state favorable to the complete reduction of CEs.

The objective of this study was to find out the causes of short-distance discrepancies in the fate of CEs in a contaminated site which displayed *a priori* favorable environmental conditions for OHR. A synergistic approach was used to evaluate the impact of factors that were likely to explain the causes of the heterogeneity of the OHR process. We therefore analyzed different components of the ecosystem to acquire a comprehensive understanding of the aquifer functioning prior to answering more specific questions. Hydrological, geological, chemical, and microbiological aspects were assessed and further combined using multivariate statistical techniques to get a comprehensive picture of mechanisms governing the aquifer functioning.

5.2 Materials and Methods

5.2.1 Test site

The contaminated site is located in Lyss (BE, Switzerland), and is part of the Seeland aquifer. The aquitard is composed of the ground-moraine of glacio-riparian depositional origin and is located around 8 m below surface and deeper. It is overlaid by two superimposed lithological layers of Quaternary depositional origin. The lower one is composed of coarse gravels of alluvial origin, showing a high hydraulic conductivity (1.10^{-3} ms^{-1}). This layer displays high dissolved oxygen concentrations (DO) and low organic matter content, as reflected by low dissolved organic carbon concentration (DOC). The upper layer is composed of finer material coming from flooding events, such as fine sand and silt, mixed with low amounts of residual peat material and forming a heterogeneous geological body. It is anoxic, showing high organic carbon content in the sediment (0.4 to 2% total organic carbon) and a low hydraulic conductivity (5.10^{-5} ms^{-1}). The boundary between the two layers is gradual and gives rise to a vertical gradient of grain size distribution, DO, and DOC. The aquifer is partly recharged with water originating from the nearby Lyssbach River, located South East ca. 20 m upstream of the test site. However the main recharge is carried out with water from the global Seeland aquifer. The upper layer reaches a depth of approximately 4 to 5 m below surface upstream and gets thinner downstream. The most probable source zone of PCE contamination is located on the location of a former dry cleaning facility with PCE which has been in use between 1956 and 1999. The plume of contamination extends towards the North-West (Figure 1 and Figure 2).

The site was assessed using monitoring wells situated in two zones, I and II. Zone I was located close to the source zone, in the upstream part of the contamination plume. The aquifer in this area was impermeable to

surface infiltration. Zone II, located further downstream of the contamination plume was covered by lawn. Fourteen pre-packed screened piezometers (P) with 0.75 inches (1.91 cm) internal diameter (3/4" PrePack Well, AMS, USA), as well as 4 multilevel piezometers (ML; Ducommun et al., in prep) were used in Zone I and II. Furthermore, four reference piezometers (R1 to R4) located on the other bank of to the Lyssbach River were monitored (Figure 1 and Figure 2).

5.2.2 Groundwater sampling

Fifty-five groundwater samples were collected during four sampling campaigns carried out between March 2009 and August 2010. Groundwater samples were labeled with the name of the monitoring well they were collected in, followed by the labels A to D in brackets, according to their sampling date. Prior to the collection of groundwater samples, each well was purged (at least one well volume) using a peristaltic pump (Type P2.52, Eijkelkamp, Giesbeek, Netherlands) through a PTFE tube (inner diameter 4 mm, Semadeni SA, Switzerland) at a flow rate of 100 mL/min until physical-chemical parameters measured on-site (temperature, pH, electrical conductivity, oxidation-reduction potential) reached stable readings. Groundwater aliquots for the analysis of DOC, Fe^{2+}, Mn^{2+}, major anions (Cl^-, F^-, Br^-, NO_3^-, NO_2^-, SO_4^{2-}, and PO_4^{3-}) and cations (Li^+, Na^+, K^+, NH_4^+, Ca^{2+}, Mg^{2+}, and Sr^{2+}), and sulfide were filtered on-site with 0.45 µm nitrocellulose membranes (Millipore, USA). Filtered groundwater samples for the analysis of DOC were collected in Falcon-like tubes (Sarstedt AG&Co, Germany) and stored at 4°C in the dark until analysis, which was carried out within one day after sampling. For the analysis of Fe^{2+} and Mn^{2+}, 20 mL filtered groundwater was acidified to pH < 2 with 10% (v/v) HNO_3 (Sigma-Aldrich, USA) in 20-mL VOC vials (BGB Analytik AG, Switzerland) and analysed within one week after sampling. For the analysis of major anions and cations, 60 mL filtered groundwater was collected separately into two 30-mL polypropylene bottles (Sarstedt AG&Co, Germany). Groundwater aliquots for cations were acidified to pH < 2 with 10% (v/v) HNO_3 (Sigma-Aldrich, USA). Groundwater samples for the analysis of anions and cations were stored at 4 °C and analysed within one week after sampling. Approximately 50 mL of groundwater were collected in 50-mL Falcon-like tubes for the quantification of S^{2-} (Sarstedt AG&Co, Germany), complemented with 2 drops of 2N zinc acetate (Sigma-Aldrich, USA), basified to pH > 9 with 1M NaOH (Sigma-Aldrich, USA), and analyzed within one week after sampling. Groundwater samples for the quantification of VOC concentrations were collected into 40-mL glass vials sealed without headspace using PTFE-lined screw caps (BGB Analytik AG, Switzerland) using a foot-valve pump (Waterra Pumps Ltd., Canada) to prevent degasing of the chlorinated compounds.

Groundwater samples for microbiological analyses were collected using the peristaltic pump at a flow rate of 100 mL/min from multilevel piezometers and at 1000 mL/min from prepacked piezometers. Samples were stored in sterile 1-L propylene Nalgene bottles (Thermo Fisher Scientific, USA) at 4°C in the dark, and were processed within 2 days after sampling.

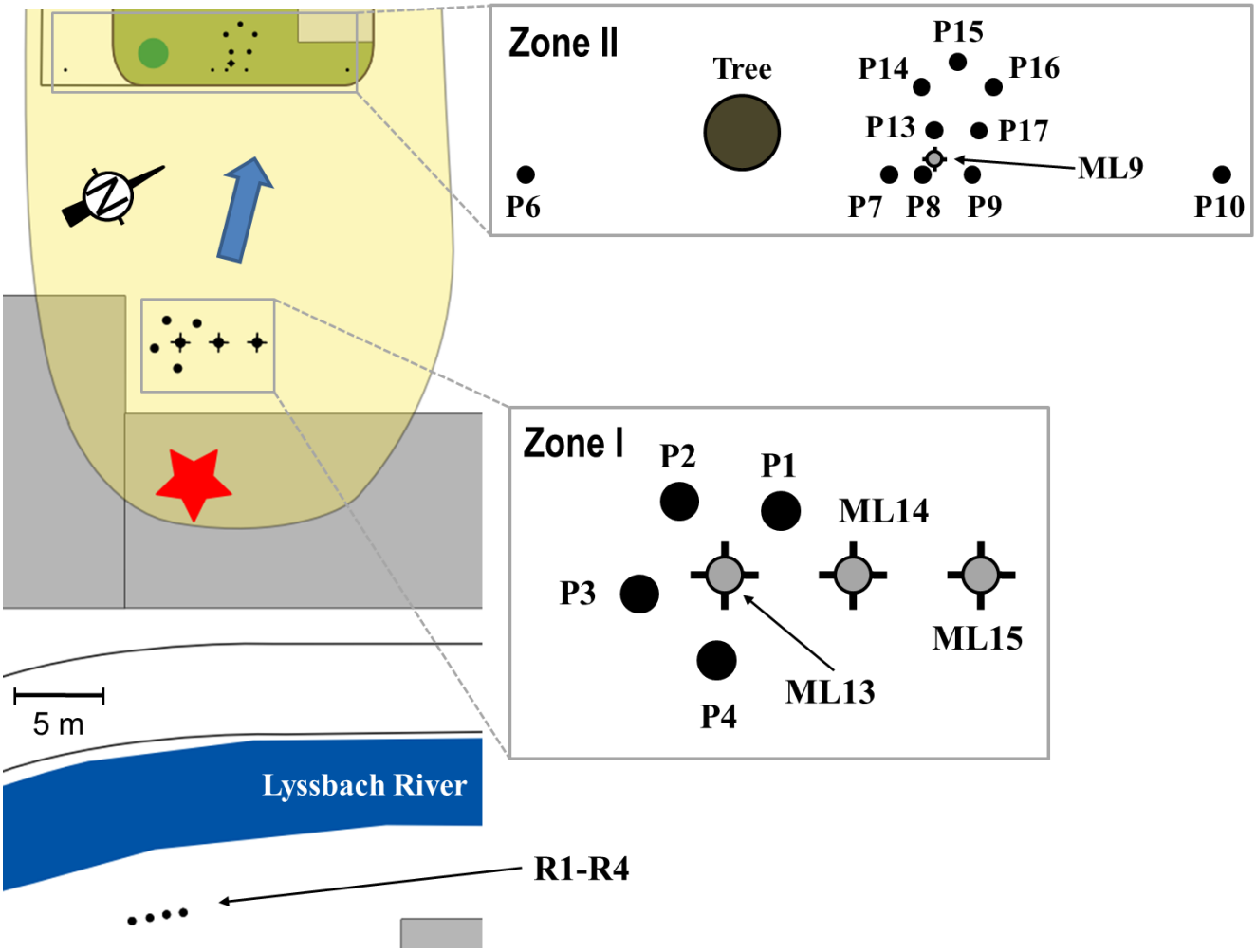


Figure 1 modified from Ducommun et al., in prep. Map of the chloroethene-contaminated site of the Lyss aquifer indicating the most probable location of the contamination source (red star) and the locations of the monitoring wells (P: prepacked piezometers; ML: multilevel piezometers; R: reference wells) in the investigated Zones I and II. The blue arrow indicates the main groundwater flow direction, and the contamination plume is indicated with yellow color. Zones I and II are enlarged on the right side of the figure. For more details, such as sampling depths and contaminant concentrations, please see Figure 2 and Tables S1 and S2.

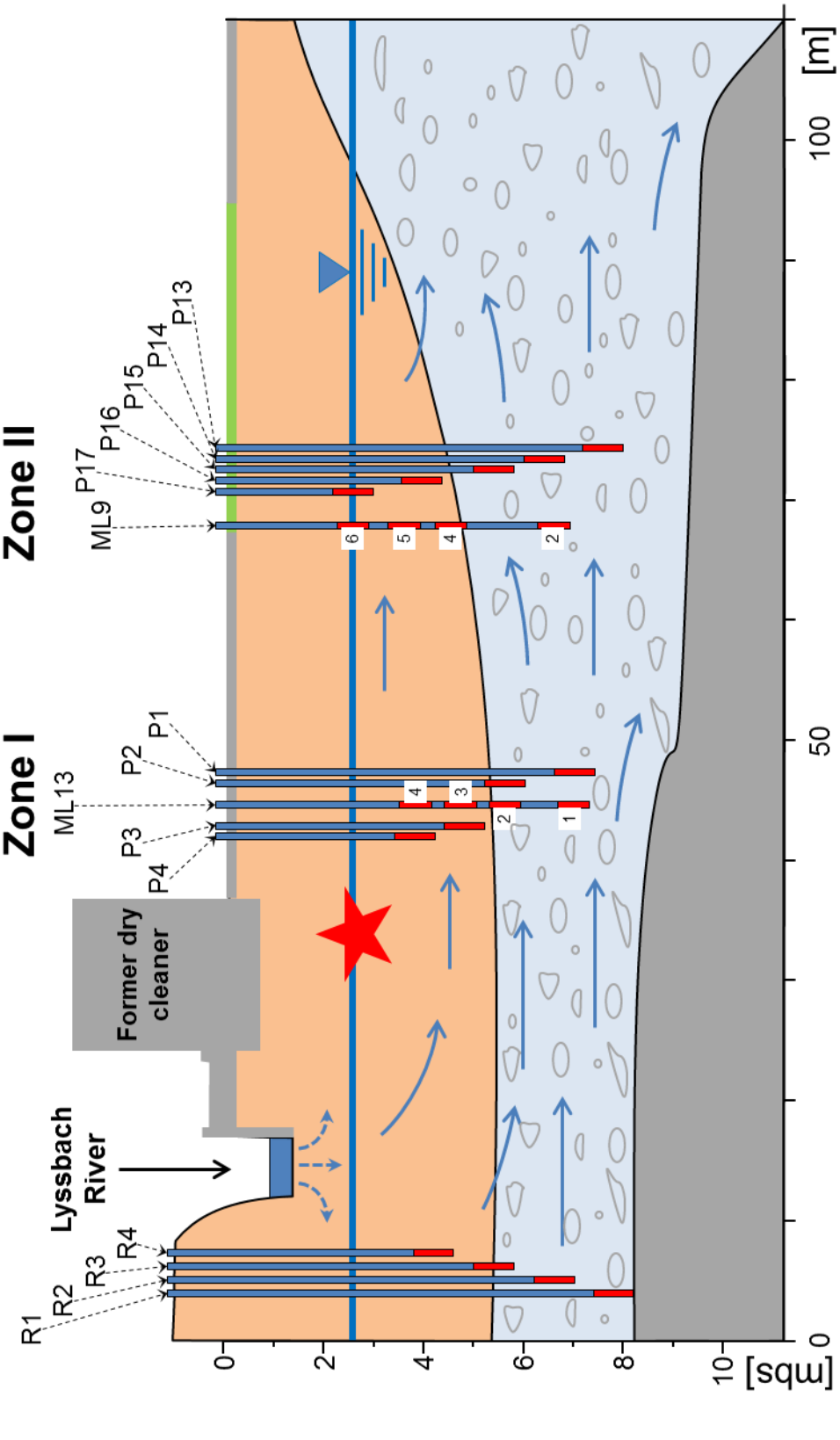


Figure 2, modified from Ducommun et al (in prep.). Section of the Lyss aquifer showing the geological layers, the Lyssbach River, and the former dry cleaner. Orange: upper lithological layer with low hydraulic conductivity, low dissolved oxygen concentration, and high organic carbon content; Blue: lower lithological layer with high hydraulic conductivity, high dissolved oxygen concentration, and low organic carbon content; Grey: aquitard. The main monitoring wells of Zones I and II discussed in this chapter are shown, as well as the reference wells (R1-R4). The red rectangles represent the screens of the piezometers. For multilevel piezometers ML9 and ML13, the numbers on the side of the screens refer to the extension of the well name. Depth is indicated on the y axis; mbs: meters below surface. The distance along the main groundwater flow is indicated on the x axis. The estimated contaminant source zone is shown with a red star. Blue arrows indicate estimated groundwater fluxes within the aquifer.

5.2.3 Analytical procedures

DOC concentrations were analyzed in a Shimadzu ASI-5000 TOC-analyzer (Shimadzu Corp., Japan). Concentrations of Fe^{2+} were measured colorimetrically following the phenantroline method (*20*) with a detection limit of 0.05 mg/L and an uncertainty of ± 5%. Mn^{2+} concentrations were measured colorimetrically using a spectrophotometer Lambda 10 (Perkin Elmer, Massachusetts, USA) following the FAD method (*21*) with a detection limit of 0.05 mg/L and an uncertainty of ± 3.5%. Major anions and cations were analysed by ion chromatography (Dionex DX-120) with a detection limit of 0.1 mg/L and an uncertainty of ± 5%. Sulfide concentrations were determined colorimetrically (*22*) with a detection limit of 0.01 mg/L and an uncertainty of ± 10%. All colorimetric analyses were carried out with a spectrophotometer Lambda 10 (Perkin Elmer, USA). Concentrations of chlorinated compounds in groundwater were determined using an in-vial purge-and-trap (P&T) extraction (VSP 4000, IMT Innovative Messtechnik GmbH, Germany) followed by gas chromatography analysis using a flame ionization detector (GC-FID, CP-3800 Varian Inc., USA) equipped with a 30 m x 0.32 mm Rt™-QSPLOT plot fused silica capillary column (RESTEK, USA). The detector temperature was at 250°C. Oven was first heated at 100°C for 4 min. Temperature was then increased at 10°C/min and kept stable at 240°C for 5 min. The carrier gas was He (purity 99.99%). The P&T procedure consisted in a 20 min purge of 20 mL (N_2)/min, followed by a trap at -35°C with a Tenax® TA adsorbent resin (Scientific Instrument Services, Inc., USA), 7 min desorption at 240°C, and 140 min transfer time at 1000 mbar pressure. Detection and quantification limits for the chlorinated compounds were 0.2 µg/L and 0.6 µg/L, respectively. Concentrations of ethene, ethane and methane were determined by headspace gas chromatography analysis using a flame ionization detector (GC-FID, CP-3800 Varian Inc., USA) equipped with a 15 m x 0.32 mm, 1.5 µm carboPLOT column (Agilent Technologies, USA). The detector temperature was at 250°C. Oven was heated at 40°C for 10 min and carrier gas was He (purity 99.99%). The loop volume was 100 µL. Five milliliters of gas (argon, nitrogen, or helium) was added to the sample by taking out the equal quantity of water using syringes of 10 mL, by volume compensation. Vials were shaken upside-down overnight at room temperature and stored in 4°C until the analysis. The analyses of VOCs were performed using the external standard method. Resulting calibration curves were linear ($r^2 > 0.999$) and F-test showed 99% confidence. Detection and quantification limits for the chlorinated compounds were 0.2 µg/L and 0.6 µg/L, respectively. For ethene, ethane and methane the detection limit was 2.5 µg/L. The uncertainty of the analytical method was ± 9, 11, 10 and 8% for PCE, TCE, cDCE and VC, respectively and ± 12% for ethene, ethane and methane.

5.2.4 DNA extraction

Water samples were filtered through 0.2-µm autoclaved polycarbonate membranes (Isopore™ Membrane Filters, Millipore) with a mobile filtration system (Filter Funnel Manifolds, Pall Corporation). DNA was extracted using the PowerSoil™ DNA Extraction Kit (Mo Bio Laboratories, Inc.) following the manufacturer instructions,

except that the samples were processed in a bead-beater (Fastprep FP120, Bio101) at 4.5 m/s for 30 s after the addition of solution C1. Extracted DNA was quantified with a ND-1000 Nanodrop® spectrophotometer (Thermo Fisher Scientific, USA).

5.2.5 Terminal-restriction fragment length polymorphism (T-RFLP)

The analysis of the bacterial communities using terminal-restriction fragment length polymorphism (T-RFLP) as well as their numerical treatment were based on the protocol described in (*23*) with the following modifications: (i) 30 µl PCR reactions were composed of 3 µL 10X Y buffer, 2.4 µL 10 mM dNTPs, 1.5 µL of each 10 µM primer, 6 µL 5X enhancer P, 1.5 U PeqGold Taq polymerase and 0.2 ng/µl template DNA (final concentration), completed with autoclaved and UV-treated Milli-Q water (Millipore, USA); (ii) for each DNA extract, PCR amplification was carried out in triplicates; (iii) the PCR products were purified with the purification kit Montage® PCR Centrifugal Filter Devices (Millipore, USA) according to the manufacturer's instructions.

5.2.6 Multivariate statistical analyses based on the T-RFLP profiles

All multivariate statistical analyses were carried out with R (*24*) and the additional package Vegan (*25*). In a preliminary step, the bacterial community dataset was constructed as follows: Ruzicka dissimilarities were calculated between replicate T-RFLP profiles; the replicate at the centroid (i.e. displaying the lowest dissimilarity with the other two) was selected as the most representative of each sample. The bacterial community dataset was transformed using a Hellinger transformation (*26*), whereas the environmental dataset was standardized (*27*). Multifactorial analysis (MFA) (*28, 29*) was used to investigate the relationships between environmental data and T-RFLP profiles in order to point out major ongoing processes. RV multivariate correlation coefficients and p-values (*30*) were computed between the bacterial community dataset and single environmental variables in order to get numerical correlations between the environmental dataset and the bacterial communities. Partial redundancy analysis (pRDA) (*27*) was carried out on the Hellinger transformed binary T-RFLP profiles to remove any impact of the different pumping rates used during groundwater sampling. Resulting site scores were computed using Ward's minimum variance clustering to investigate the BCS in the upper geological layer where OHR was susceptible to take place. Most representative T-RFs of each group were defined using their species indicator values (IndVal) (*31*).

5.2.7 Quantitative PCR (qPCR)

One sample from the upper layer in Zone I and one from the same layer in Zone II were selected for comparison of the populations of *Dehalococcoides* sp. between Zones I and II, as well as the VC reductive dehalogenase genes *vcr*A and *bvc*A (GenBank accession numbers YP003330719 and AY563562, respectively). The total numbers of 16S rRNA genes of *Dehalococcoides* sp. and *vcr*A and *bvc*A genes were quantified in a RotorGene RG3000 real-time PCR machine (Corbett Research, Australia) and normalized to total bacterial 16S rRNA genes. Primer sets used to amplify the different genes are listed in Table 1. Plasmids for standard curves of *Dehalococcoides* 16S rRNA gene, *vcr*A, and *bvc*A were prepared using DNA amplicons from the PCE-to-ethene dechlorinating MAROC11b consortium. Plasmids pDhc (16S rRNA gene of *Dehalococcoides* sp.), pvcrA (*vcr*A gene) and pbvcA (*bvc*A gene) were isolated from 50 mL LB culture using the Qiagen Plasmid Midiprep kit and linearized with *Nco*I (Promega). The digest was purified using the MSB® Spin PCRapace (Stratec Molecular GmbH, Germany) according to the manufacturer's instructions. The DNA was quantified in a Nanodrop spectrophotometer (Nanodrop®, ND-1000, Thermo Fisher Scientific, USA) and copy numbers per microliter were calculated from the number of base pairs of the plasmid, the average molecular weight of a base pair in double-stranded DNA (660 Da) and the concentration in μg/μL obtained from the quantification. pDhc first standard sample contained $3.44 \cdot 10^6$ copies/μL, and ten-fold dilutions were prepared down to $3.44 \cdot 10^2$ copies/μL. pvcrA first standard sample contained $4 \cdot 10^6$ copies/μL, and ten-fold dilutions were prepared down to $4 \cdot 10^2$ copies/μL. pbvcA first standard sample contained $8 \cdot 10^6$ copies/μL, and ten-fold dilutions were prepared down to $8 \cdot 10^2$ copies/μL. Reactions for qPCR were prepared as follows: 5.0 μL KAPA® SYBR FAST 2X qPCR Master Mix (Kapa Biosystems, Inc., USA), 0.2 μL of each primer, 2.1 μL sterile Milli-Q water (Millipore, USA), and 2.5 μL template DNA solution were used per 10 μL reaction volume. Samples were analyzed in triplicate, whereas standard curves were measured in duplicate.

Table 1. List of primers used for the PCR amplifications of bacterial 16S rRNA gene pool, 16S rRNA gene of *Dehalococcoides* sp., and vinyl chloride reductive dehalogenase genes *vcr*A and *bvc*A.

Target gene	Name of primer	Sequence (5'-3')	Reference
Eubacteria 16S rRNA gene	Eub-27f	AGAGTTTGATCMTGGCTCAG	(*32*)
	Univ-518r	ATTACCGCGGCTGCTGG	(*33*)
Dehalococcoides sp. 16S rRNA gene	DHC443f	GGTAATACGTAGGGAAGCAAGCG	This study
	DHC611r	GTTTTCTCCTGCTGTACTCTAGTCC	modified from (*1*)
*vcr*A	vcrA880f	CCCTCCAGATGCTCCCTTTA	(*34*)
	vcrA1018r	ATCCCCTCTCCCGTGTAACC	(*34*)
*bvc*A	bvcA227f	TGGGGACCTGTACCTGAAAA	(*34*)
	bvcA523r	CAAGACGCATTGTGGACATC	(*34*)

5.2.8 Pyrosequencing

DNA samples selected above were sent for pyrosequencing (Research and Testing Laboratory LLC, USA; using a 454/Roche GS-FLX Titanium Instrument (Roche, NJ, USA)) for comparison of the BCS present in the upstream and downstream sections (Zones I and II, respectively) of the aquifer. Amplification was carried out with the same primer set as used for T-RFLP analysis, Eub-27f (*32*) and Univ-518r (*33*). Pyrosequencing data were pre-processed in QIIME (*35*) on the Vital-IT platform (Vital-IT, Swiss Institute of Bioinformatics, Switzerland). Briefly, the script split_library.py was used to keep only sequences above 300 bp with a PHRED sequencing quality score above 20, and to remove the tags and primers. Denoising for the removal of classical 454 pyrosequencing flowgram errors such as homopolymers (*36*, *37*) was carried out with the script denoise_wrapper.py. Denoised pyrosequences were then processed using the script inflate_denoiser_output.py in order to generate clusters of similar pyrosequences and to define one representative pyrosequence for each cluster (centroid) as well as non-clustering sequences (singletons). A new file was created containing singletons and cluster centroids inflated according to the original cluster size at the species level 97% similarity.

5.2.9 Identification of T-RFs

Only reads ranging between 300 and 500 bp selected with QIIME were used in the homemade software for pyrosequencing terminal-restriction fragments identification (PyroTRF-ID; see CHAPTER 3). Briefly, mapping of the sequences was performed using Burrows-Wheeler Aligner's Smith-Waterman Alignment (BWA-SW) (*38*) against the Greengenes database (*39*). Digital T-RFLP profiles were produced *in silico* using the aligned sequences and the palindromic GGCC sequence specific for the enzyme *Hae*III. Both experimental and digital T-RFLP profiles (eT-RFLP and dT-RFLP, respectively) were synchronized using autocorrelation algorithms based on dominant terminal-restriction fragments (T-RFs) so as to correct possible experimental fragment length drift (*40*, *41*). A consistent match was then obtained between experimental T-RFs (eT-RFs) and digital T-RFs (dT-RFs), and the putative corresponding sequence affiliation of eT-RFs representative of the clusters defined above was thus obtained.

5.2.10 Comparison of bacterial communities from Zones I and II

To assess differences among the BCS, results of the pyrosequencing analysis were annotated against the Greengenes database (*39*) on the MG-RAST online server (http://metagenomics.nmpdr.org/) (*42*). Hits with an e-value $< 10^{-5}$ and with an alignment length > 100 bp were used for the annotation with 97% minimum identity cutoff. Statistical comparison of BCS from Zones I and II was carried out using the STAMP software (*43*).

5.3 Results

5.3.1 Hydrochemical conditions in the Lyss aquifer

The investigated fraction of the Lyss aquifer was composed of two superimposed lithological layers. The lower one displayed a high hydraulic conductivity (1.10^{-3} ms^{-1}) and oxic conditions. The upper one displayed very low (5.10^{-5} ms^{-1}) hydraulic conductivity and was mainly anoxic. The source zone of PCE contamination was located in the upper layer. Samples were collected during four sampling campaigns and were labeled accordingly (A to D).

Hydrochemical conditions notably diverged between the two lithological layers, as observed in groundwater samples collected during the last sampling campaign. Groundwater samples originating from the lower layer were highly comparable in terms of chemical composition. They were clearly under oxic conditions (DO above 2mg/L), with NO_3^- concentrations above 15 mg/L, SO_4^{2-} concentrations above 18 mg/L, and Fe^{2+} and Mn^{2+} below detection limit (Table S1). In Zone I, a vertical gradient of redox conditions was present in the aquifer, going from clearly oxic in the deeper and coarser section to strongly reducing near the surface. Fe^{2+} and Mn^{2+} were detected in the most reduced fraction of the layer (respectively 6.2 and 0.83 mg/L in well P4), as well as high CH_4 concentrations (11.5 mg/L in well P4), whereas NO_3^- and SO_4^{2-} were below detection limit and at 3.1 mg/L, respectively. The same gradient of redox conditions was observed further downstream in Zone II, but reached less reducing conditions than in Zone I at comparable depths. Indeed, CH_4 was not detected in well P16, NO_3^- was below detection limit, whereas SO_4^{2-} concentration was still above 18 mg/L. Fe^{2+} and Mn^{2+} were present at low concentrations (0.1 mg/L for both). CH_4 and ethane were detected in monitoring well ML9, located in the upstream part of Zone II close to the tarred surface area. However, the concentrations of these two compounds were respectively 300 and two times lower in well ML9 compared with well P4 reaching a similar depth (3.6 and 4 m depth, respectively) in Zone I. CE concentrations were distributed along the same vertical gradient as described above for redox species. Total CE concentrations were lowest in the lower lithological layer and reached high values close to the surface. In Zone I, all CEs were detected, whereas no VC was detected in Zone II (Figure 3). The proportion of cDCE among total CE concentrations (%cDCE) was increasing together with the distance to the source zone, reaching 88% in well P16 (Figure 3, Table S2). In well P17, concentrations of cDCE were higher than in P4 where total CE concentrations reached their maximal value (Table S2). Furthermore, ethene was detected in Zone I exclusively. In Zone II, only ethane was detected in well ML9.

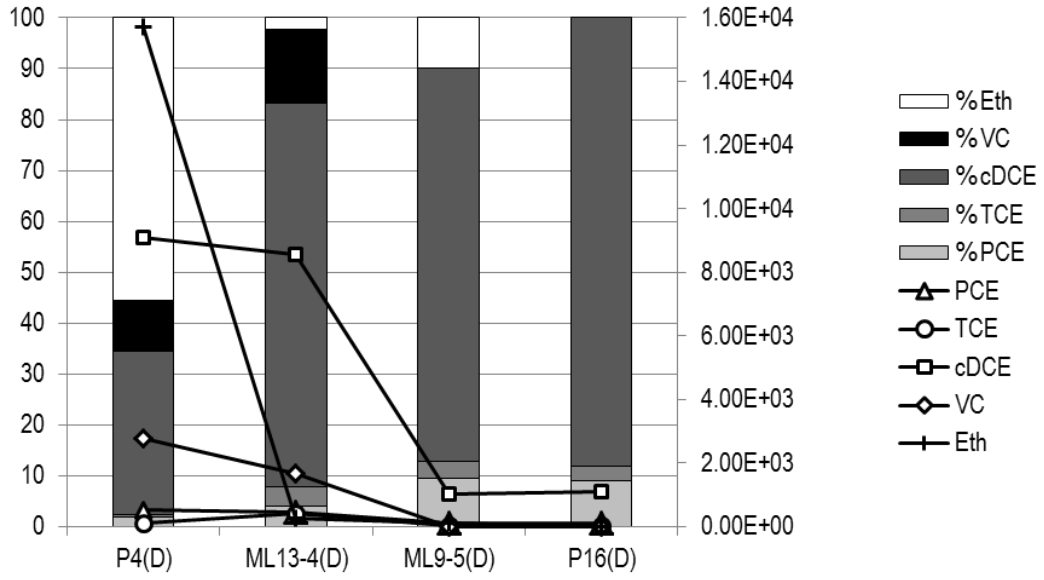


Figure 3. Evolution of the proportions of chlorinated ethenes (CEs) along a longitudinal transect through the contaminated aquifer. Investigated monitoring wells all target the same depth with similar geological properties. CE proportions (left y axis) were calculated among the cumulated concentrations (nM; right y axis) of CEs, ethene, and ethane. Eth and %Eth refer to both ethene and ethane cumulative concentrations and proportions, respectively.

5.3.2 Correlations between bacterial community structures and environmental variables

Reliable CE quantification results were obtained for the last sampling campaign only, due to a technical failure in the quantification process for samples of the other campaigns. Accordingly, and since CEs were presumed to be putative important environmental variables in the Lyss test site, analysis of the relationships between BCS and environmental variables was carried out using data from the last sampling campaign only. The MFA plot showed that environmental variables were oriented in the direction of two axes that can be interpreted as two main ecological gradients present in this section of the aquifer (Figure 4A). CEs and DO were aligned on one axis whereas variables related to oxidation-reduction state ("NO_3^-", "SO_4^{2-}", "Fe^{2+}", "Mn^{2+}", "CH_4", "Eh") were aligned on another one. With the exception of "%PCE" which was positively correlated with "DO", all other CEs were anti-correlated with this variable. "DO" and "%cDCE" displayed the two highest RV correlation coefficients with BCS (RV = 0.29, $p < 0.001$, and RV = 0.28, $p < 0.001$, respectively; Table 2). More generally, variables indicative of CE concentrations ("PCE", "TCE", and "cDCE") displayed slightly significant RV correlation coefficients with BCS, except "VC" which was not significant. Variables indicative of CE proportions were globally more significantly correlated with BCS, except "%VC", which displayed only a slightly significant RV correlation coefficient (RV = 0.20, $p = 0.044$; Table 2). The variable "Eh" representing

the oxidation-reduction potential measured on-site, was the third most strongly correlated with BCS (RV = 0.27, $p < 0.001$). It was positively correlated with oxidized electron acceptors such as "NO_3^-" and "SO_4^{2-}", and negatively correlated with reduction products such as "Fe^{2+}", "Mn^{2+}", "Sulfide", and "CH_4". All these variables were in turn moderately to highly significantly correlated with BCS, except "Sulfide" which was only slightly significant. The reduction products variables grouped together with the variables "Ethene", "Ethane", "%Eth", "CH_4", and "DOC".

The MFA plot shows a set of samples forming a dense cluster (Figure 4B, C1) originating from the lower lithological layer, and characterized by high DO concentrations and high proportion of PCE, as shown in Figure 4A. Samples R3 and R4 were exceptions, as they originated from the upper lithological layer, but in the uncontaminated area. A looser cluster was composed of a second set of samples (Figure 4B, C2) originating from (i) all depths of the upper lithological layer of Zone II, and (ii) the deepest fraction of upper layer of Zone I. Finally, a loose cluster was composed of 5 samples (Figure 4B, C3) originating from the upper lithological layer of Zone I at around 4 m depth and from the same layer in Zone II, but very close to the surface, between 2 and 4 m depth. This cluster was characterized by the most reducing conditions, with highest concentrations of CH_4, ethane, Fe^{2+}, Mn^{2+} and, for samples from monitoring wells ML15-4 and P4, with the highest concentrations of ethene. Samples from this cluster were also characterized by the highest DOC concentrations.

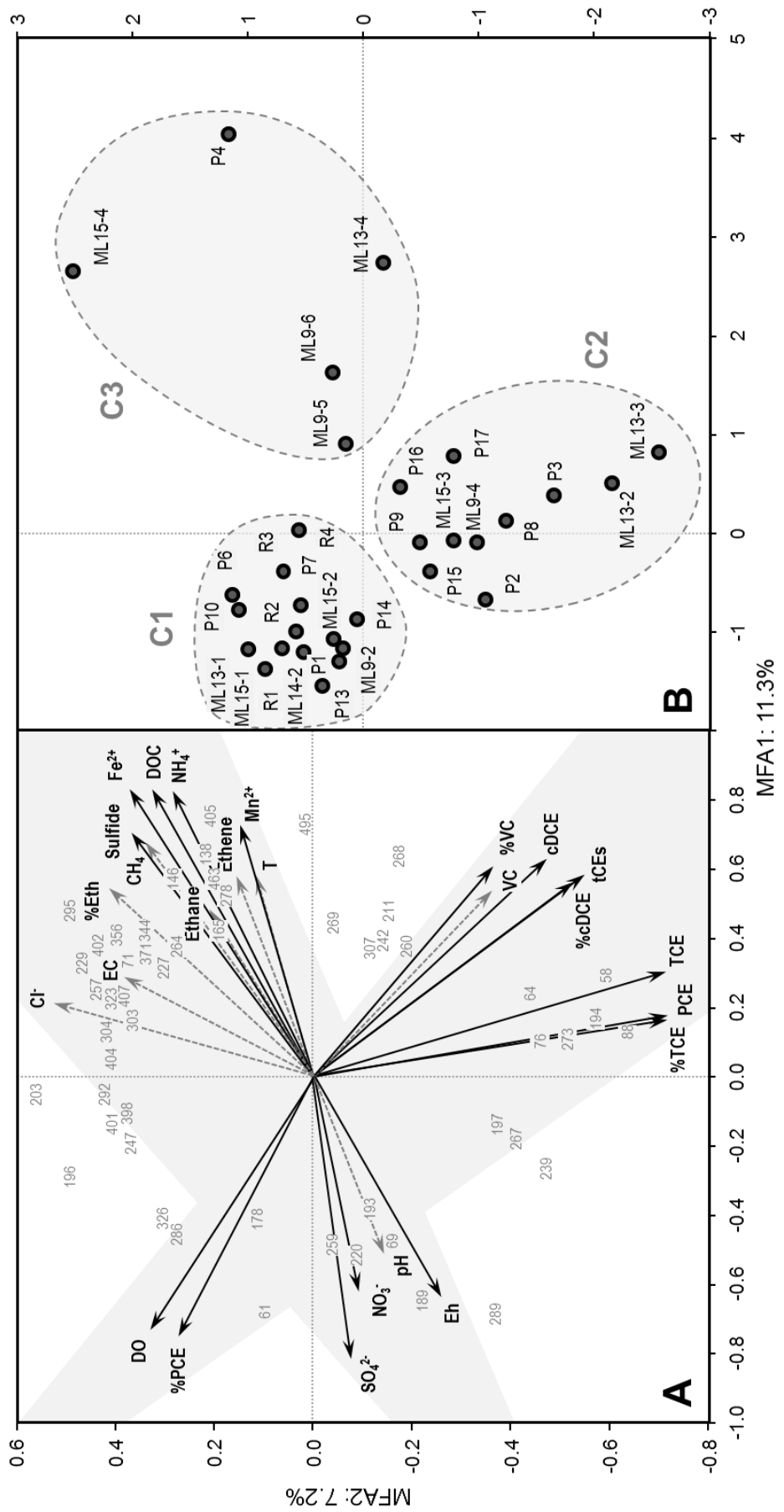


Figure 4. Multifactorial Analysis (MFA) carried out on environmental variables and bacterial community structures from the last sampling campaign, showing the first two principal components. Note that scales of the plots are adapted to facilitate readability. **A**: Significant and non-significant environmental variables are shown with black and dashed grey arrows, respectively. T-RFs significantly correlated with MFA axes are displayed with grey numbers indicating their size in bp. Two virtual axes on which environmental variables are aligned are shown with gray shadow to help interpreting the plot. **B**: MFA plot of the samples. Clusters C1 to C3 were arbitrarily delineated and denote the origin of samples in the aquifer related to the redox state. Note that samples from well ML14 have not been computed because they were lacking DOC values.

Table 2. RV correlation coefficients calculated between the bacterial community dataset and each environmental variable.

Variable	RV	p-value[a]	Significance
T[b]	0.20	0.054	ns
pH	0.17	0.259	ns
EC[c]	0.16	0.558	ns
Eh[d]	0.27	<0.001	***
DO[e]	0.29	<0.001	***
DOC[f]	0.22	0.008	**
Fe^{2+}	0.25	<0.001	***
Mn^{2+}	0.22	0.007	**
Cl^-	0.16	0.523	ns
Sulfide	0.21	0.029	*
NH_4^+	0.24	0.002	**
NO_3^-	0.21	0.015	*
SO_4^{2-}	0.22	0.010	*
PCE	0.21	0.033	*
%PCE	0.26	<0.001	***
TCE	0.21	0.022	*
%TCE	0.23	0.004	**
cDCE	0.20	0.043	*
%cDCE	0.28	<0.001	***
VC	0.19	0.136	ns
%VC	0.20	0.044	*
tCEs	0.20	0.043	*
Ethene	0.17	0.318	ns
Ethane	0.19	0.076	ns
%Eth[g]	0.17	0.302	ns
CH_4	0.18	0.189	ns

[a] The p-values were calculated using a permutation test with 1000 permutations. ***: highly significant; **: significant; *: moderately significant and ns: non-significant.
[b] Temperature.
[c] Electrical conductivity.
[d] Oxidation-reduction potential.
[e] Dissolved oxygen.
[f] Dissolved organic carbon.
[g] Cumulated proportion of both ethene and ethane concentrations among total concentration of chlorinated and non-chlorinated ethenes and ethanes in the samples.

5.3.3 Analysis of the bacterial communities of the upper lithological layer

Clustering of the bacterial communities from the upper anoxic lithological layer resulted in 4 clusters (Figure 5). Representative T-RFs of each cluster were designated using the IndVal analysis (Figure 5, Table 3), and some of the representative T-RFs were successfully identified using the PyroTRF-ID software (Table 4). Cluster cl1 was composed of samples originating exclusively from the top fraction of the upper lithological layer of Zone I (3.5 to 4 m below surface).

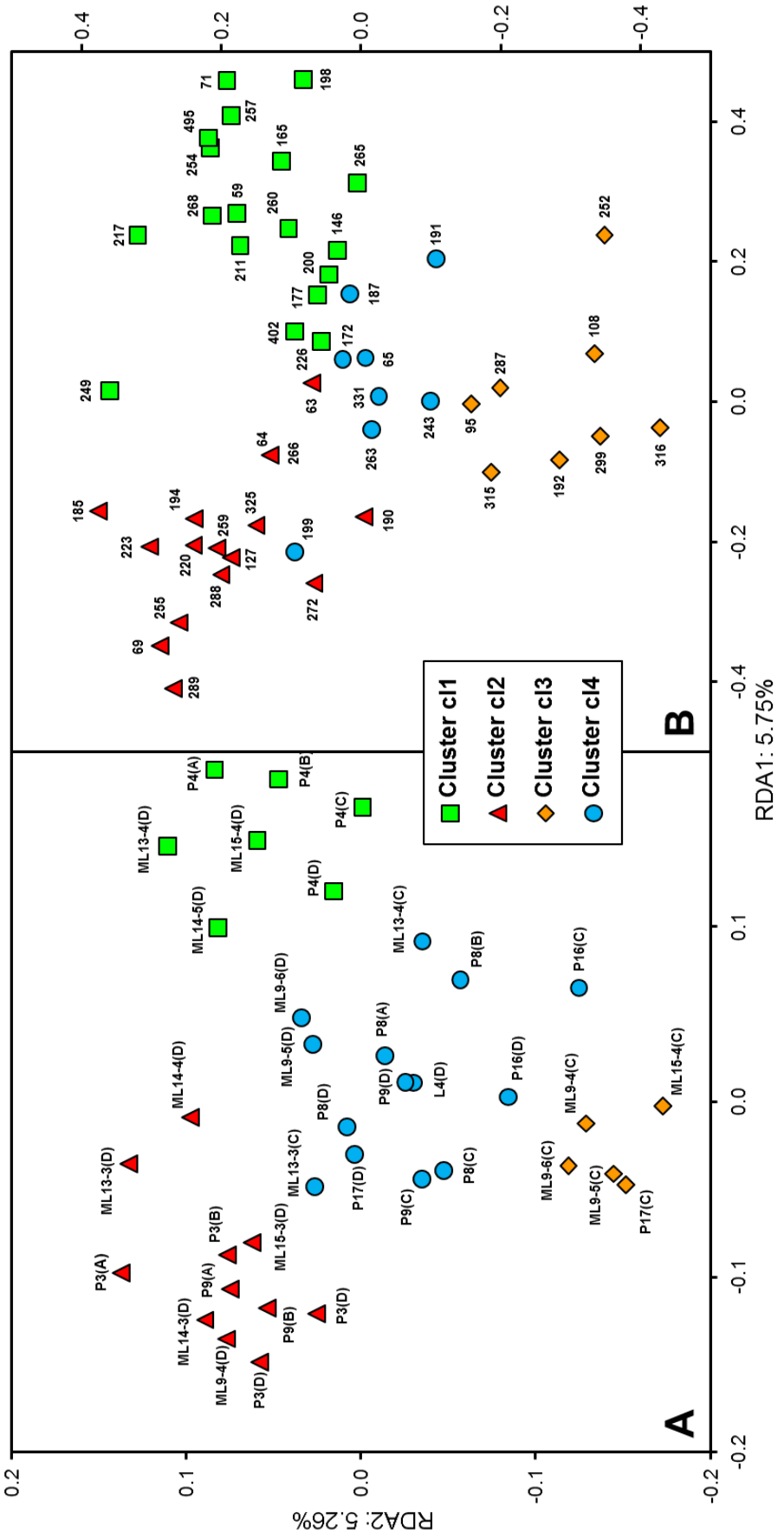


Figure 5. Superposition of (**A**) a PCA plot and clustering of all samples obtained from the upper anoxic geological layer of the Lyss aquifer, and (**B**) a plot of IndVal representative T-RFs of each cluster (see Table 3) in the same reduced space. The same cluster symbols are used for both samples and T-RFs. The reduced space of the PCA plot represents 11% of the total variance of the multifactorial space. Clusters were built using Ward's minimum variance agglomerative hierarchical clustering method on scores of a partial redundancy analysis (RDA) in which the potential effect of the pumping flow rate was removed. Samples are labeled according to their monitoring well followed by a reference of their sampling date (A to D). The T-RFs are indicated by their size (bp) only, for readability purpose.

Bacterial communities from all groundwater samples collected in well P4 (sampling campaigns A to D) were found in this group. The IndVal analysis designated 18 representative T-RFs for this cluster, 10 of which were successfully identified. Among them, only T-RF 165 was affiliated to an OHRB, namely *Dehalococcoides* sp. This T-RF was present in monitoring wells in the upper lithological layer of Zone I and was never detected in Zone II. Other T-RFs were identified as belonging to the families *Rhodocyclaceae*, *Desulfobulbaceae*, *Comamonadaceae*, *Syntrophaceae*, *Catabacteriaceae*, as well as to the genus *Geobacter* and to the *Sphingobacteriales* order. Cluster cl2 was composed of samples taken also from Zone I, but from a deeper fraction of the overlaying surface geological deposit, in which less reducing conditions prevailed. All groundwater samples collected in well P3 were grouped in this cluster, together with some groundwater samples from downstream wells (P9(A), P9(B), and ML9-4(D)). Sixteen representative T-RFs were designated for this cluster, of which only T-RF 194 was successfully affiliated to the family *Comamonadaceae*. Cluster cl3 was mainly composed of samples originating from Zone II taken from the upper lithological layer (ML9-4(C), ML9-5(C), ML9-6(C), P17(C)), but also contained one sample originating from the upper lithological layer of Zone I (ML15-4(C)). A single T-RF, out of 8 representative T-RFs of this cluster, was identified as a member of the order *Bacteroidales* (T-RF 252). Cluster cl4 was composed of samples originating from Zone II, except two samples from Zone I (ML13-3(C), ML13-4(C)). The two samples from well P16 were found in this cluster, together with samples from wells P8, P9 and P17. Eight representative T-RFs were designated for this cluster, 3 of which were affiliated to the family *Desulfobulbaceae* (T-RF 187), the genus *Acidovorax* (T-RF 191), and to *Geobacter uraniireducens* (T-RF 199).

Table 3. IndVal indicative T-RFs of each cluster defined based on the T-RFLP profiles of samples collected in the upper anoxic geological layer of the Lyss aquifer, showing T-RFs that are both more frequent and more abundant in one cluster as compared with others.

T-RF	Cluster	IndVal	p-value[a]
71		1	0.001
257		0.90	0.001
495		0.70	0.001
165		0.70	0.001
198		0.67	0.004
211		0.66	0.002
265		0.64	0.001
254	Cluster cl1	0.57	0.001
217		0.56	0.006
226		0.55	0.007
260		0.54	0.003
59		0.53	0.006
249		0.53	0.005
200		0.45	0.033
177		0.38	0.018
402		0.37	0.029
268		0.36	0.030
146		0.36	0.024
289		0.81	0.001
185		0.73	0.001
63		0.66	0.001
69		0.65	0.001
223		0.65	0.002
259		0.56	0.006
64	Cluster cl2	0.55	0.005
127		0.55	0.001
190		0.51	0.011
255		0.51	0.007
194		0.47	0.008
325		0.46	0.004
266		0.45	0.038
220		0.44	0.020
288		0.42	0.026
272		0.34	0.049
316		0.62	0.001
252		0.52	0.020
299	Cluster cl3	0.48	0.011
192		0.48	0.002
315		0.42	0.010
108		0.35	0.027
95		0.33	0.030
287		0.30	0.021
187		0.52	0.027
191		0.47	0.018
199	Cluster cl4	0.41	0.042
263		0.36	0.020
331		0.36	0.022
243		0.32	0.025
65		0.32	0.042
172		0.30	0.048

[a] Significance of the IndVal value indicated by the p-value calculated *a posteriori* with 999 permutations.

Table 4. Identification of IndVal representative T-RFs of defined clusters using PyroTRF-ID software. Only those T-RFs that were successfully identified are shown.

eT-RF[a]	Cluster	Sample	Relative abundance of eT-RF in sample[b]	Corresponding dT-RF[c] (bp)	Dominant phylotype[d]	Normalized SW score[e]
71	cl1	P4(B)	0.4	71	*Rhodocyclaceae* (29.4%; 5/17)	0.70-0.83
165	cl1	P4(D)	0.3	164	*Dehalococcoides* sp. (100%; 22/22)	0.55-1
177	cl1	P4(D)	0.3	177	*Geobacter* sp. (50%; 5/10)	0.63-0.78
198	cl1	P4(B)	0.8	197	*Desulfobulbaceae* (50.0%; 12/24)	0.61-0.95
200	cl1	P4(B)	2.2	199	*Geobacter* sp. (55.6%; 6/11)	0.94
211	cl1	P4(B)	1.6	211	*Comamonadaceae* (60.7%; 88/145)	0.66-0.96
					Limnohabitans sp. (37.9%; 56/145)	0.698-0.96
217	cl1	P4(D)	5.6	217	*Rhodocyclaceae* (26.2%; 11/42)	0.60-1
249	cl1	P4(B)	2.1	249	*Catabacteriaceae* (100%; 9/9)	0.97
257	cl1	P4(B)	1.3	256	*Sphingobacteriales* (90.5%; 19/21)	0.84-0.91
268	cl1	P4(D)	0.3	268	*Syntrophaceae* (50.0%; 3/6)	0.69-0.87
194	cl2	P4(B)	0.6	194	*Comamonadaceae* (44.2%; 80/181)	0.72-0.93
					Rhodoferax sp. (16.0%; 29/181)	0.73-0.89
252	cl3	P4(B)	1.3	252	*Bacteroidales* (60%; 6/10)	0.45-0.95
187	cl4	P16(C)	0.3	187	*Desulfobulbaceae* (85.7%; 6/7)	0.74-0.80
191	cl4	P4(B)	1.9	192	*Comamonadaceae* (91.8%; 90/98)	0.76-1
					Acidovorax sp. (89.8%; 88/98)	0.81-1
199	cl4	P16(C)	0.3	199	*Geobacter uraniireducens* (76.9%; 10/13)	0.81-1

[a] Experimental Terminal-Restriction Fragment (eT-RF).
[b] Relative abundance (%) of the eT-RF in the sample used for PyroTRF-ID computation.
[c] Digital Terminal-Restriction Fragment corresponding to the eT-RF.
[d] Main phylotype contributing to the dT-RF. In brackets: relative contribution of the phylotype to the dT-RF; number of reads showing this affiliation among the total number of reads contributing to the dT-RF.
[e] Range of Smith-Watermann mapping scores of reads contributing to the dT-RF normalized by the sequence length.

5.3.4 Comparison of bacterial communities from Zones I and II

The information obtained from pyrosequencing with 4380 reads per sample in average that had a mean length of 372 bp was used to reconstruct the composition of the bacterial communities at the time of sampling. Significant differences were found between the relative abundances at the genus level between Zone I (monitoring well P4) and Zone II (monitoring well P16) at similar depth (Figure 6 & Figure S1, Table 5). The genera significantly more abundant in P4 were composed of strict and facultative anaerobes that were not detected or only accounted for a very low proportion of sequences in P16. Among them, the most abundant genera were, in decreasing order, the obligate anaerobic OHRB *Dehalococcoides* sp. (29.9% relative abundance), the facultative anaerobe *Dechloromonas* sp. (21.9%), sequences affiliated to the strict anaerobic *Geothermobacter* sp. (9.2%), *Spirochaeta* sp. (5.0%) which was exclusively composed of the strict anaerobic species *S. stenostrepta* (data not shown), the strict anaerobic facultative OHRB *Desulfuromonas* sp. (4.7%) composed of *D. michiganensis* (4.7%), the microaerophilic genus *Leptothrix* sp. (3.1%), the strict anaerobic sulfate reducer *Desulfobacter* sp. (2.4%), the strict anaerobic *Sulfurimonas* sp. (2.3%) which was exclusively composed of the nitrate reducer *S. denitrificans*, the facultative fermenter *Zoogloea* sp. (2.1%), and the cellular parasite *Coxiella* sp. (1.7%). Strict and facultative aerobes were found in significantly higher abundance in P16. The most abundant genera in P16 were, in decreasing order, the strict anaerobe *Geobacter* sp. (23.1%), the strict aerobe *Janthinobacterium* sp. (22.4%), the facultative aerobe *Rhodoferax* sp. (12.4%), the strict aerobe *Flavobacterium* sp. (9.5%), the facultative aerobe *Pseudomonas* sp. (7.2%), the strict anaerobic sulfate reducer *Desulfosporosinus* sp. (5.0%), the aerobic *Acidovorax* sp. (3.3%), the strict aerobe *Collimonas* sp. (1.6%), the strict anaerobic facultative OHRB *Desulfitobacterium* sp. (1.0%), the strict aerobic ammonium oxidizer *Nitrosomomas* sp. (0.7%), and the phototrophic cyanobacterium *Arthrospira* sp. (0.7%). Apart from the OHRB mentioned above, a noticeable difference was also detected in the abundance, although low, of *Dehalobacter* sp., which was detected in P4 (0.4%) and absent in P16 (Table 5).

Quantitative PCR (qPCR) provided additional information on the differences between Zones I and II (Table 5). The copy number of the *Dehalococcoides* 16S rRNA gene was more than 200 times higher in Zone I than in Zone II (Table 5). The gene *vcrA* was abundant in Zone I and below detection limit in Zone II whereas the gene *bvcA* was about four times more abundant in Zone I than in Zone II.

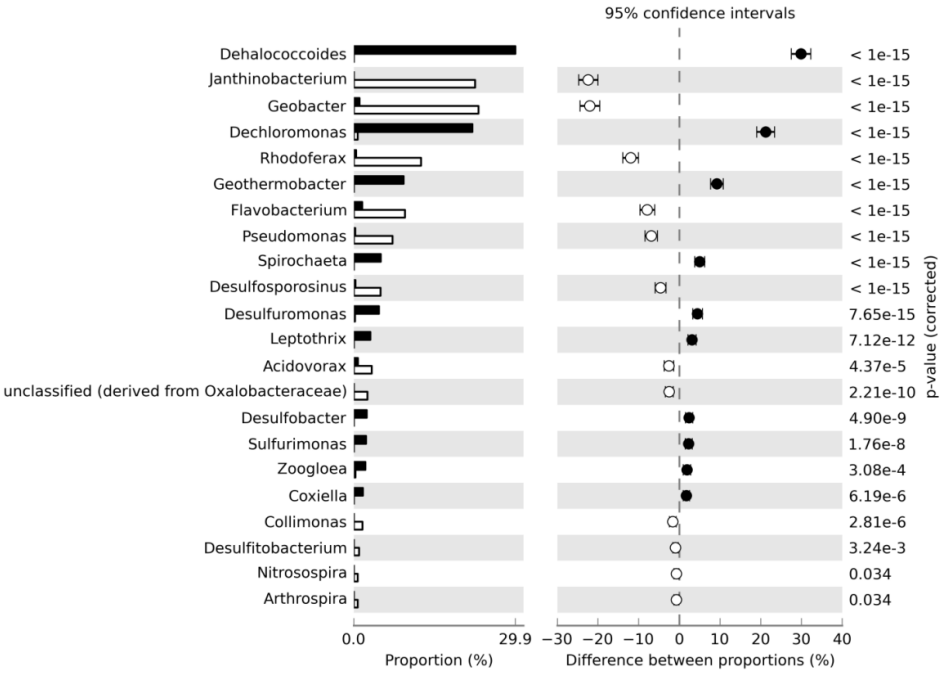


Figure 6. Significant differences at the $p < 0.05$ level between relative abundances of bacterial genera detected with pyrosequencing of bacterial 16S rRNA gene pool from Zones I (monitoring well P4, black bars and dots) and II (P16, white bars and dots).

Table 5. Organohalide respiration potential in Zones I (well P4) and II (well P16), defined using (i) quantification of *Dehalococcoides* sp. 16S rRNA, *vcr*A, and *bvc*A genes using quantitative PCR (qPCR) and expressed in gene copy numbers per liter (cp/L), (ii) relative abundance of organohalide respiring bacteria (OHRB) detected using the pyrosequencing approach and expressed in percentage of reads (%). The gene copy numbers were normalized to the total 16S rRNA gene copy numbers quantified in the samples using qPCR.

		P4	P16
qPCR (cp/L)	*Dehalococcoides* sp.[a]	1.4E+08	6.1E+05
	*vcr*A[b]	2.7E+07	bdl
	*bvc*A[c]	2.0E+06	5.6E+05
Pyrosequencing	*Dehalococcoides* sp.	29.9%	0.0%
	Desulfuromonas sp.	4.7%	0.2%
	Desulfitobacterium sp.	0.0%	1.0%
	Dehalobacter sp.	0.4%	0.0%

[a] Amplification efficiency = 0.963; R^2 = 0.976; detection limit = 7.7E+04 copies/L
[b] Amplification efficiency = 0.856; R^2 = 0.996; detection limit= 4.1E+03 copies/L
[c] Amplification efficiency = 0.924; R^2 = 0.977; detection limit = 5.3E+05 copies/L

5.4 Discussion

5.4.1 OHR was only detected in the vicinity of the source zone

CE contents

Detection of cDCE and VC, together with large amounts of ethene and ethane (accounting together for more than 50% of the molar concentration of cumulated chlorinated and non-chlorinated ethenes and ethanes) indicated that complete organohalide respiration (OHR) of CEs was taking place in the upper layer of Zone I, close to the contamination source of the aquifer. In the lower lithological layer, CE concentrations rapidly decreased with depth and no ethene was detected any more. At the bottom of this layer, only PCE was found at rather low concentrations. Dissolved oxygen concentration (DO) above 2 mg/L in this layer was inhibitory to OHR (*44*), and therefore TCE and cDCE present in the lower layer were probably not produced there. Concentrations of PCE detected upstream of the test site in the same layer were either close to detection limit or below and no TCE and cDCE were detected. We therefore assumed that CEs detected in the lower layer of Zone I originated from the upper layer, where OHR was active. The carbon isotopic signatures of CEs corroborated this hypothesis (Pascale Ducommun, personal communication). The situation in the upper lithological layer in Zone II, around 15 m downstream in the contamination plume, contrasted with the situation in Zone I in that neither VC nor ethene was detected. In addition, compound specific isotope analysis indicated that cDCE detected in Zone II originated from upstream, suggesting that no OHR was occurring in Zone II and not much between Zone I and II either (P. Ducommun, personal communication).

Molecular results are in accordance with the distribution pattern of CEs

The difference in OHR between Zones I and II was corroborated by molecular results. In Zone I, high copy numbers of *Dehalococcoides* sp. 16S rRNA and vinyl chloride reductive dehalogenase (*vcr*A) genes were detected. To date, cDCE and VC have been shown to be reduced only by bacteria belonging to the genus *Dehalococcoides*, and *vcr*A has been shown to be responsible for the reduction of both cDCE and VC (*45*). Therefore the detection of large numbers of both *Dehalococcoides* sp. 16S rRNA and *vcr*A gene copies was a clear indication that the potential for the reduction of lower CEs was present in Zone I and the presence of VC and ethene an indication of the expression of these genes. In contrast, the *Dehalococcoides* sp. 16S rRNA gene was approximately 200 times less abundant in Zone II, and the *vcr*A gene was below detection limit. The *bvc*A gene encoding a putative VC reductive dehalogenase responsible for the degradation of VC (*46*) was also detected in higher proportion in Zone I as compared with Zone II, although the difference was less marked.

The same trend was also observed in the proportion of *Dehalococcoides* sp. in the bacterial community obtained with the pyrosequencing method and was consistent with the low abundance of this obligate OHRB in Zone II. Pyrosequencing results revealed that the bacterial communities in Zones I and II were composed of different facultative and obligate OHRB. The two most abundant OHRB, *Dehalococcoides* sp. and *Desulfuromonas michiganensis*, were exclusively detected in Zone I. *Desulfuromonas michiganensis* has been previously described as a PCE-to-cDCE dechlorinating bacterium able to use Fe(III) and S^0 as alternative electron acceptors. Isolates of this species were unable to use NO_3^- or SO_4^{2-} as electron acceptors (*47*). The presence of this species in relatively high proportion (4.7%) in the zone of OHR together with *Dehalococcoides* sp. suggests that it was responsible for at least part of the reductive dechlorination of higher CEs, whereas *Dehalococcoides* sp. was responsible for the OHR of lower CEs mainly. *Dehalobacter* sp., although detected in Zone I, represented only a small proportion of the community (0.4%) and was probably only marginally involved in the reduction of PCE and TCE. In Zone II, the low proportions of potential OHRB together with the absence of VC and ethene indicated that OHR was not complete and not a major process. Microbiological results however suggest that the reductive dechlorination of PCE and TCE was occurring, at least occasionally. Indeed, two facultative OHRB genera, *Desulfitobacterium* sp. and *Desulfuromonas chloroethenica*, were detected in Zone II and were hypothesized to use CEs only when conditions became favorable for OHR, whereas other electron acceptors could be used as well. Both genera have been shown to be able to dechlorinate up to cDCE. Their ability to use other electron acceptors than organohalides may bring them a competitive advantage over obligate OHRB, especially in constantly fluctuating environments. *Desulfitobacterium* sp. is well-known for its versatile metabolism and its wide range of potential electron acceptors, including S^0, Fe(III), fumarate, and NO_3^- (*48*). *Desulfuromonas chloroethenica* has been shown to use Fe(III) and S^0 as alternative electron acceptors in laboratory conditions (*49*). Molecular investigations thus did not indicate a complete repression of OHR in Zone II, since facultative OHRB were present. However, the drastic differences observed in the *Dehalococcoides* sp. 16S rRNA, *vcr*A, and *bvc*A gene copy numbers, as well as in the relative abundance of *Dehalococcoides* sp. 16S rRNA gene in pyrosequencing data were a strong indication that OHR of lower CEs in Zone II was not active or not conducted to such an extent that all CEs could be removed. Analyses based on RNA extracts could not be carried out here, but they would provide further clues concerning the activities of OHRB and metabolic genes.

5.4.2 Causes of the absence of OHR in Zone II

Investigation of the relationships between environmental variables and BCS

Multivariate statistical investigation of the relationships between environmental variables and BCS showed that dissolved oxygen (DO) was the most strongly correlated environmental variable. The influence of oxygen over BCS was explained by the pronounced oxygen zonation in the aquifer. The fact that oxygen is the first utilized

electron acceptor in the ecological succession of TEAPs (*6*) explained the negative correlations found between "DO" and the oxidation products of other TEAPs, such as Fe^{2+}, Mn^{2+}, and S^{2-}, and the positive correlation found with electron acceptors such as NO_3^- and SO_4^{2-}. The environmental variable showing the second largest correlation with BCS was the proportion of cDCE ("%cDCE"), which was reflecting a relative accumulation of cDCE. This result indicated that fluctuations of BCS were intimately related to variations of "%cDCE". In the Lyss aquifer, "%cDCE" fluctuated as a function of (i) depth reflecting the amount of DO, and (ii) distance to the source zone.

Oxidation-reduction conditions

Similarly, the pattern of redox species fluctuated as a function of depth, especially in the upper lithological layer, where the vertical gradient of redox species was observed. In Zone I, ethene and CH_4 concentrations increased progressively towards the top of the upper layer, following in this sense the gradient of redox species. The latter changed together with the distance to the source, as the upper lithological layer in Zone II was reaching a less reduced state as compared with Zone I, which is probably a consequence of the thinning of this geological layer from Zone I to Zone II (Figure 6). Indeed, at comparable depths, SO_4^{2-} was present in higher concentrations in Zone II. In monitoring well P16, no NO_3^- was detected during the last sampling campaign, however low concentrations were detected during the previous campaigns (data not shown). In addition, concentrations of NO_3^- in the monitoring well just underneath were in the same range as concentrations found in the lower lithological layer, whereas they were much lower at the same depth in Zone I. SO_4^{2-} concentrations in well P16 were also much higher as the ones observed at similar depth in Zone I, indicating that SR was less active. Furthermore, no CH_4 was detected in monitoring well P16, showing that methanogenesis was not taking place in this zone, whereas high concentrations were observed in Zone I at the same depth. SO_4^{2-} concentrations were even higher in well P17 located above P16, in a zone a priori as distant from the lower lithological layer as was P4. Several hypotheses for the more oxidized state in P17 were formulated. First, as this well was the closest to the lawn surface, it may have been influenced by surface infiltration. Second, preferential fluxes bringing oxidized water from the lower lithological layer cannot be ruled out. Monitoring well ML9-5, located in Zone II, represented an interesting intermediate between the highly reducing conditions of Zone I and the less reducing conditions around the monitoring wells P16 and P17 in Zone II. Indeed, neither VC, nor ethene was detected in this well, but low amounts of CH_4 were observed, indicative of a reduced environment. The DOC concentrations were lower than the ones detected in Zone I, but higher than the ones detected around P16 and P17. On the MFA plot, samples from well ML9-5 and the overlying one (ML9-6) were positioned between samples of similar depth from the rest of Zone II and the most reduced samples of Zone I, in which OHR was taking place.

Two joint environmental mechanisms are proposed to explain the less reducing conditions encountered in the upper lithological layer in Zone II: the lower availability of electron donors and the higher proneness to recharge

with oxidized water from the lower geological layer. The amount of organic matter (DOC) was lower in Zone II compared with Zone I, and, on the MFA plot, "DOC" was characterizing the samples with the strongest reducing conditions, together with redox-sensitive species. Substrates and electron donors for TEAPs originate from the decomposition of organic matter by fermenting and syntrophic bacteria (*7*). Depletion of organic compounds generally results in less available electron donors and increased competition between TEAPs. On the contrary, high organic matter content enhances the reduction of electron acceptors, rapidly leading to highly reducing conditions if the flow of oxidized electron acceptors is slow. During rainy events or snow break the hydraulic pressure in the lower lithological layer increases, as shown by pressure probes (P. Ducommun, personal communication), and pushes up the groundwater level, probably resulting in a recharge of the upper layer with oxidized water containing O_2, NO_3^-, and SO_4^{2-}. The conditions in the upper lithological layer in Zone I render this zone less susceptible to re-oxidation by recharge from underneath. Since the upper layer is thicker in Zone I and more electron donors are available for reduction of unfavorable electron acceptors present in the oxidized water, it is less susceptible to re-oxidation by recharge from underneath than Zone II. In the latter, contrastingly, electron acceptors such as O_2, NO_3^-, and SO_4^{2-} may have periodically reached the top fraction of the upper layer, due to lower availability of electron donors and a lower thickness of the upper layer. Furthermore, infiltration from the lawn surface cannot be fully ruled out as an explanation to the less reduced state of the upper layer in Zone II.

Clusters of BCS

The clustering of BCS showed that samples clustered according to their location in the aquifer, that is, either in the upper or lower lithological layer, and either in Zone I or II. In particular cluster cl1 contained only samples from Zone I in the vicinity of the source, taken close to the surface. This location was among the most severely contaminated of the investigated area, the most active in terms of OHR as indicated by the higher ethene concentrations, and the most reduced, as shown by the higher concentrations of Fe^{2+}, Mn^{2+}, S^{2-}, and the lower concentrations of NO_3^- and SO_4^{2-}, as compared to other investigated zones of the aquifer. In addition, highest DOC concentrations were observed in this location. These observations were confirmed from a microbiological point of view by the fact that T-RF 165, which was statistically representative of this cluster, was unambiguously identified as *Dehalococcoides* sp. Furthermore, as mentioned above, pyrosequencing revealed that the genus *Dehalococcoides* was dominant in one sample (P4(D)) collected in this zone. In a former study (*3*), the abundance of *Dehalococcoides* sp. was related to the amount of organic matter. Our results based on quantitative PCR tend to the same conclusion, as the genus was found to be much more abundant in the zone with higher DOC content. This result is based on two samples only, and is not sufficient to conclude on the correlation. However, the fact T-RF 165, which was affiliated to *Dehalococcoides* sp., was clearly correlated with DOC (Figure 4A) was an additional clue in this sense. Nevertheless, it is not yet clear whether the differences between the two investigated zones in terms of *Dehalococcoides* sp. abundance arose from the difference in DOC content

or from the difference in redox conditions, as reflected by the anti-correlation of vectors "Eh" and "DOC" in Figure 4A. Probably both features were combined. Other representative T-RFs from Zone I were affiliated to sulfate-reducing bacteria (SRB) of the *Desulfobulbaceae* family, and to metal-reducing bacteria of the genus *Geobacter*. The fact that they were indicative of BCS from the most reduced investigated zone, together with the highest measured concentrations of Fe^{2+} and Mn^{2+}, and the lowest SO_4^{2-} concentrations, suggested that metal reduction and SR were active concomitantly to OHR. Concomitant reductions of Fe(III) and sulfate has been suggested previously when electron donors are not limited (*50*), as seems to be the case in the upper layer top fraction in Zone I.

PyroTRF-ID allowed identifying only part of T-RFs representative of cluster cl3 and cl4, which were mainly composed of BCS originating from Zone II in the upper lithological layer. Additional pyrosequencing data from samples of these clusters would be needed.

5.4.3 Bacterial processes in Zones I and II based on bacterial community structure data

Results of pyrosequencing provided further insight into the differences between the upper layers in Zones I and II in terms of bacterial community composition. The two zones were indeed characterized by distinct bacterial communities. The dominant population in the bacterial community of Zone I was unambiguously related to OHR, as *Dehalococcoides* sp. was the most abundant genus and as the bacterial potential for complete OHR of CEs was covered in the upper layer, with *Desulfuromonas* sp. being another genus involved in OHR of PCE and TCE. The high organic matter content, together with its partial isolation from oxidized groundwater, made this zone a well-adapted habitat for the complete reduction of CEs. The abundance of *Dehalococcoides* sp. reflected the strongly reducing and strict anoxic conditions of its habitat. This was sustained by the fact that no strict aerobic bacterium, except the microaerophilic *Lepthothrix* sp., displayed a significant abundance in Zone I. *Dechloromonas* sp. was the second most abundant genus in this zone, and was mainly represented by the species *D. aromatica*. This species displays a versatile metabolism, with the ability to degrade aromatic compounds aerobically and anaerobically combined with the reduction of perchlorate among others (*51*). The presence of this contaminant was not assessed in the present study.

The upper lithological layer in Zone II was characterized by bacterial guilds typical of more oxidized environments, and no strict anaerobic genus was found significantly more abundant than in Zone I. On the contrary, strict aerobes were detected, such as *Janthinobacterium* sp. (*52*), *Flavobacterium johnsoniae* (*53*), *Collimonas* sp. (*54*), and *Nitrosospira* sp., described as a nitrifyer (*55*, *56*). Facultative anaerobes were found as well, most of them characteristic of nitrate-reducing conditions, such as *Pseudomonas fluorescens*, a facultative anaerobic nitrate reducer (*57*) and *P. putida*, a species able to perform aerobic nitrate reduction as well as nitrification (*58*, *59*). The only detected species of the genus *Acidovorax*, *A. avenae*, has also been described as a

nitrate reducer (*60*), with the ability to use chlorobenzenes as a carbon source in a polluted aquifer (*61*). Diverse metabolisms have been described for the genus *Rhodoferax*, which was abundant in sample P16. *Rhodoferax* species can be iron reducers (*62*, *63*), whereas others can grow heterotrophically as fermenting bacteria or with oxygen or nitrate as electron acceptors, and even as anoxygenic phototrophs (*62*, *64*). Since low concentrations of Fe^{2+} were detected in well P16 in the last sampling campaign, and low concentrations of NO_3^- in a previous one, both metabolisms were plausible for this genus. Finally, *Geobacter* sp. was the most abundant genus in the upper layer of Zone II. This genus also shows a versatile metabolism and is not restricted to fully anoxic environments, as indicated by its ability to grow with O_2 as final electron acceptor (*65*).

In addition to the competition with nitrate-reducing bacteria, the fact that cDCE was present in large concentrations in Zone II and that VC and ethene were not detected, suggested that these compounds were degraded along the path between Zones I and II. DO concentrations below the detection limit of the device used here can also exert an adverse effect on OHR (*44*, *66*), either by means of a toxic effect on OHRB and their enzymes (*67*, *68*) or via competition for the electron donors in favor of aerobic bacteria (*7*). According to Bradley et al. (*44*), the field standard detection limit of 0.1 to 0.5 mg/L may hinder concluding on the absence of aerobic oxidation DCE and VC. These authors have estimated that only highly reducing conditions such as SO_4^{2-}-reducing ones and methanogenesis can really be associated with the absence of aerobic metabolism. As explained above, if these conditions were reached in Zone II, they were punctuated by sporadic events of re-oxidation. During the re-oxidation of the zone, aerobic oxidation of lower CEs could be expected to occur, leading to the complete removal of VC. The conditions between the two Zones were not analyzed in this study, but if strict anoxic conditions were found, anaerobic oxidation may also represent a plausible reason for the absence of VC (*69*, *70*).

5.4.4 Conclusion: is MNA a solution for this site?

The conditions in the upper lithological layer between Zones I and II are not known, since no monitoring well targeting discrete depths was installed in this zone. Analysis of CEs and ethene indicated however that cDCE produced in Zone I or a little further reached Zone II, whereas VC and ethene did not. The conditions in the upper lithological layer may progressively have become slightly more oxidized as set out above. This could have led to inhibition of OHR and further production of VC and ethene. As mentioned above, the already produced VC and ethene could in addition have been degraded by anaerobic and perhaps even aerobic oxidation processes while migrating through the upper lithological layer from Zone I to Zone II (*44*, *70*). Hence, the structure and functioning of the Lyss aquifer turned out to be interesting from a bioremediation point of view, in that the conditions were met for a naturally-occurring sequential anaerobic/aerobic biodegradation of CEs (*71*). Complete natural sequential anaerobic/aerobic biodegradation of CEs has been previously reported for a site in Germany (*72*). In the Lyss aquifer, PCE and TCE were first reduced in the anoxic Zone I. The dechlorination products cDCE, VC, and ethene were possibly oxidized in the upper lithological layer that gets progressively

thinner and more oxidized, and that is finally replaced by the oxic Aareschotter layer where aerobic oxidation of the remaining cDCE could occur (*73*).

However, the fact that higher CEs also reached the lower lithological layer in Zone I, and were detected all along the contamination plume, although at low concentrations, counter-balanced the positive aspects of the aquifer lithological structure. PCE and TCE are strong oxidants and only the latter can be degraded in a co-metabolic way under oxic conditions, which requires auxiliary substrates that are lacking in the lower lithological layer. Therefore, PCE and TCE will probably persist in the lower lithological layer of the Lyss aquifer, and only dispersion and dilution phenomena will lead to a decrease of concentrations. It can also not be completely ruled out that part of PCE DNAPL was drained from the upper layer and migrated from the primary contamination plume downwards to accumulate on the aquitard, which would constitute a problematic long-term contamination source.

In this work, global insight of the aquifer structure and behavior led to appreciating the reasons for local variations in the fate of OHR. However, the major achievement was probably not the understanding of cDCE accumulation *per se*, but rather a change of paradigm in designing a remediation strategy. Indeed, accumulation of cDCE was first considered as the main disadvantage in the use of MNA at the present site. However, the hydrogeological functioning of the aquifer led to understanding that the major problem from a bioremediation perspective was the persistence of PCE and TCE, mainly in the lower lithological layer, which was completely oxic. Therefore, remediation strategies for this site should focus on the complete elimination of PCE and TCE, by enhancing OHR in Zone I through the injection of organic matter, before these compounds reach the lower lithological layer rather than on the accumulation of cDCE. The latter is indeed expected to be oxidized quickly after it reaches the oxic zone of the aquifer.

Acknowledgements

This project was carried out in collaboration with Pascale Ducommun and Daniel Hunkeler, Center for Hydrogeology of the University of Neuchâtel (CHYN). I warmly thank Pascale for her help in the field work and the understanding of the hydrogeological aspects of the contaminated site, Xanthippe Boutsiadou for the chemical analyses of groundwater, Jürgen Abrecht (Geotest AG) for test site sampling authorization, and Ioannis Xenarios for support and access to the Vital-IT high performance computing cluster of the Swiss Institute for Bioinformatics (Lausanne, Switzerland). Many thanks to François Gillet (Université de Franche-Comté, France) for his precious advice in multivariate statistics. I cheerfully thank Cyril Cicekli (CEL, EPFL, Switzerland) for the setup of the qPCR technique. I also acknowledge the decisive technical contributions of Julien Maillard, Guillaume Laulan and Amandine Parent (LBE, EPFL, Switzerland).

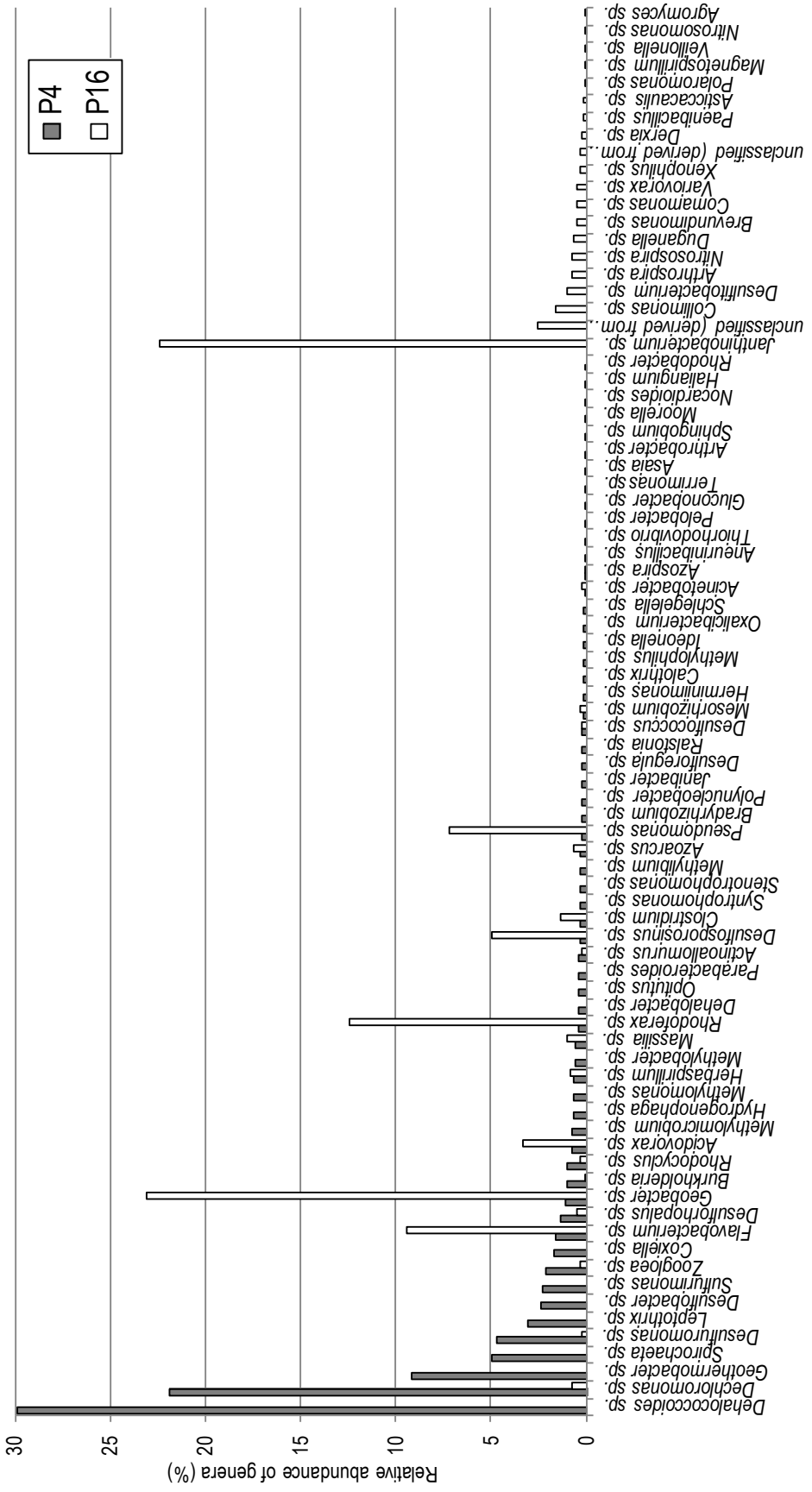


Figure S1. Relative abundances of all genera detected in Zones I (well P4) and II (well P16) using the pyrosequencing approach and expressed as a proportion to all reads (%). The genera were classified in decreasing order of their relative abundance in well P4.

Table S1. Groundwater characteristics during the last sampling campaign (D). The proportion of each chlorinated ethane (CE) was calculated among total molar concentration of chlorinated and non-chlorinated ethenes and ethanes.

Monitoring well	Mid-screen depth [mbs]	Temp[a] [°C]	pH	EC[b] [µS cm^{-1}]	DO[c] [mg l^{-1}]	DOC[d] [mg l^{-1}]	Fe^{2+} [mg l^{-1}]	Mn^{2+} [mg l^{-1}]	S^{2-} [mg l^{-1}]	Na^{+} [mg l^{-1}]	K^{+} [mg l^{-1}]	NH_4^{+} [mg l^{-1}]	Mg^{2+} [mg l^{-1}]	Ca^{2+} [mg l^{-1}]	Sr^{2+} [mg l^{-1}]	Cl^{-} [mg l^{-1}]	NO_3^{-} [mg l^{-1}]	SO_4^{2-} [mg l^{-1}]
P1	7.0	17.2	6.97	642	4.29	2.02	<0.05	<0.05	<0.01	11.8	<0.1	<0.1	14.9	111.0	<0.1	23.0	26.1	21.4
P2	5.9	16.4	6.97	603	2.43	1.99	<0.05	<0.05	<0.01	10.6	<0.1	<0.1	14.5	98.2	0.2	21.0	24.1	18.8
P3	4.9	16.4	6.98	620	0.57	2.37	<0.05	0.06	<0.01	10.6	1.8	1.8	13.6	103.3	0.2	25.2	12.4	14.7
P4	4.0	17.6	6.66	719	0.71	9.19	6.2	0.83	0.11	9.3	3.3	7.6	6.6	117.9	<0.1	33.6	<0.1	3.1
ML13-1	7.0	15.9	7.29	652	3.64	1.62	<0.05	<0.05	0.02	11.6	<0.1	<0.1	14.7	110.2	0.3	26.3	19.9	23.3
ML13-2	5.9	17.5	7.32	613	2.54	1.97	<0.05	<0.05	0.03	10.1	0.7	0.7	13.4	94.2	0.6	25.4	15.1	18.8
ML13-3	4.9	16.4	7.32	603	0.45	1.92	<0.05	<0.05	0.04	11.2	0.9	0.9	14.1	100.6	0.2	24.5	14.1	18.7
ML13-4	4.0	17.7	7.07	646	0.49	4.91	6.02	0.25	0.05	12.4	7.0	7.0	6.6	102.7	0.2	31.9	<0.1	12.8
ML14-2	6.5	16.8	7.14	655	3.31	n.a.	<0.05	<0.05	0.03	11.8	<0.1	<0.1	14.8	111.8	0.4	26.9	19.6	23.6
ML14-3	5.5	17.2	7.35	630	1.84	n.a.	<0.05	0.06	0.04	11.6	<0.1	<0.1	15.0	105.4	0.3	25.6	18.6	21.9
ML14-4	4.5	18.8	7.45	563	0.49	n.a.	<0.05	<0.05	0.03	12.7	4.9	4.9	9.1	89.3	0.4	24.9	10.8	20.3
ML14-5	3.5	18.8	7.00	930	0.56	n.a.	7.58	0.29	0.08	37.4	10.7	10.7	7.2	123.7	0.3	77.7	1.6	9.2
ML15-1	7.5	15.6	7.26	253	3.64	1.74	<0.05	<0.05	0.03	11.8	<0.1	<0.1	14.9	110.8	<0.1	27.9	20.1	24.3
ML15-2	6.5	15.7	7.20	656	3.55	1.63	<0.05	<0.05	0.03	11.9	<0.1	<0.1	14.9	111.8	0.7	27.5	20.0	24.0
ML15-3	5.5	16.0	7.26	604	0.54	1.73	<0.05	<0.05	0.03	12.7	0.4	0.4	13.5	99.5	0.7	26.7	14.0	20.1
ML15-4	4.5	17.1	7.30	702	0.56	5.21	5.97	0.20	0.10	20.3	7.6	7.6	9.0	103.5	0.4	44.4	<0.1	12.9
P6	5.9	n.a.	n.a.	n.a.	2.27	1.64	<0.05	<0.05	0.02	11.4	<0.1	<0.1	13.6	104.7	0.2	23.2	25.7	20.6
P7	5.9	18.4	7.28	615	2.98	1.47	<0.05	<0.05	0.02	11.8	<0.1	<0.1	14.6	108.9	0.2	25.7	26.8	22.6
P8	4.0	16.0	7.24	655	0.40	1.44	<0.05	<0.05	<0.01	9.9	<0.1	<0.1	11.0	88.9	<0.1	17.9	16.6	20.6
P9	4.6	16.4	7.31	558	0.42	1.33	<0.05	<0.05	<0.01	9.9	<0.1	<0.1	11.5	92.5	0.2	20.1	5.1	22.7
P10	5.9	15.7	7.24	663	3.28	1.67	<0.05	<0.05	<0.01	12.9	<0.1	<0.1	14.4	110.4	0.3	28.0	26.2	22.2
P13	7.6	14.8	7.22	683	3.74	1.53	<0.05	<0.05	<0.01	12.3	<0.1	<0.1	14.2	114.6	0.3	28.9	28.4	26.0
P14	6.4	14.8	7.24	653	2.93	1.46	<0.05	<0.05	<0.01	12.0	<0.1	<0.1	14.5	109.0	0.3	26.2	26.8	22.7
P15	5.1	15.0	7.30	636	2.09	1.52	<0.05	<0.05	0.03	11.7	<0.1	<0.1	14.1	105.8	0.4	24.9	24.6	23.3
P16	3.9	16.0	7.15	649	0.23	1.51	0.10	0.10	0.03	17.3	<0.1	<0.1	13.5	107.7	0.3	22.5	<0.1	18.0
P17	2.6	17.1	7.04	640	0.29	1.77	<0.05	0.64	0.03	15.4	<0.1	<0.1	12.1	105.5	0.1	21.5	<0.1	22.4
ML9-2	6.6	14.7	7.26	644	2.95	1.43	<0.05	<0.05	<0.01	12.0	<0.1	<0.1	14.8	110.4	0.3	27.7	19.8	23.4
ML9-4	4.6	16.3	7.27	576	1.13	1.46	<0.05	<0.05	<0.01	11.0	<0.1	<0.1	12.4	97.6	0.2	22.3	14.8	21.6
ML9-5	3.6	16.3	7.37	544	0.39	2.45	0.11	0.37	<0.01	10.5	<0.1	<0.1	10.6	89.0	0.2	21.0	6.4	22.6
ML9-6	2.6	18.6	7.07	652	0.51	2.45	2.00	0.59	<0.01	15.3	0.6	0.6	11.4	109.4	0.3	21.9	<0.1	16.1
R1	7.9	16.8	7.13	688	3.31	1.50	<0.05	<0.05	0.03	12.7	<0.1	<0.1	13.8	116.4	0.7	30.3	<0.1	25.1
R2	6.7	16.8	6.95	696	4.09	1.65	<0.05	<0.05	0.02	12.9	<0.1	<0.1	14.1	117.3	0.7	30.9	<0.1	24.6
R3	5.4	17.3	7.13	569	2.28	2.47	<0.05	<0.05	0.03	10.8	<0.1	<0.1	14.5	92.6	0.6	12.0	0.8	12.6
R4	4.2	18.1	7.18	588	0.47	2.97	0.10	<0.05	0.05	9.0	<0.1	<0.1	15.0	96.8	0.7	9.2	<0.1	13.5

[a] Temperature
[b] Electrical conductivity
[c] Dissolved oxygen
[d] Dissolved organic carbon

Table S2. Groundwater characteristics during the last sampling campaign (D). The proportion of each chlorinated ethane (CE) was calculated among total molar concentration of chlorinated and non-chlorinated ethenes and ethanes.

Monitoring well	PCE [µg l^{-1}]	TCE [µg l^{-1}]	cDCE [µg l^{-1}]	VC [µg l^{-1}]	Total CEs [µg l^{-1}]	Ethene [µg l^{-1}]	Ethane [µg l^{-1}]	Methane [µg l^{-1}]	%PCE	%TCE	%cDCE	%VC	%Eth
P1	16.6	<0.2	<0.2	<0.2	16.6	<0.5	<0.5	<0.4	100.00	0.00	0.00	0.00	0.00
P2	165.1	28.2	203.0	2.5	398.8	<0.5	<0.5	1.7	29.77	6.42	62.62	1.18	0.00
P3	504.4	92.2	1337.5	235.6	2169.7	51.7	12.5	1.5	12.90	2.98	58.54	16.00	9.59
P4	86.5	15.8	883.2	173.3	1158.8	433.7	7.7	11537.5	1.85	0.43	32.26	9.82	55.65
ML13-1	6.1	<0.2	<0.2	<0.2	6.1	1.2	<0.5	138.0	47.09	0.00	0.00	0.00	52.91
ML13-2	484.1	141.5	1101.6	117.0	1844.2	19.2	3.4	222.4	16.19	5.97	63.02	10.39	4.43
ML13-3	624.1	187.5	840.6	94.1	1746.3	3.2	1.2	118.5	24.25	9.20	55.86	9.70	1.00
ML13-4	76.2	55.1	829.5	103.5	1064.3	1.1	6.2	2.3	4.05	3.70	75.47	14.60	2.18
ML14-2	16.1	<0.2	<0.2	<0.2	16.1	<0.5	<0.5	2.6	100.00	0.00	0.00	0.00	0.00
ML14-3	121.6	16.1	45.5	<0.2	183.2	<0.5	0.7	54.7	54.39	9.09	34.82	0.00	1.70
ML14-4	158.4	34.1	217.2	7.7	417.4	<0.5	5.6	2.9	25.38	6.89	59.55	3.27	4.91
ML14-5	42.9	18.0	102.7	8.0	171.6	<0.5	<0.5	1.5	16.34	8.68	66.94	8.04	0.00
ML15-1	6.7	<0.2	<0.2	<0.2	6.7	<0.5	<0.5	64.3	100.00	0.00	0.00	0.00	0.00
ML15-2	6.7	<0.2	<0.2	<0.2	6.7	<0.5	<0.5	33.1	100.00	0.00	0.00	0.00	0.00
ML15-3	131.3	26.1	121.9	2.5	281.8	<0.5	1.0	0.0	34.12	8.56	54.20	1.74	1.38
ML15-4	9.4	9.2	179.5	12.6	210.7	7.0	42.6	4096.2	1.47	1.82	48.13	5.25	43.33
P6	9.1	0.5	3.0	<0.2	12.6	<0.5	<0.5	11.7	61.62	3.99	34.39	0.00	0.00
P7	22.7	0.6	4.8	<0.2	28.1	<0.5	<0.5	<0.4	71.50	2.51	25.99	0.00	0.00
P8	26.0	4.8	25.9	<0.2	56.7	<0.5	<0.5	4.5	34.02	7.95	58.03	0.00	0.00
P9	33.1	3.7	79.7	<0.2	116.5	<0.5	<0.5	3.2	19.02	2.70	78.28	0.00	0.00
P10	3.6	<0.2	<0.2	<0.2	3.6	<0.5	<0.5	140.1	100.00	0.00	0.00	0.00	0.00
P13	151.8	<0.2	1.2	<0.2	153.0	<0.5	<0.5	14.0	98.66	0.00	1.34	0.00	0.00
P14	14.9	0.6	5.5	<0.2	21.0	<0.5	<0.5	9.5	59.10	3.21	37.69	0.00	0.00
P15	35.6	4.7	24.3	<0.2	64.6	<0.5	<0.5	1.7	42.91	7.08	50.01	0.00	0.00
P16	19.0	4.8	108.4	<0.2	132.2	<0.5	<0.5	<0.4	9.04	2.89	88.08	0.00	0.00
P17	18.8	23.8	989.2	<0.2	1031.8	<0.5	<0.5	<0.4	1.08	1.72	97.19	0.00	0.00
ML9-2	25.9	0.5	3.5	<0.2	29.9	<0.5	<0.5	4.1	79.46	2.08	18.45	0.00	0.00
ML9-4	34.8	5.3	34.7	<0.2	74.8	<0.5	<0.5	13.6	34.54	6.59	58.87	0.00	0.00
ML9-5	20.8	5.6	98.6	<0.2	125.0	<0.5	3.9	38.4	9.52	3.27	77.35	0.00	9.87
ML9-6	3.9	6.5	1033.7	<0.2	1044.1	<0.5	75.7	534.3	0.18	0.37	80.46	0.00	18.99
R1	1.8	<0.2	<0.2	<0.2	1.8	<0.5	<0.5	0.9	100.00	0.00	0.00	0.00	0.00
R2	1.8	<0.2	<0.2	<0.2	1.8	<0.5	<0.5	0.6	100.00	0.00	0.00	0.00	0.00
R3	<0.2	<0.2	<0.2	<0.2	<0.2	<0.5	<0.5	7.0	0.00	0.00	0.00	0.00	0.00
R4	<0.2	<0.2	<0.2	<0.2	<0.2	<0.5	<0.5	180.5	0.00	0.00	0.00	0.00	0.00

References

1. **Hendrickson, E. R., Payne, J. A., Young, R. M., Starr, M. G., Perry, M. P., Fahnestock, S., Ellis, D. E., and Ebersole, R. C.** (2002) Molecular analysis of *Dehalococcoides* 16S ribosomal DNA from chloroethene-contaminated sites throughout North America and Europe, *Applied and Environmental Microbiology* 68 (2):485–495.
2. **Bradley, P. M.** (2003) History and ecology of chloroethene biodegradation: A review, *Bioremediation Journal* 7 (2):81–109.
3. **Abe, Y., Aravena, R., Zopfi, J., Parker, B., and Hunkeler, D.** (2009) Evaluating the fate of chlorinated ethenes in streambed sediments by combining stable isotope, geochemical and microbial methods, *Journal of Contaminant Hydrology* 107 (1-2):10–21.
4. **Aeppli, C., Hofstetter, T. B., Amaral, H. I. F., Kipfer, R., Schwarzenbach, R. P., and Berg, M.** (2010) Quantifying in situ transformation rates of chlorinated ethenes by combining compound-specific stable isotope analysis, groundwater dating, and carbon isotope mass balances, *Environmental Science & Technology* 44 (10):3705–3711.
5. **Wiedemeier, T. H., Swanson, M. A., Moutoux, D. E., Gordon, E. K., Wilson, J. T., Wilson, B. H., Kampbell, D. H., Haas, P. E., Miller, R. N., Hansen, J. E., and Chapelle, F. H.** (1998) Technical protocol for evaluating natural attenuation of chlorinated solvents in ground water, USEPA, Washington D. C.
6. **McMahon, P. B., and Chapelle, F. H.** (2008) Redox processes and water quality of selected principal aquifer systems, *Ground Water* 46 (2):259–271.
7. **Heimann, A., Jakobsen, R., and Blodau, C.** (2010) Energetic constraints on H_2-dependent terminal electron accepting processes in anoxic environments: a review of observations and model approaches, *Environmental Science & Technology* 44 (1):24–33.
8. **Vogel, T. M., Criddle, C. S., and McCarty, P. L.** (1987) ES Critical Reviews: Transformations of halogenated aliphatic compounds, *Environmental Science & Technology* 21 (8):722–736.
9. **Lee, P. K. H., Warnecke, F., Brodie, E. L., Macbeth, T. W., Conrad, M. E., Andersen, G. L., and Alvarez-Cohen, L.** (2011) Phylogenetic microarray analysis of a microbial community performing reductive dechlorination at a TCE-contaminated site, *Environmental Science & Technology* 46 (2):1044–1054.
10. **Maillard, J., Charnay, M.-P., Regeard, C., Rohrbach-Brandt, E., Rouzeau-Szynalski, K., Rossi, P., and Holliger, C.** (2011) Reductive dechlorination of tetrachloroethene by a stepwise catalysis of different organohalide respiring bacteria and reductive dehalogenases, *Biodegradation* 22 (5):949–960.
11. **Rouzeau-Szynalski, K., Maillard, J., and Holliger, C.** (2011) Frequent concomitant presence of *Desulfitobacterium* spp. and "*Dehalococcoides*" spp. in chloroethene-dechlorinating microbial communities, *Applied Microbiology and Biotechnology* 90 (1):361–368.
12. **Taş, N., van Eekert, M. H. A., de Vos, W. M., and Smidt, H.** (2010) The little bacteria that can - diversity, genomics and ecophysiology of "*Dehalococcoides*" spp. in contaminated environments, *Microbial Biotechnol*ogy 3 (4):389–402.
13. **Maymó-Gatell, X., Chien, Y., Gossett, J. M., and Zinder, S. H.** (1997) Isolation of a bacterium that reductively dechlorinates tetrachloroethene to ethene, *Science* 276 (5318):1568–1571.
14. **Kube, M., Beck, A., Zinder, S. H., Kuhl, H., Reinhardt, R., and Adrian, L.** (2005) Genome sequence of the chlorinated compound-respiring bacterium *Dehalococcoides* species strain CBDB1, *Nature Biotechnology* 23 (10):1269–1273.

15. **Seshadri, R., Adrian, L., Fouts, D. E., Eisen, J. A., Phillippy, A. M., Methe, B. A., Ward, N. L., Nelson, W. C., Deboy, R. T., Khouri, H. M., Kolonay, J. F., Dodson, R. J., Daugherty, S. C., Brinkac, L. M., Sullivan, S. A., Madupu, R., Nelson, K. E., Kang, K. H., Impraim, M., Tran, K., Robinson, J. M., Forberger, H. A., Fraser, C. M., Zinder, S. H., and Heidelberg, J. F.** (2005) Genome sequence of the PCE-dechlorinating bacterium *Dehalococcoides ethenogenes*, *Science* 307 (5706):105–108.

16. **Vieites, J. M., Guazzaroni, M.-E., Beloqui, A., Golyshin, P. N., and Ferrer, M.** (2009) Metagenomics approaches in systems microbiology, *FEMS Microbiology Reviews* 33 (1):236–255.

17. **ITRC**. (2008) In situ bioremediation of chlorinated ethene: DNAPL source zones, The Interstate Technology & Regulatory Council.

18. **Vroblesky, D. A., and Chapelle, F. H.** (1994) Temporal and spatial changes of terminal electron-accepting processes in a petroleum hydrocarbon-contaminated aquifer and the significance for contaminant biodegradation, *Water Resources Research* 30 (5):1561–1570.

19. **McGuire, J. T., Long, D. T., and Hyndman, D. W.** (2005) Analysis of recharge-induced geochemical change in a contaminated aquifer, *Ground Water* 43 (4):518–530.

20. **Eaton, A. D., Franson, M. A. H., Association, A. W. W., and Federation, W. E.** (2005) Standard methods for the examination of water & wastewater, American Public Health Association.

21. **Goto, K., Komatsu, T., and Furukawa, T.** (1962) Rapid colorimetric determination of manganese in waters containing iron : A modification of the formaldoxime method, *Analytica Chimica Acta* 27 :331–334.

22. **Cline, J. D.** (1969) Spectrophotometric determination of hydrogen sulfide in natural waters, *Limnology and Oceanography* 14 (3):454–458.

23. **Rossi, P., Gillet, F., Rohrbach, E., Diaby, N., and Holliger, C.** (2009) Statistical assessment of variability of terminal restriction fragment length polymorphism analysis applied to complex microbial communities, *Applied and Environmental Microbiology* 75 (22):7268–7270.

24. **R development Core Team** (2008), R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.

25. **Jari Oksanen, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Helene Wagner** (2008), Vegan: Community Ecology Package. R package version 1.16-22, *http://R-Forge.R-project.org/projects/vegan/*

26. **Legendre, P., and Gallagher, E.** (2001) Ecologically meaningful transformations for ordination of species data, *Oecologia* 129 (2):271–280.

27. **Legendre, P., and Legendre, L.** (1998) Numerical Ecology, Volume 24, Second Edition, Elsevier Science.

28. **Escofier, B., and Pages, J.** (1994) Multiple factor analysis (AFMULT package), *Computational Statistics & Data Analysis* 18 (1):121–140.

29. **Bécue-Bertaut, M., and Pagès, J.** (2008) Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data, *Computational Statistics & Data Analysis* 52 (6):3255–3268.

30. **Robert, P., and Escoufier, Y.** (1976) A unifying tool for linear multivariate statistical methods: The RV coefficient, *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 25 (3):257–265.

31. **Dufrêne, M., and Legendre, P.** (1997) Species assemblages and indicator species: The need for a flexible asymmetrical approach, *Ecological Monographs* 67 (3):345–366.

32. **Lane, D. J.** (1991) Nucleic acid techniques in bacterial systematics, John Wiley & Sons.

33. **Muyzer, G., de Waal, E. C., and Uitterlinden, A. G.** (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA., *Applied and Environmental Microbiol*ogy 59 (3):695–700.

34. **Behrens, S., Azizian, M. F., McMurdie, P. J., Sabalowsky, A., Dolan, M. E., Semprini, L., and Spormann, A. M.** (2008) Monitoring abundance and expression of "*Dehalococcoides*" species chloroethene-reductive dehalogenases in a tetrachloroethene-dechlorinating flow column, *Applied And Environmental Microbiology* 74 (18):5695–5703.

35. **Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J., and Knight, R.** (2010) QIIME allows analysis of high-throughput community sequencing data, *Nature Methods* 7 (5):335–336.

36. **Quince, C., Lanzen, A., Davenport, R. J., and Turnbaugh, P. J.** (2011) Removing noise from pyrosequenced amplicons, *BMC Bioinformatics* 28:12-38.

37. **Balzer, S., Malde, K., and Jonassen, I.** (2011) Systematic exploration of error sources in pyrosequencing flowgram data, *Bioinformatics* 27 (13):304–309.

38. **Li, H., and Durbin, R.** (2010) Fast and accurate long-read alignment with Burrows–Wheeler transform, *Bioinformatics* 26 (5):589–595.

39. **McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., Desantis, T. Z., Probst, A., Andersen, G. L., Knight, R., and Hugenholtz, P.** (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea, *The ISME Journal* 6 (3):610–618.

40. **Liu, W. T., Marsh, T. L., Cheng, H., and Forney, L. J.** (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA., *Applied and Environmental Microbiology* 63 (11):4516–4522.

41. **Kaplan, C. W., and Kitts, C. L.** (2003) Variation between observed and true Terminal Restriction Fragment length is dependent on true TRF length and purine content, *Journal of Microbiological Methods* 54 (1):121–125.

42. **Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., and Edwards, R. A.** (2008) The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes, *BMC Bioinformatics* 9 (1):386.

43. **Parks, D. H., and Beiko, R. G.** (2010) Identifying biologically relevant differences between metagenomic communities, *Bioinformatics* 26 (6):715–721.

44. **Bradley, P. M.** (2011) Reinterpreting the importance of oxygen-based biodegradation in chloroethene-contaminated groundwater, *Ground Water Monitoring & Remediation* 31 (4): 50–55.

45. **Müller, J. A., Rosner, B. M., von Abendroth, G., Meshulam-Simon, G., McCarty, P. L., and Spormann, A. M.** (2004) Molecular identification of the catabolic vinyl chloride reductase from *Dehalococcoides* sp. Strain VS and its environmental distribution, *Appliend and Environmental Microbiology* 70 (8):4880–4888.

46. **Krajmalnik-Brown, R., Hölscher, T., Thomson, I. N., Saunders, F. M., Ritalahti, K. M., and Löffler, F. E.** (2004) Genetic identification of a putative vinyl chloride reductase in *Dehalococcoides* sp. Strain BAV1, *Applied and Environmental Microbiology* 70 (10):6347–6351.

47. **Sung, Y., Ritalahti, K. M., Sanford, R. A., Urbance, J. W., Flynn, S. J., Tiedje, J. M., and Löffler, F. E.** (2003) Characterization of two tetrachloroethene-reducing, acetate-oxidizing anaerobic bacteria and their description as *Desulfuromonas michiganensis* sp. nov, *Applied and Environmental Microbiology* 69 (5):2964–2974.

48. **Villemur, R., Lanthier, M., Beaudet, R., and Lépine, F.** (2006) The *Desulfitobacterium* genus, *FEMS Microbiology Reviews* 30 (5):706–733.

49. **Krumholz, L. R.** (1997) *Desulfuromonas chloroethenica* sp. nov. uses tetrachloroethylene and trichloroethylene as electron acceptors, *International Journal of Systematic Bacteriol*ogy 47 (4):1262–1263.

50. **Chapelle, F. H., Bradley, P. M., Thomas, M. A., and McMahon, P. B.** (2009) Distinguishing iron-reducing from sulfate-reducing conditions, *Ground Water* 47 (2):300–305.

51. **Chakraborty, R., O'Connor, S. M., Chan, E., and Coates, J. D.** (2005) Anaerobic degradation of benzene, toluene, ethylbenzene, and xylene compounds by *Dechloromonas* strain RCB, *Applied and Environmental Microbiology* 71 (12):8649–8655.

52. **Gillis, M., and Ley, J. D. E.** (2006) The genera *Chromobacterium* and *Janthinobacterium*, In *The Prokaryotes* third edition:737–746.

53. **Bernardet, J.-F., Segers, P., Vancanneyt, M., Berthe, F., Kersters, K., and Vandamme, P.** (1996) Cutting a gordian knot: emended classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatis* nom. nov. (Basonym, *Cytophaga aquatilis* Strohl and Tait 1978), *International Journal of Systematic Bacteriol*ogy 46 (1):128–148.

54. **de Boer, W., Leveau, J. H. J., Kowalchuk, G. A., Klein Gunnewiek, P. J. A., Abeln, E. C. A., Figge, M. J., Sjollema, K., Janse, J. D., and van Veen, J. A.** (2004) *Collimonas fungivorans* gen. nov., sp. nov., a chitinolytic soil bacterium with the ability to grow on living fungal hyphae, *International Journal of Systematic and Evolutionary Microbiology* 54 (3):857–864.

55. **de Boer, W., Gunnewiek, P. J. A. K., Veenhuis, M., Bock, E., and Laanbroek, H. J.** (1991) Nitrification at low pH by aggregated chemolithotrophic bacteria, *Applied and Environmental Microbiol*ogy 57 (12):3600–3604.

56. **Jiang, QQ, and Bakken, LR** (1999) Comparison of *Nitrosospira* strains isolated from terrestrial environments, *FEMS Microbiology Ecology* 30(2):171–186.

57. **Almeida, J. S., Reis, M. A., and Carrondo, M. J.** (1995) Competition between nitrate and nitrite reduction in denitrification by *Pseudomonas fluorescens*, *Biotechnology and Bioengineering* 46 (5):476–484.

58. **Carter, J. P., Richardson, D. J., and Spiro, S.** (1995) Isolation and characterisation of a strain of *Pseudomonas putida* that can express a periplasmic nitrate reductase, *Archives of Microbiology* 163 (3):159–166.

59. **Daum, M., Zimmer, W., Papen, H., Kloos, K., Nawrath, K., and Bothe, H.** (1998) Physiological and molecular biological characterization of ammonia oxidation of the heterotrophic nitrifier *Pseudomonas putida*, *Current Microbiology* 37 (4):281–288.

60. **Willems, A., Goor, M., Thielemans, S., Gillis, M., Kersters, K., and De Ley, J.** (1992) Transfer of several phytopathogenic *Pseudomonas* species to *Acidovorax* as *Acidovorax avenae* subsp. *avenae* subsp. nov., comb. nov., *Acidovorax avenae* subsp. *citrulli*, *Acidovorax avenae* subsp. *cattleyae*, and *Acidovorax konjaci*, *International Journal of Systematic Bacteriology* 42 (1):107–119.

61. **Monferrán, M. V., Echenique, J. R., and Wunderlin, D. A.** (2005) Degradation of chlorobenzenes by a strain of *Acidovorax avenae* isolated from a polluted aquifer, *Chemosphere* 61 (1):98–106.

63. **Finneran, K. T., Johnsen, C. V., and Lovley, D. R.** (2003) *Rhodoferax ferrireducens* sp. nov., a psychrotolerant, facultatively anaerobic bacterium that oxidizes acetate with the reduction of Fe(III), *International Journal of Systematic and Evolutionary Microbiology* 53 (3):669–673.

63. **Zhuang, K., Izallalen, M., Mouser, P., Richter, H., Risso, C., Mahadevan, R., and Lovley, D. R.** (2011) Genome-scale dynamic modeling of the competition between *Rhodoferax* and *Geobacter* in anoxic subsurface environments, *The ISME Journal* 5, (2):305–316.

65. **Hiraishi, A., Hoshino, Y., and Satoh, T.** (1991) *Rhodoferax fermentans* gen. nov., sp. nov., a phototrophic purple nonnsulfur bacterium previously referred to as teh "*Rhodocyclus gelatinosus*-like" group, *Archives of Microbiology 155*, 330–336.

65. **Lin, W. C., Coppi, M. V., and Lovley, D. R.** (2004) *Geobacter sulfurreducens* can grow with oxygen as a terminal electron acceptor, *Applied and Environmental Microbiology* 70 (4):2525–2528.

66. **Bradley, P. M., and Chapelle, F. H.** (2011) Microbial mineralization of dichloroethene and vinyl chloride under hypoxic conditions, *Ground Water Monitoring & Remediation* 31 (4):39–49.

67. **Amos, B. K., Ritalahti, K. M., Cruz-Garcia, C., Padilla-Crespo, E., and Löffler, F. E.** (2008) Oxygen effect on *Dehalococcoides* viability and biomarker quantification, *Environmental Science & Technology* 42 (15):5718–5726.

68. **He, J., Ritalahti, K. M., Yang, K.-L., Koenigsberg, S. S., and Löffler, F. E.** (2003) Detoxification of vinyl chloride to ethene coupled to growth of an anaerobic bacterium, *Nature* 424 (6944):62–65.

69. **Bradley, P. M., and Chapelle, F. H.** (1998) Microbial mineralization of VC and DCE under different terminal electron accepting conditions, *Anaerobe* 4 (2), 81–87.

70. **Smits, T. H. M., Assal, A., Hunkeler, D., and Holliger, C.** (2011) Anaerobic degradation of vinyl chloride in aquifer microcosms, *Journal of Environmental Quality* 40 (3):915–922.

71. **Tiehm, A., and Schmidt, K. R.** (2011) Sequential anaerobic/aerobic biodegradation of chloroethenes--aspects of field application, *Current Opinion in Biotechnology* 22 (3):415–421.

72. **Tiehm, A., Schmidt, K. R., Pfeifer, B., Heidinger, M., and Ertl, S.** (2008) Growth kinetics and stable carbon isotope fractionation during aerobic degradation of cis-1,2-dichloroethene and vinyl chloride, *Water Research* 42 (10-11):2431–2438.

73. **Schmidt, K. R., Augenstein, T., Heidinger, M., Ertl, S., and Tiehm, A.** (2010) Aerobic biodegradation of cis-1,2-dichloroethene as sole carbon source: Stable carbon isotope fractionation and growth characteristics, *Chemosphere* 78 (5):527–532.

Concluding remarks and outlook

6 Concluding remarks and outlook

6.1 Summary and conclusions

The main objective of this thesis was to elucidate the reasons for the accumulation of intermediate products induced by the incomplete biodegradation of chlorinated ethenes (CEs) in contaminated aquifers, such as dichloroethene (DCE) and vinyl chloride (VC). The investigations carried out in this thesis focused on natural processes taking place *in situ* using an ecological approach so as to address the complexity of the processes and their interactions, prior to deducing the possible reasons for the accumulation of lower CEs. Two contaminated sites undergoing monitored natural attenuation (MNA) were investigated for this purpose. Both were showing incomplete reductive biodegradation of CEs.

6.1.1 Sampling approach

Although groundwater sampling does not enable drawing a comprehensive picture of aquifer bacterial communities, it is convenient for low-cost and easy monitoring of the evolution of the contaminant biodegradation as well as for describing overall aquifer physico-chemical factors and microbial communities. Since this work was based on groundwater samples, it was important to gain knowledge on the impact of pumping parameters on apparent bacterial community structures (BCS). This study showed that selected parameters related to the tubing, such as lift, flow regime within the tubing, and tubing material, had no detectable impact on BCS, as was expected from previous experiments carried out with chemical compounds. Significant impact of pumping flow rate on BCS was detected when groundwater was extracted from low permeability zones. Differential impact was related to the lithological composition of the aquifer matrix, underlining the importance of low-flow and standardized sampling.

6.1.2 Pyrosequencing-based terminal-restriction fragments identification (PyroTRF-ID)

The core analysis of groundwater BCS has been obtained with terminal-restriction fragment length polymorphism (T-RFLP). This technique enabled obtaining rapid fingerprints of a relatively high number of samples, which were conveniently exploited in further statistical investigations. However, identities of the bacteria significantly expressed in the T-RFLP profiles were not known and it would have required huge efforts and time to unambiguously identify terminal-restriction fragments (T-RFs) of interest with classical cloning and sequencing approach. The bioinformatics approach described in CHAPTER 3 enabled assigning a phylogenetic affiliation to several T-RFs of interest via a combination of pyrosequencing and T-RFLP data, and was a convenient tool throughout this work. More generally, this tool represented a good compromise between laboratory efforts required for the description of bacterial communities (in time and space) and the

financial and infrastructural costs related to multiple datasets reprocessing and may reveal essential for future investigations of environmental microbial biocenoses.

6.1.3 Zuchwil aquifer: community ecology

VC, and to a lesser extent cDCE, was known from previous work (*1*) to accumulate in the aquifer investigated in CHAPTER 4. Thorough investigation of BCS and physico-chemical environmental variables, as well as their correlations, indicated that the two main terminal electron-accepting processes (TEAPs) shaping BCS were Fe(III)- and VC reductions. Statistical findings suggested that these two TEAPs were in competition for electron donor(s), indicating the presence of a strong heterogeneous zonation of the process. Furthermore, sulfate reduction might have favored VC reduction by enhancing Fe(III) reduction locally. Identification of a T-RF strongly correlated with VC revealed that a bacterium closely related to a previously reported PCE-degrading OHR guild member (*2*) may be involved in the reduction of lower CEs. In general, this study showed the need for a clear understanding of the respiratory mechanisms present within the aquifer, which includes also determining the possible occurrence of competitive processes. In this sense, the reductive dechlorination of lower CEs was itself considered as a TEAP, which acted as a supplementary source of electron acceptors.

6.1.4 Lyss aquifer: characterization of the habitat

Discrepancies were observed in organohalide respiration (OHR) on the second contaminated site in zones displaying *a priori* similar hydrogeological conditions. Profiling of the BCS provided evidence of the presence of radically different bacterial communities in these zones, which were strongly influenced by the hydrological functioning of the aquifer. A selection of microbiological, chemical, hydrogeological, and statistical indications attributed the reason for local cDCE accumulation to the occasional decrease of reducing conditions. The observed OHR discrepancy, and the subsequent incomplete degradation of the CEs, was related more precisely to the occasional upward recharge of the anoxic lithological layer with oxidized groundwater. As a consequence, local discrepancies in the porosity of the aquifer, coupled with differences in local organic carbon content, impacted differentially the OHR process. These conclusions pointed out the need of in-depth description and understanding of the hydrogeological features of contaminated sites.

6.2 The contaminated aquifer ecosystem – ecological considerations

6.2.1 The aquifer, a dynamic habitat hosting OHRB

In aquifers, as it is the case for all ecosystems, microorganisms live and interact according to strict ecological rules. These rules are induced implicitly by underlying energetic mechanisms. In this sense, the chemical values measured from the surface, such as pH, redox potential, and electron donors and acceptors, form a context conducive to the development of a balanced biocenose. These values in turn are strongly dependent

on the geological and hydrodynamic context of these habitats, such as the connectivity with the surface and the recharge of the aquifer. All the above mentioned environmental variables determine the framework of the development and the activities of the biocenose. However, and because of their metabolic activities, microorganisms are likely to impact and modify permanently environmental variables in these environments. The groundwater habitat and the related biocenose must be considered as a complex dynamic assemblage of physical and chemical variables influencing and in turn impacted by biological activity. Furthermore, and due to their nature, these habitats can be considered as geographic islands which are isolated spatially, showing relatively slow and globally unidirectional water fluxes, with no or reduced interconnectivity.

From another point of view, it is remarkable that a contaminant such as PCE or TCE is systematically degraded when environmental conditions are favorable. It is all the more singular that this degradation occurs in often pristine habitats that have previously never been exposed to this type of contaminant. This remark is naturally valid also for any other type of contaminants and denotes the profound lack of knowledge that characterize the catabolic potential present in these habitats.

6.2.2 Chlorinated ethenes are an integral part of the habitat

Degradation pathway of CEs is related to their degree of chlorination (*3*, *4*), i.e. oxidation becomes more favorable as the number of chlorine substituents decreases. Thus, PCE and TCE are more easily reduced than DCE and VC. In groundwater, as it is the case in all ecosystems, thermodynamically more favorable and kinetically more efficient TEAPs preferentially occur. The spatial and temporal structuring of microbial processes therefore takes place according to the so-called thermodynamic ladder (*5*). Showing no exception to the rule, OHR of CEs strictly follows this ladder. Apart from the source zone, where dissolved CEs reach toxic concentrations (*6*–*9*), CEs can be used as terminal electron acceptors and their reduction thus constitutes a TEAP *per se*. CEs are often considered from an anthropocentric point of view as an external disturbance to the aquifer system. However, from an ecological point of view, they merely represent an additional terminal electron acceptor that will or will not be used by the indigenous microorganisms depending on the energy they could retrieve and on the competitive advantage that these compounds would provide over other TEAPs.

The differential degradation capabilities of higher and lower CEs gives rise to their characteristic behavior in aquifers. Indeed, the two CEs most often released in the environment, PCE and TCE, are highly oxidized molecules recalcitrant to oxidation. On the contrary, they are prone to be reduced in anoxic environments. Their degradation products are consequently released in anoxic, reduced environments where oxidation is not an obvious task. On the other hand, reductive dechlorination of DCE and VC requires highly reducing conditions which are hardly reached as long as relatively oxidized terminal electron acceptors are present and feed competing TEAPs. Consequently, sequential reductive dechlorination often stalls after TCE reduction, generating lower CEs accumulation. Despite this simple explanation of stalling reasons, the phenomenon is difficult to predict under the complex conditions usually prevailing in aquifers.

6.2.3 Towards an ecological engineering of contaminated sites

The concept underlying the approach adopted in this thesis was that aquifer environments host microbial systems that are "intelligent" in the sense that they can (i) feel the temporal and local variations of the habitat, (ii) react to these variations, and (iii) optimize their functioning in adopting the most efficient energetic behavior. In this sense, assessing incomplete OHR of CEs in the light of the presence/absence of *Dehalococcoides* sp. is a reductionist view that may lead to erroneous conclusions. In this thesis, the presence of *Dehalococcoides* sp. was not considered as a prerequisite for complete OHR of CEs, since this hypothesis would have meant keeping aside microorganisms showing a potentially, yet unknown, ability to reductively dechlorinate lower CEs. For instance, CHAPTER 4 provided statistical evidence that bacteria phylogenetically different from *Dehalococcoides* sp. are able to degrade higher and lower CEs. Furthermore, it was hypothesized that only favorable environmental conditions lead lower CE-reducing microorganisms such as *Dehalococcoides* sp. to proliferate and express their full OHR potential. However, since the aquifer ecosystem is highly complex, deducing whether the conditions are favorable or not by considering only redox conditions is relatively uncertain. In the two aquifers reported in this thesis (CHAPTER 4 and CHAPTER 5), overall redox conditions were apparently favorable to the complete reductive dechlorination *a priori*, though stalling of OHR was observed. In CHAPTER 5, the geological structure of the aquifer itself was suspected to be responsible for seasonal changes of the redox conditions in the lithological layer where OHR was stalling at DCE. Based on these considerations, insufficient ecological background may preclude engineered remediation approaches such as bioaugmentation, due to inaccurate predictions of the behavior of the microorganisms injected into the aquifer ecosystem (*10*). It has been stated elsewhere that to understand and predict microbial ecosystem functioning, reductionist approaches should be avoided (*11*). The ecological approach used in this thesis allowed pointing out potential reasons for lower CEs accumulation that would perhaps have been missed or hardly detected otherwise. Multivariate statistics were a powerful tool to build an integrated picture of the bacterial community and its relationships with environmental variables. Based on this global picture, hypotheses were formulated to explain the accumulation of lower CEs. The hypotheses were confronted to data obtained from the field for verification.

The proposed methodology converged with the integrative approach followed by Imfeld (2009) to evidence and follow-up *in situ* biodegradation processes in wetland systems (*12*). Apart from the fact that we focused on the specific issue of lower CEs accumulation, the methodology proposed here was similar in terms of integration of techniques for an in-depth description of the system. However, in his integrative approach, Imfeld focused on the biodegradation issue only in a process-oriented way. Here, it was rather proposed to first conceptualize the global ecosystem functioning, considering OHR as a TEAP among others. In other words, we proposed to assess biodegradation aspects starting from an ecological way of thinking the aquifer. This implies that hydrogeological aspects of the aquifer (groundwater fluxes, geological structure, lithology) are described as well.

Dynamic aspects of the aquifer ecosystem have been only briefly approached in this thesis. They may however represent a major feature in the fate of OHR, as suggested in CHAPTER 5. From a practical point of view, diagnosis of OHR potential on a contaminated site should take this aspect into account. Indeed, general conditions may be favorable to OHR at a certain time point of the investigation, but since the aquifer is a dynamic habitat, conditions may rapidly change and become unfavorable. To date, microorganisms known to reductively dechlorinate lower CEs, i.e. *Dehalococcoides* sp., were shown to be very sensitive to variations of the surrounding environmental conditions. For instance, it has been demonstrated in laboratory experiments that a slight increase of O_2 concentrations repressed the reductive dehalogenases. These enzymes did not recover their activity when the conditions became favorable again (*13*). Such changes of the redox conditions occur for example as a result of aquifer recharge (*14*, *15*). Thus, temporal follow-up is crucial to determine whether MNA would be an adapted engineering strategy on a particular contaminated site.

6.3 Outlook: an ecological approach of contaminated sites accumulating lower chlorinated ethenes

6.3.1 General recommendations to investigate contaminated sites with an ecological approach

Due to the inherent complexity of aquifer ecosystems, each contaminated site displays specific characteristics and will respond in a specific way to the introduction of the contaminant. Specific geological structure, the connectivity with the surface, and the grain size distribution will impact groundwater flows within the aquifer. Groundwater flows in turn will impact the fluxes of matter, such as terminal electron acceptors, carbon sources and electron donors. Geochemical and mineralogical compositions can prelude potential TEAPs. For instance, Fe(III) reduction (CHAPTER 4) requires iron in the solid phase of the sediment. Microorganisms coping with OHR in different aquifers will have to find their niche in the complex framework that will take place in these complex ecosystems. Based on these considerations, each contaminated site should be handled as a specific case, since conclusions obtained for one site may be irrelevant for another.

To understand the particularities of each contaminated site, multidisciplinary approaches are indispensable for describing the different components of the aquifer. Hydrogeological expertise is needed to characterize the structure of the aquifer and its hydrological functioning, as well as characterizing the contaminant distribution and the evolution of the plume. Hydrochemistry provides essential knowledge on the chemical composition of groundwater, and thus on the ongoing processes. Microbiology provides knowledge on the bacterial communities involved in the different hydrogeochemical processes. The synthesis of the collected datasets must be reordered and analyzed in an integrative way, forming a comprehensive picture of the aquifer ecosystem. Multivariate statistical tools provide numerous techniques for this purpose. They must

however be handled with care, and the choice of the techniques to be used for specific questions is crucial. For instance, it was decided along the present work to assess the relationships between BCS and the environmental variables using multifactorial analysis (MFA), because the hypotheses underlining this technique were fitting our biological understanding of the aquifer ecosystem. It was preferred to another powerful technique, namely redundancy analysis, because the latter presupposes that the environmental variables unilaterally influence the bacterial communities, which was not totally true in the aquifer ecosystem, where environmental variables are themselves greatly impacted by the bacterial communities.

An untouched aspect of this work that still need to be exploited is about the putative role displayed by other organisms in the elimination of the CEs contaminants. In this thesis, only the bacterial part of the aquifer biocenosis was considered. To provide a more comprehensive picture, *Archaea* could be included in the investigation as it has been shown already that these organisms may be involved in different aspects of the elimination of the CEs (*16*, *17*). The importance of other organisms such as Protozoans must also be investigated further, whereas the role of Invertebrates can be considered as negligible in anaerobic contaminated aquifers (*18*).

Functional aspects of the bacterial community provide precious information, such as the activity of enzymatic systems. Quantitative PCR targeting specific catabolic genes or bacterial protagonists is an adapted tool for these more specific aspects. The study of RNA provides further information on the activity of targeted catabolic genes, such as reductive dehalogenases. Widening up our view to yet unknown reductive dehalogenase genes, by using degenerate primers, could bring additional knowledge on the catabolic potential of contaminated aquifers. For instance a whole section of analysis carried out on reductive dehalogenases (*rdh*A) and on the design of degenerate primers targeting up to 90% of all sequences found in references databases (more than 350 in January 2012) is not presented in the current thesis and still need to be exploited.

6.3.2 Detailed description of the proposed methodology

As mentioned above, the goal of this study was to assess the reasons for lower CEs accumulation based on the understanding of the aquifer ecosystem. By definition, an ecosystem is composed of both abiotic (the habitat) and biotic (the organisms, and more specifically the microorganisms) components. The first step of the proposed methodology (Figure 1) is therefore to describe the geological and mineralogical components of the habitat (I). As mentioned above, hydrochemical features have to be characterized as well as detailed mapping of the contaminant distribution. For assessing the temporal evolution of the site, this characterization has to be carried out at regular intervals.

The second step of the methodology is the description of the microbial biocenosis that lives in and interacts with the habitat (II). Fingerprinting provides useful multiple pictures of the bacterial communities in the different sampling locations and metagenomics provides the phylogenetic diversity. Raw data arising from the descriptions of the habitat and the microbial biocenosis should be analyzed in an integrative way (III).

Knowledge from scientific fields, as well as dedicated analytical tools, is required at this stage. Multivariate statistics derived from numerical ecology tools play a central role, enabling linking the different datasets. Interrelationships between bacterial communities and environmental variables can be evaluated using unambiguous statistical techniques. In this thesis, MFA was adopted for the reasons described above. Another interesting feature of MFA is the possible calculation of correlation coefficients between single environmental variables and the entire bacterial community. This indication is of great use in the interpretation of the relationships between environmental variables and BCS. Expertise in hydrogeology, hydro-and geochemistry is required for interpreting data related to the habitat. Other tools such as sequences databases for the analysis of metagenomic data, as well as bioinformatics techniques and competence are needed as well. The software developed in CHAPTER 3 represents an advantageous tool at this stage. Indeed, it combines T-RFLP profiling, which is particularly well adapted for the follow-up in space and time of BCS at relatively low costs, with pyrosequencing, which provides more depth into the data, by bringing phylogenetic information from which functional aspects can be hypothesized.

Interpretation of the output of this integrative data analysis in light of microbiological and hydrogeochemical competences leads to the building-up of a conceptual model of the dynamic aquifer ecosystem (IV). In this model, OHR is placed in the framework of interactions between TEAPs. Deciphering whether MNA is a realistic remediation strategy (V) relies on the observation that OHR leads to the formation of ethene, and that no accumulation of intermediates occurs. If OHR results in lower CEs accumulation, hypotheses are formulated to explain the reasons of these phenomena (VI) based on up-to-date scientific knowledge of energetic mechanisms, such as the ecological succession of TEAPs, thermodynamic and kinetic aspects, as well as minimum hydrogen threshold requirements. Experience and intuition of the investigator may be helpful at this stage. Hypotheses for the reasons for lower CEs accumulation are then tested (VII) by confronting them to available data and statistical findings. At this stage, the most relevant variables pointed out by RV correlation coefficients with the bacterial community designate the major TEAPs within the contaminated site. This, together with the hydrodynamic functioning and genetic data, such as the amount of functional genes and the identity of selected T-RFs correlated with the TEAP-indicative environmental variables, allow checking the validity of the hypothesis. If necessary, additional information can be gathered from further field measurements for alleviating possible doubts about the behavior of incongruent variables. If the hypotheses fit available data and congruent statistical indications, a scenario explaining the accumulation is drawn (VIII).

At this final stage, dedicated strategies for the bioremediation of the contaminated site can be proposed and tested *in situ* (IX). Alternatively, small laboratory (*ex situ*) experiments could be set up, reducing the efforts and allowing for the testing of multiple remediation strategies simultaneously. This strategy is however reductionist and on-site experiments are preferred. Validation of the scenarios and designing of remediation strategies was left aside in this study for reasons of time and scope, and could be an interesting topic for further research.

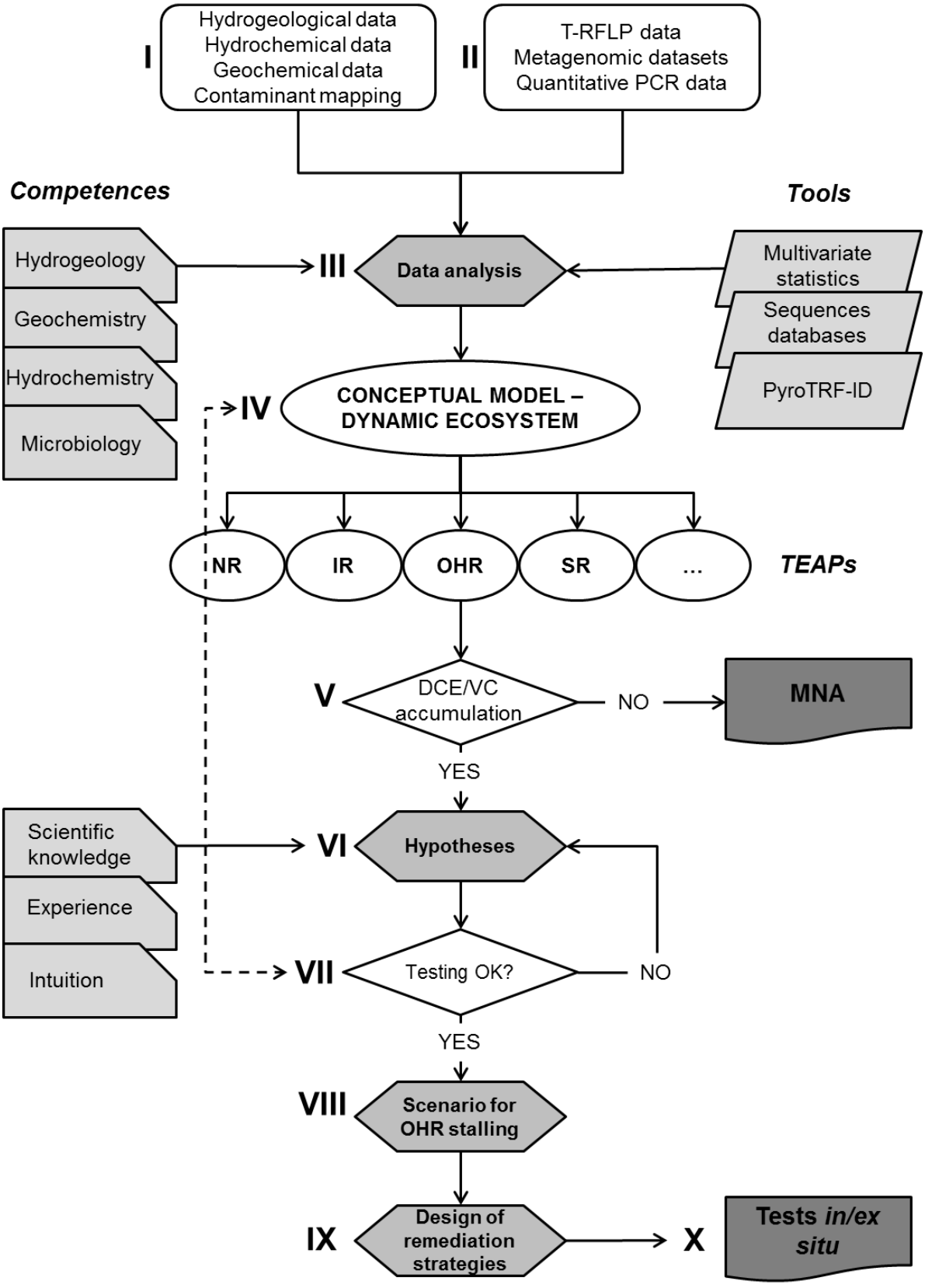


Figure 1. Flowchart describing the proposed ecological methodology of investigation of the reasons for lower CEs accumulation *in situ*. The flowchart is divided into 9 parts (Roman numbers) which are described in the text. The list of represented terminal electron-accepting processes (TEAPs) is not exhaustive and includes nitrate reduction (NR), iron reduction (IR), organohalide reduction (OHR), and sulfate reduction (SR).

References

1. **Aeppli, C., Hofstetter, T. B., Amaral, H. I. F., Kipfer, R., Schwarzenbach, R. P., and Berg, M.** (2010) Quantifying in situ transformation rates of chlorinated ethenes by combining compound-specific stable isotope analysis, groundwater dating, and carbon isotope mass balances, *Environmental Science & Technology* 44 (10):3705–3711.
2. **Kittelmann, S., and Friedrich, M. W.** (2008) Identification of novel perchloroethene-respiring microorganisms in anoxic river sediment by RNA-based stable isotope probing, *Environmental Microbiology* 10 (1):31–46.
3. **Vogel, T. M., Criddle, C. S., and McCarty, P. L.** (1987) ES Critical Reviews: Transformations of halogenated aliphatic compounds, *Environmental Science & Technology* 21 (8):722–736.
4. **Bossert, I. D., Häggblom, M. M., and Young, L.** (2004) Microbial ecology of dehalogenation, *Dehalogenation*: 33–52.
5. **Bethke, C. M., Sanford, R. A., Kirk, M. F., Jin, Q., and Flynn, T. M.** (2011) The thermodynamic ladder in geomicrobiology, *American Journal of Science* 311:183–210.
6. **Yu, S., and Semprini, L.** (2004) Kinetics and modeling of reductive dechlorination at high PCE and TCE concentrations, *Biotechnology and Bioengineering* 88 (4):451–464.
7. **Amos, B. K., Christ, J. A., Abriola, L. M., Pennell, K. D., and Loffler, F. E.** (2007) Experimental evaluation and mathematical modeling of microbially enhanced tetrachloroethene (PCE) dissolution, *Environmental Science & Technol*ogy 41 (3):963–70.
8. **Yu, S., and Semprini, L.** (2009) Enhanced reductive dechlorination of PCE DNAPL with TBOS as a slow-release electron donor, *Journal Of Hazardous Materials* 167 (1-3):97–104.
9. **Haest, P. J., Springael, D., and Smolders, E.** (2010) Dechlorination kinetics of TCE at toxic TCE concentrations: Assessment of different models, *Water Research* 44 (1):331–339.
10. **El Fantroussi, S., and Agathos, S. N.** (2005) Is bioaugmentation a feasible strategy for pollutant removal and site remediation? *Current Opinion in Microbiology* 8 (3):268–275.
11. **Vieites, J. M., Guazzaroni, M.-E., Beloqui, A., Golyshin, P. N., and Ferrer, M.** (2009) Metagenomics approaches in systems microbiology, *FEMS Microbiology Reviews* 33 (1):236–255.
12. **Imfeld, G.** (2009) Assessment of in situ biodegradation solvents in aquifers and constructed wetlands using an integrative approach. PhD thesis, Faculty of Science, University of Neuchâtel, Switzerland, 178 pp.
13. **Amos, B. K., Ritalahti, K. M., Cruz-Garcia, C., Padilla-Crespo, E., and Löffler, F. E.** (2008) Oxygen effect on *Dehalococcoides* viability and biomarker quantification, *Environmental Science & Technology* 42 (15):5718–5726.
13. **Vroblesky, D. A., and Chapelle, F. H.** (1994) Temporal and spatial changes of terminal electron-accepting processes in a petroleum hydrocarbon-contaminated aquifer and the significance for contaminant biodegradation, *Water Resources Research* 30 (5):1561–1570.
14. **McGuire, J. T., Long, D. T., and Hyndman, D. W.** (2005) Analysis of recharge-induced geochemical change in a contaminated aquifer, *Ground Water* 43 (4):518–530.
15. **Fathepure, B. Z., and Boyd, S. A.** (1988) Reductive dechlorination of perchloroethylene and the role of methanogens, *FEMS Microbiology Letters* 49 (2):149–156.
16. **Jablonski, P. E., and Ferry, J. G.** (1992) Reductive dechlorination of trichloroethylene by the CO-reduced CO dehydrogenase enzyme complex from Methanosarcina thermophila, *FEMS Microbiology Letters* 96 (1):55–60.
17. **Foulquier, A., Malard, F., Mermillod-Blondin, F., Montuelle, B., Doledec, S., Volat, B., and Gibert, J.** (2011) Surface water linkages regulate trophic interactions in a groundwater food web, *Ecosystems* 14 (8):1339–1353.

Table of content

1. General Introduction 7
1.1. General aspects of groundwater pollution with chlorinated ethenes 8
1.1.1. General considerations 8
1.1.2. Naturally occurring organohalides 8
1.1.3. Industrial use and abuse of chlorinated ethenes 9
1.1.4. Dense non-aqueous phase liquids (DNAPLs) and plume of dissolved contaminants in aquifers 11
1.2. Organohalide respiration of chloroethenes 11
1.2.1. CEs can be degraded biologically 11
1.2.2. Organohalide respiration (OHR) and organohalide-respiring bacteria (OHRB) 12
1.2.3. *In situ* interactions of the OHR guild with other microbial guilds 15
1.2.4. Oxidation of chlorinated ethenes 17
1.3. Existing remediation technologies and related problems 18
1.3.1. General aspects of groundwater remediation 18
1.3.2. Monitored natural attenuation 19
1.3.3. The aquifer, a complex ecosystem 20
1.3.4. Stalling OHR at cDCE or VC 20
1.4. Objectives and approach of the thesis 22
2. Impact assessment of pumping parameters on groundwater bacterial communities 31
2.1. Introduction 33
2.2. Materials & Methods 34
2.2.1. Laboratory experiment 34
2.2.2. Field experiment 36
2.2.3. Molecular analyses of microbial communities 37
2.2.4. Multivariate statistical analyses 38
2.3. Results 38
2.3.1. Controlled laboratory experiment with stable BCS of a homogenous water body 38
2.3.2. Field experiment 39
2.4. Discussion 41
2.5. Conclusion 42
3. PyroTRF-ID: a novel bioinformatics approach for the identification of terminal-restriction fragments using microbiome pyrosequencing data 45
3.1. Introduction 47
3.2. Materials & Methods 49
3.2.1. Samples 49
3.2.2. DNA extraction 49
3.2.3. Experimental T-RFLP 49
3.2.4. Cloning and sequencing 51
3.2.5. Pyrosequencing 51
3.2.6. Description of PyroTRF-ID 52
3.3. Results 56
3.3.1. PyroTRF-ID software and performance 56
3.3.2. Pyrosequencing quality assurance and read length limitation 57
3.3.3. Denoising 57
3.3.4. Mapping 59

3.3.5. Digital T-RFLP 59
3.3.6. Mirror plots and cross-correlations between digital and experimental T-RFLP profiles 64
3.3.7. Impact of sequence processing steps, pyrosequencing methods, and sample types 64
3.3.8. Peak annotation and phylogenetic affiliation of T-RFs 67
3.3.9. Screening of restriction enzymes 69
3.4. Discussion 71
3.4.1. Setup of the standard PyroTRF-ID procedure 71
3.4.2. Aspects of the PyroTRF-ID standard procedure 72
3.5. Conclusion 74
4. Accumulation of lower chlorinated ethenes in a contaminated aquifer due to competition between iron- and organohalide-respiring bacterial guilds 89
4.1. Introduction 91
4.2. Materials and Methods 92
4.2.1. Test site 92
4.2.2. Groundwater sampling 92
4.2.3. Chemical analyzes 94
4.2.4. DNA extraction 94
4.2.5. Specific detections of OHRB and reductive dehalogenase genes 94
4.2.6. Terminal-restriction fragment length polymorphism (T-RFLP) 96
4.2.7. Multivariate statistical analyses based on the T-RFLP profiles 96
4.2.8. Identification of T-RFs 97
4.3. Results 98
4.3.1. Environmental variables 98
4.3.2. Pairwise correlations between environmental variables 99
4.3.3. Bacterial communities 100
4.3.4. Relationships between environmental variables and bacterial community structures 102
4.3.5. Correlations between T-RFs and environmental variables 103
4.3.6. Identification of T-RFs 106
4.4. Discussion 109
4.4.1. Incomplete CE dechlorination cannot be explained by the lack of electron donors and the absence of biochemical potential 109
4.4.2. TEAPs shape the whole bacterial community present 110
4.4.3. CEs were dechlorinated by several *Chloroflexi* 111
4.4.4. Iron reduction is another main TEAP in the Zuchwil aquifer 112
4.4.5. Competition between VC reduction and iron reduction in the Zuchwil aquifer 113
4.4.6. Sulfate reduction potentially helps VC to be dechlorinated 114
4.5. Conclusion 115
5. Incomplete reductive dechlorination of tetrachloroethene as a result of hydrogeological characteristics of the Lyss aquifer 129
5.1. Introduction 131
5.2. Materials and Methods 132
5.2.1. Test site 132
5.2.2. Groundwater sampling 133
5.2.3. Analytical procedures 136
5.2.4. DNA extraction 136

5.2.5. Terminal-restriction fragment length polymorphism (T-RFLP) 137
5.2.6. Multivariate statistical analyses based on the T-RFLP profiles 137
5.2.7. Quantitative PCR (qPCR) 138
5.2.8. Pyrosequencing 139
5.2.9. Identification of T-RFs 139
5.2.10. Comparison of bacterial communities from Zones I and II 139
5.3. Results 140
5.3.1. Hydrochemical conditions in the Lyss aquifer 140
5.3.2. Correlations between bacterial community structures and environmental variables 141
5.3.3. Analysis of the bacterial communities of the upper lithological layer 144
5.3.4. Comparison of bacterial communities from Zones I and II 149
5.4. Discussion 151
5.4.1. OHR was only detected in the vicinity of the source zone 151
5.4.2. Causes of the absence of OHR in Zone II 152
5.4.3. Bacterial processes in Zones I and II based on bacterial community structure data 155
5.4.4. Conclusion: is MNA a solution for this site? 156
6. Concluding remarks and outlook 167
6.1. Summary and conclusions 168
6.1.1. Sampling approach 168
6.1.2. Pyrosequencing-based terminal-restriction fragments identification (PyroTRF-ID) 168
6.1.3. Zuchwil aquifer: community ecology 169
6.1.4. Lyss aquifer: characterization of the habitat 169
6.2. The contaminated aquifer ecosystem – ecological considerations 169
6.2.1. The aquifer, a dynamic habitat hosting OHRB 169
6.2.2. Chlorinated ethenes are an integral part of the habitat 170
6.2.3. Towards an ecological engineering of contaminated sites 171
6.3. Outlook: an ecological approach of contaminated sites accumulating lower chlorinated ethenes 172
6.3.1. General recommendations to investigate contaminated sites with an ecological approach 172
6.3.2. Detailed description of the proposed methodology 173

Remerciements

Je ne saurais conclure ce travail sans témoigner ma gratitude aux nombreuses personnes qui m'ont accompagné et soutenu durant ces quelques années de recherche.

Je n'ai pas de mots assez forts pour remercier mon directeur et mon co-directeur de thèse Christof Holliger et Pierre Rossi. Christof m'a offert un cadre de travail idéal, a toujours été présent et disponible pour partager son savoir et son esprit critique, et m'a accordé sa confiance et une grande liberté. Pierre quant à lui m'a apporté un soutien quotidien tout au long de ce projet, ainsi que des moments de détente, de discussion, d'échange et d'amitié. Dans les moments difficiles, il a toujours su poser les bonnes questions et trouver les mots justes. La complémentarité de leurs styles a sans aucun doute contribué à la forme qu'aura prise ma thèse, et la qualité de leur encadrement aura contribué à faire de mon travail un grand plaisir. Un grand merci à tous les deux !

Je remercie le président de mon jury de thèse, Prof. Fernando Porté-Agel, ainsi que les rapporteurs, Prof. Hauke Smidt, Prof. Timothy M. Vogel, et Prof. Josef Zeyer.

Diverses collaborations ont permis à cette thèse de voir le jour. Je tiens à ce titre à remercier chaleureusement Pascale Ducommun pour notre stimulante collaboration, ainsi que pour avoir partagé avec moi ses connaissances d'hydrogéologue et de précieux moments sur le terrain. Merci à Christoph Aeppli pour sa collaboration dans le cadre de l'étude de l'aquifère de Zuchwil. Je tiens à exprimer ma gratitude à David Weissbrodt, non seulement pour la motivante et fructueuse collaboration à la création du software bioinformatique et à l'écriture du chapitre correspondant, mais également pour les moments de bonne humeur et de rigolade. Merci aussi à Lucas Sinclair pour sa grande implication dans ce même projet, ainsi qu'aux deux étudiants de Master qui ont travaillé dans le cadre de ma thèse, Romain Coulon et Guillaume Laulan. Je ne saurais oublier de remercier Vasantha Aravinthan pour sa précieuse collaboration et sa grande motivation.

Je tiens à remercier très chaleureusement Julien Maillard pour les moments d'amitié partagés ainsi que pour tous ses conseils scientifiques avisés, pour sa collaboration à divers projets qui n'ont pas trouvé place dans ce document mais qui trouveront j'en suis sûr leur place ailleurs.

Un merci tout spécial à Nevenka Adler pour ses conseils rédactionnels et son enthousiasme sans faille. J'adresse également un merci chaleureux à Martina Praveckova pour le brin de folie qu'elle a apporté lors de son passage dans notre bureau.

Un très grand merci à toutes les personnes du LBE, qui ont façonné mon quotidien et ont apporté une ambiance de travail idéale durant ces quelques années : Emmanuelle Rohrbach, Nouhou Diaby, Laure Prat, Aurélie Duret, Yoan Rappaz, Audrey Ducret, Jean-Pierre Kradolfer, Marc Deront, Sonja Desplos, Samuel Lochmatter, Elsa Lacroix, Esther von Arx, et tous les autres.

Je remercie ma famille et mes amis pour leur soutien au jour le jour et leurs encouragements, ainsi que Cheewanon Migliorini et ses collègues du Laténium pour m'avoir trouvé une salle de travail pour les dimanches et jours fériés.

Merci à mes filles Jade et Olive pour avoir pimenté un peu plus mon labeur quotidien, pour m'avoir apporté des moments riches en émotions, et pour m'avoir contraint à quitter l'ordinateur le soir de retour à la maison.

Je tiens finalement à adresser un merci tout particulier à mon épouse Monchaya, qui m'a encouragé à me lancer dans ce projet et qui m'a ensuite apporté de nombreux encouragements et un soutien inconditionnel tout au long de ces années de recherche.

Curriculum Vitae

Noam Shani

PERSONAL DATA

Private address:	Guillaume Ritter 9 CH-2000 Neuchâtel Switzerland
Professional address:	Laboratory for Environmental Biotechnology (LBE) ENAC – IIE Ecole Polytechnique Fédérale de Lausanne (EPFL) Station 6, Bâtiment CH CH-1015 Lausanne Switzerland
Tel:	+41 (0)21 693 57 58
E-mail:	noam.shani@epfl.ch

PROFESSIONAL EXPERIENCE

Since 2007	**PhD thesis**, Laboratory for Environmental Biotechnology (LBE), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. Thesis director: Prof. Christof Holliger Thesis co-director: Dr. Pierre Rossi Thesis title: Assessing the bacterial ecology of organohalide respiration for the design of bioremediation strategies.
2005 – 2007	Laboratory technician, Johnson & Johnson, Le Locle, Switzerland
2005 – 2006	Scientific associate (civil service), LBE – EPFL, Lausanne, Switzerland
2005	Scientific associate, Microbiology Laboratory, Neuchâtel University, Switzerland
2004	Biology teacher, Secondary school Les Cerisiers, Gorgier, Switzerland / Internship in the company DLK Technologies, Le Locle, Switzerland
2003	Scientific associate (civil service), ecology consulting firm Aquarius, Gorgier, Switzerland; Centre Suisse de Cartographie de la Faune, Neuchâtel, Switzerland
2002	Scientific associate (civil service), Swiss Tropical Institute, Basel, Switzerland
2002	Scientific associate, Microbiology Laboratory, Neuchâtel University, Switzerland

EDUCATION

Since 2007	**PhD thesis**, Laboratory for Environmental Biotechnology (LBE), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. Thesis director: Prof. Christof Holliger Thesis co-director: Dr. Pierre Rossi Thesis title: Assessing the bacterial ecology of organohalide respiration for the design of bioremediation strategies
2002	**M.Sc. Biology**, Neuchâtel University, Switzerland Specializations in microbiology and parasitology. Supervisors: Prof. Michel Aragno, Dr. Pierre Rossi M.Sc. Title: 2,4-diacetylphloroglucinol (DAPG) producers in the rhizosphere of wheat : development of probes and primers to assess their presence and abundance in Indian rice-wheat crop rotations.
1996	High school federal diploma, type B (literature), Lycée Cantonal de Porrentruy

PRESENTATIONS AT SCIENTIFIC CONFERENCES

2011	Oral presentation at the Swiss Microbial Ecology meeting, February 2-4, Engelberg, Switzerland. Competition between Fe(III)- and vinyl chloride-reducing bacteria.
2010	Poster presentation at the 69th Annual Assembly of the Swiss Society of Microbiology, June 24-25, Zürich, Switzerland. Bacterial life in a stagnant PCE-contaminated aquifer.
2009	Poster presentation at the 19th Goldschmidt Conference, June 21-26, Davos, Switzerland. Influence of environmental factors on the structures of bacterial communities in a chloroethene-contaminated aquifer. Poster presentation at the 68th Annual Assembly of the Swiss Society of Microbiology, June 4-5, Lausanne, Switzerland. Presence and activity of bacterial communities in a contaminated aquifer: influence of environmental variables. Poster presentation at the 3rd Meeting of Swiss Microbial Ecology, January 28-30, Einsiedeln, Switzerland. Presence and activity of bacterial communities in a contaminated aquifer: influence of environmental variables.
2008	Poster presentation at the 67th Annual Assembly of the Swiss Society of Microbiology, March 1-2, Interlaken, Switzerland. Variability of bacterial communities associated with aquifers contaminated with chlorinated ethenes: role of environmental variables.
2006	Oral presentation at the 65th Annual Assembly of the Swiss Society of Microbiology, March 7-8, Lausanne, Switzerland. Effects of a sudden feed composition change on the microbial community composition and activity of an anaerobic digester.

PUBLICATIONS

Shani N., V. Aravinthan, P. Rossi, G. Laulan, P. Ducommun, and C. Holliger. Impact assessment of pumping parameters on groundwater bacterial communities. Submitted to *Groundwater Monitoring & Remediation*.

Shani N. & Weissbrodt D., L. Sinclair, G. Lefebvre, P. Rossi, J. Maillard, J. Rougemont, and C. Holliger. PyroTRF-ID: a novel bioinformatics approach for the identification of terminal-restriction fragments using microbiome pyrosequencing data. In preparation.

Imfeld G., H. Pieper, **N. Shani**, P. Rossi, M. Nikolausz, I. Nijenhuis, H. Paschke, H. Weiss, and H.H. Richnow (2011) Characterization of groundwater microbial communities, dechlorinating bacteria, and in situ biodegradation of chloroethenes along a vertical gradient. *Water Air & Soil Pollution* 221 (1-4): 107-122.

Imfeld G, **N. Shani**, D. Roesti, **N.** Fromin, B.N. Johri, R. Gaur, P. Rossi, L. Locatelli, F. Poly, and M. Aragno (2006) Diversity of *phl*D alleles in the rhizosphere of wheat cropped under annual rice-wheat rotation in fields of the Indo-Gangetic plains: influence of cultural conditions. *Current Science* 90 (11): 1521-1525.

Gaur R., **N. Shani**, Kawaljeet, B.N.Johri, P. Rossi, and M. Aragno (2004). Diacetylphloroglucinol-producing pseudomonads do not influence AM fungi in wheat rhizosphere. *Current Science* 86 (3): 453-457.