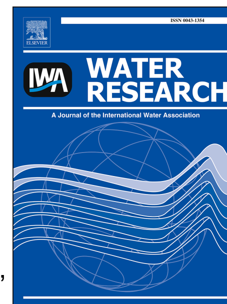
Accepted Manuscript



Successful aerobic bioremediation of groundwater contaminated with higher chlorinated phenols by indigenous degrader bacteria

Anu Mikkonen, Kati Yläranta, Marja Tiirola, Lara Ambrosio Leal Dutra, Pauliina Salmi, Martin Romantschuk, Shelley Copley, Jukka Ikäheimo, Aki Sinkkonen

PII: S0043-1354(18)30229-X

DOI: 10.1016/j.watres.2018.03.033

Reference: WR 13652

To appear in: *Water Research*

Received Date: 17 October 2017

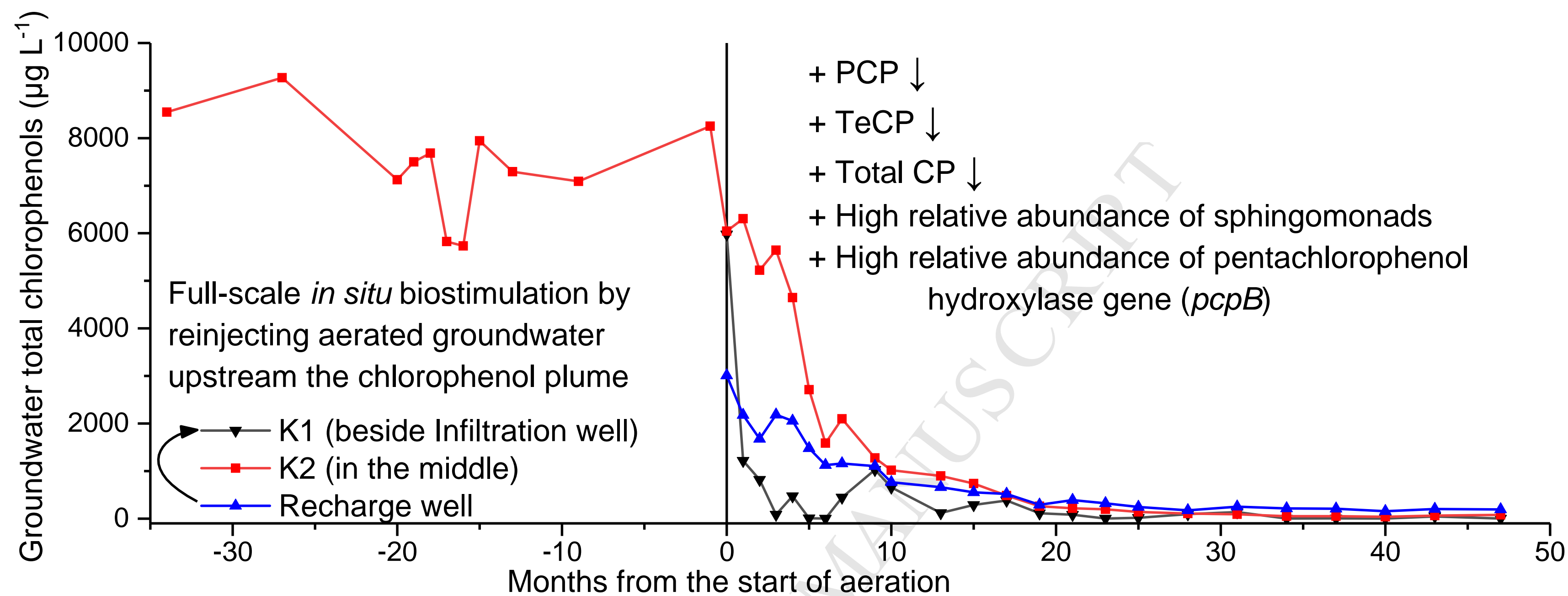
Revised Date: 10 March 2018

Accepted Date: 12 March 2018

Please cite this article as: Mikkonen, A., Yläranta, K., Tiirola, M., Dutra, L.A.L., Salmi, P., Romantschuk, M., Copley, S., Ikäheimo, J., Sinkkonen, A., Successful aerobic bioremediation of groundwater contaminated with higher chlorinated phenols by indigenous degrader bacteria, *Water Research* (2018), doi: 10.1016/j.watres.2018.03.033.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

provided by Helsingin yliopiston digitaalinen arkisto
View metadata, citation and similar papers at core.ac.uk
brought to you by CORE



ACCEPTED MANUSCRIPT

1 # Successful aerobic bioremediation of groundwater contaminated with
2 higher chlorinated phenols by indigenous degrader bacteria

3

4 Anu Mikkonen[a,b]*, Kati Yläranta[a], Marja Tiirola[b], Lara Ambrosio Leal Dutra[b], Pauliina
5 Salmi[b], Martin Romantschuk[a,c], Shelley Copley[d], Jukka Ikäheimo[e], Aki Sinkkonen[a]

6

7 [a]Department of Environmental Sciences, University of Helsinki, Niemenkatu 73, 15140,
8 Lahti, Finland;
9 [b]Department of Biological and Environmental Science, University of Jyväskylä, Survontie 9
10 C, 40500, Jyväskylä, Finland;
11 [c]Institute of Environmental Sciences, Kazan Federal University, 420008, Kremlevskaya 18,
12 Kazan, Russia;
13 [d]Cooperative Institute for Research in Environmental Sciences and Molecular, Cellular and
14 Developmental Biology, University of Colorado, CO 80309 Boulder, The United States;
15 [e]Pöyry Finland Ltd, PL 50, 01621 Vantaa, Finland
16 *Corresponding author: E-mail: anu.mikkonen@jyu.fi

17

18 **Abstract:**
19 The xenobiotic priority pollutant pentachlorophenol has been used as a timber preservative in
20 a polychlorophenol bulk synthesis product containing also tetrachlorophenol and
21 trichlorophenol. Highly soluble chlorophenol salts have leaked into groundwater, causing
22 severe contamination of large aquifers. Natural attenuation of higher-chlorinated phenols
23 (HCPs: pentachlorophenol + tetrachlorophenol) at historically polluted sites has been
24 inefficient, but a 4-year full scale *in situ* biostimulation of a chlorophenol-contaminated
25 aquifer by circulation and re-infiltration of aerated groundwater was remarkably successful:

ACCEPTED MANUSCRIPT

pentachlorophenol decreased from $400\ \mu g\ L^{-1}$ to $<1\ \mu g\ L^{-1}$ and tetrachlorophenols from $4000\ \mu g\ L^{-1}$ to $<10\ \mu g\ L^{-1}$. The *pcpB* gene, the gene encoding pentachlorophenol hydroxylase - the first and rate-limiting enzyme in the only fully characterised aerobic HCP degradation pathway - was present in up to 10% of the indigenous bacteria already 4 months after the start of aeration. The novel quantitative PCR assay detected the *pcpB* gene *in situ* also in the chlorophenol plume of another historically polluted aquifer with no remediation history. Hotspot groundwater HCPs from this site were degraded efficiently during a 3-week microcosm incubation with one-time aeration but no other additives: from $5400\ \mu g\ L^{-1}$ to $1200\ \mu g\ L^{-1}$ and to $200\ \mu g\ L^{-1}$ in lightly and fully aerated microcosms, respectively, coupled with up to 2400% enrichment of the *pcpB* gene. Accumulation of lower-chlorinated metabolites was observed in neither *in situ* remediation nor microcosms, supporting the assumption that HCP removal was due to the aerobic degradation pathway where the first step limits the mineralisation rate. Our results demonstrate that bacteria capable of aerobic mineralisation of xenobiotic pentachlorophenol and tetrachlorophenol can be present at long-term polluted groundwater sites, making bioremediation by simple aeration a viable and economically attractive alternative.

Keywords: *in situ* bioremediation; *pcpB* gene; pentachlorophenol hydroxylase; quantitative PCR; Ion PGM amplicon sequencing; *Sphingomonas sensu lato*

ACCEPTED MANUSCRIPT

51 **1. INTRODUCTION**

52 Pentachlorophenol (PCP) is a xenobiotic with no known natural sources (Crawford et al.
53 2007). Due to its high toxicity and poor biodegradability PCP is a commonly used model
54 pollutant in environmental research. Large-scale production and use as a timber preservative
55 from the 1930s to the 1980s, typically in a chlorophenol mixture with 2,3,4,6-
56 tetrachlorophenol (TeCP) as the main component, resulted in persistent environmental
57 pollution (Bryant and Schultz 1994; Männistö *et al*. 2001, Rautio 2011). In the phenolic form
58 PCP is very poorly soluble in water, but with pK_a of 4.74 the majority of it is present as the
59 phenolate form in near-neutral environments, rendering the sodium and potassium salts over
60 four orders of magnitude more soluble and prone to polluting large aquifers (Olaniran and
61 Igbinosa 2011). The same applies to 2,3,4,6-TeCP with pK_a of 5.38. Some 30 years after
62 banning, these higher chlorinated phenols (HCPs) are still found to contaminate soil and
63 groundwater at Finnish sawmill sites where chlorophenol impregnants were used (Sinkkonen
64 *et al*. 2013).
65
66 One reason for the environmental persistence of PCP may be the lack of microbial degraders
67 capable of catabolising the xenobiotic thermodynamically stable molecule at the prevailing
68 conditions. PCP biodegradation is well possible in anaerobic conditions, but it leads to lower-
69 chlorinated compounds (Bouchard *et al.* 2000; D'Angelo and Reddy 2000). The long-term
70 persistence of the compound in boreal oxygen-deficient groundwaters (constantly around
71 +8°C) and successful degradation after water aeration and fertilisation (e.g. Järvinen *et al.*
72 1994, Tiirola *et al.* 2001a) indicates that aerobic organisms may rule the process at low
73 temperatures. On the other hand, aerobic degradation of chlorophenols may compete with
74 ferrous iron oxidation for the supplied oxygen (see Langwaldt *et al.* 2007 and references
75 therein).Even though PCP degradation has been studied actively already in the 1980s, the

ACCEPTED MANUSCRIPT

76 genes encoding a PCP mineralisation pathway have been fully described only in *Sphingobium*
77 *chlorophenolicum* L-1 (Copley *et al*. 2012; Hlouchova *et al*. 2012, and references therein). Its
78 aerobic oxidative dechlorination pathway initiated by pentachlorophenol hydroxylase (PcpB)
79 is notably inefficient, with PCP turnover to tetrachlorobenzoquinone proceeding with a k_{cat} of
80 0.024 s^{-1}. Interestingly, the same enzyme initiates degradation of TeCP, hydroxylating it
81 directly to tetrachlorohydroquinone, the second metabolite in the PCP mineralisation pathway
82 (Hlouchova *et al*. 2012). TeCP may even be the preferred substrate: according to Tiirola *et al.*
83 (2002b) PcpB is only expressed in *Novosphingobium lentum* MT1 when induced by 2,3,4,6-
84 TeCP. Both degradation monitoring in bacterial liquid cultures (2,3,4,6-TeCP;
85 *Novosphingobium lentum* MT1)(Tiirola *et al.* 2002a) as well as kinetic studies with the
86 purified enzyme (2,3,5,6-TeCP; *Sphingobium chlorophenolicum* L-1)(Hlouchova *et al*. 2012)
87 have confirmed PcpB to be more reactive on TeCP than PCP. We are not aware of any other
88 pathways for TeCPs biodegradation that have been characterised on a genetic and enzymatic
89 level. Also trichlorophenol can be hydroxylated by PcpB (Hlouchova *et al*. 2012), but other
90 pathways exist including that of betaproteobacterial genus *Cupriavidus* initiated by
91 trichlorophenol monooxygenase TcpA (Sánchez and González 2007).
92
93 *pcpB* gene homologues have been detected in multiple *Sphingomonas sensu lato* (i.e.
94 sphingomonad) isolates that degrade PCP in culture, both in the United States and Canada
95 (Crawford *et al.* 2007) as well as Finland (Tiirola *et al.* 2002b). However, there are only two
96 earlier works reporting cultivation-independent detection of the *pcpB* gene (PCR
97 amplification, cloning and sequencing), in Canadian mixed slurry bioreactors (Beaulieu *et al.*
98 2000) and in Montana soil biopiles (Crawford *et al.* 2007). Quantitative polymerase chain
99 reaction (qPCR) assays have greatly aided testing, justification and monitoring of
100 bioremediation of various recalcitrant pollutants, perchloroethene-degrading *Dehalococcoides*

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

101 with reductive dehalogenase genes as one of the benchmark examples (Ritalahti *et al.* 2006).
102 For HCP degraders, no such assays have been available; to our knowledge, there are no
103 earlier reports of PCR determination of the presence, abundance or diversity of any HCP
104 degradation gene at original polluted sites or during HCP biodegradation. Investigation of
105 HCP degrader communities have relied on likely biased and often cumbersome laboratory
106 cultivation protocols (Bécaert *et al.* 2000, Männistö *et al.* 2001).
107
108 The aim of the current work was to establish protocols that enable cultivation-independent
109 investigation of sphingomonads and *pcpB* gene carrying organisms. Novel assays were set up
110 to cover the known diversity of the *pcpB* gene, amplifying shorter (qPCR) and longer
111 (sequencing) fragments of the *pcpB* gene. We hypothesised the *pcpB* gene to be directly
112 detectable at historically contaminated groundwater sites and its relative abundance in
113 bacterial community to increase upon oxygenation of the water. Firstly, the novel assays were
114 tested on plume microbial community samples from Pursiala site with long-term chlorophenol
115 contamination but no bioremediation history. Secondly, groundwater from Pursiala was
116 incubated aerobically in microcosms to quantify changes in chlorophenols and degrader
117 communities. Thirdly, we report chlorophenol dissipation and *pcpB* gene abundance in
118 Kärkölä, another long-term contaminated aquifer, that underwent 4-year *in situ* biostimulation
119 by aeration. Our study sheds light on the potential for aerobic degradation of HCPs and the
120 responsible *in situ* degrader communities.
121

122 2 MATERIALS AND METHODS

123 2.1 Sites and sampling

124 We studied two former sawmill sites in Finland where KY5, a commercial fungicide
125 consisting mostly of tetrachlorophenol (2,3,4,6-TeCP, 75-80%), pentachlorophenol (5-15%)

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

126 and trichlorophenol (2,4,6-TriCP, 5-15%), had been used for timber treatment, resulting in
127 persistent soil and groundwater pollution. In Pursiala, Mikkeli (N61.672, E27.290) KY5 was
128 used for 32 years (1954-1986, with estimated 500-1 500 kg a^{-1}) and in Kärkölä (N60.866,
129 E25.268) for 54 years (1930-1984, with estimated 7 000-10 000 kg a^{-1})(Rautio 2011). At both
130 sites, the highest reported groundwater chlorophenol concentrations have been around
131 100 000 µg L^{-1}. Groundwater chemistry at both sites has been followed up regularly through
132 an extensive system of sampling wells established to monitor the plume. The well casings
133 extended 1 m above ground, they were capped and locked and only accessed by certified
134 sampling personnel, who changed sampling hose between wells to minimize risk of cross-
135 contamination.

136 In Pursiala, site investigation and follow-up has revealed no evidence of significant
137 chlorophenol degradation or biotransformation *in situ* (Rautio 2011). Five wells P1-P5,
138 representing a 350-meter long contamination gradient towards the CP hotspot, were sampled
139 for microbial analyses in April 2013. In June 2016 Pursiala well P5 (hotspot) was sampled
140 again for well bottom sediment slurry (1 L), as well as for groundwater (4 L) 1 m above the
141 well bottom, to test the presence of aerobic PCP degradation potential. Sampling glass bottles
142 were filled to the top and stored in the dark at 4 °C for 7 d before the start of the incubations.
143 For both samplings, we received well-specific background data from the collaborating
144 consultant companies who analysed groundwater samples taken minutes prior to the sediment
145 sampling.
146

147 In Kärkölä, the hotspot (most contaminated soil) has been removed, preventing further CP
148 emissions. Aerobic full-scale closed circle *in situ* bioremediation (biostimulation) was started
149 in June 2012 to clean up the polluted aquifer, following the principle scheme in Figure 1.
150 Anaerobic contaminated water is pumped up from the recharge well, fully aerated by simple

151 mechanical ejector (10-12 mg O_2 L^{-1}), and reinjected through an infiltration well located 200
152 m upstream in the CP plume. Aerated water infiltration rate has been approx. 65 m^3 d^{-1}.
153 Changes in groundwater table and properties during the course of remediation were followed
154 up with an extensive system of monitoring wells (Figure S1). In addition, two wells were
155 sampled for microbial analyses in October 2012 and April 2013 (4 and 10 months of
156 remediation, respectively): well K1 just next to infiltration well and well K2 halfway between
157 the infiltration and recharge wells.
158
159 Standardised methods and accredited commercial laboratories were used in the groundwater
160 analyses. CPs were quantified with GC-MS according to CSN EN 12673, with a detection
161 limit of 0.05 μg L^{-1} for each congener. Groundwater O_2 was measured according to SFS-EN
162 25813, alkalinity according to SFS 3005:1981 and soluble iron according to SFS 3027:1976.
163 Groundwater humic substances were not analysed in this study, but according to earlier
164 references little organic substrates other than chlorophenols are present at either of the sites
165 (Rautio 2011; Tiirola *et al.* 2002a).
166
167 Samples for microbial analyses (DNA extraction) were collected by pumping sedimented
168 particles ('sediment slurry') from the bottom of the groundwater wells established to monitor
169 the plume, at 13-41 m depth, to get more microbial biomass than that present in the
170 groundwater alone. Another reason for sampling solids was to target sphingomonads, which
171 were assumed to be sessile instead of planktonic (Pollock and Armentrout 1999; Tiirola *et al*.
172 2002a). Well-bottom sediment slurry was collected by filling a sterile 50-ml plastic tube to
173 the top, leaving minimal headspace. Chlorophenols are assumed to be highly soluble in these
174 groundwaters of near-neutral pH, but if either limited solubility or biodegradation would
175 affect the chlorophenol concentrations experienced by groundwater (planktonic) vs. sediment

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

176 slurry (sessile) bacteria, this should show as different CP congener distributions in the two
177 matrices. An in-house modification of the standard method (Sinkkonen *et al*. 2013) was used
178 for CP test extractions from slurries from two Kärkölä wells sampled in October 2012; the
179 relative abundances of the different chlorophenol congeners were very similar in water ($\mu g\ L^{-1}$
180 - accredited commercial laboratory) and sediment slurries ($mg\ kg^{-1}$ dwt - in-house method),
181 with Pearson correlation above 0.99.

182

183 **2.2 Nucleic acid extraction**

184 The sampled slurries were allowed to settle in the dark at 4 °C for 4-6 d, after which the clear
185 supernatant was decanted and discarded. Sediment DNA was extracted in duplicate with MO
186 BIO PowerSoil DNA Isolation Kit (MO BIO, Carlsbad, CA, USA). A sufficient pellet for
187 DNA extraction was collected by centrifuging multiple 2-mL volumes of the sedimented
188 slurry at 10 000 g according to manufacturer's recommendations. To maximize DNA yield
189 from clayey low-biomass subsurface sediments, we replaced the kit's PowerBead Tubes
190 with 'G2 Beadbeating Tubes' purchased from GEUS, Copenhagen, Denmark. The G2 tubes
191 are MO BIO PowerLyzer tubes with added G2 blocking agent (modified and fragmented
192 salmon sperm DNA) to prevent immediate irreversible adsorption of released DNA on clay.
193 The sample pellets and PowerSoil Bead Solution from the PowerLyzer Tubes were added to
194 G2 tubes, vortexed briefly and allowed to stand for 5 min before addition of C1 solution and
195 continuation of the extraction according to the manufacturer's protocol. Cells were disrupted
196 by FastPrep FP120 Homogenizer (MP Biomedicals) for 40 s at 6.0 $m\ s^{-1}$. DNA was stored at -
197 20 °C in aliquots.

198

199 DNA yield and quality were analysed with the Quant-IT PicoGreen dsDNA Assay Kit
200 (Invitrogen, Thermo Fisher Scientific) and by agarose gel electrophoresis. The yields varied

201 from 0.04 to 12.6 μg g^{-1} dry sediment but did not correlate with well depth, groundwater O_2,

202 PCP or total CP concentrations. Because sediment sampling by pumping slurry from the

203 groundwater well bottom is unlikely reproducible or representative in terms of particle size

204 distribution, and because DNA extraction efficiency cannot be expected to be equal in

205 sediments of different grain sizes and organic matter contents, we report the final sediment

206 results not as absolute (per sediment dry weight) but as relative abundances.

207

208 **2.3 *pcpB* gene primer design**

209 The *pcpB* gene sequences were retrieved from GenBank, including sequences from pure

210 cultures as reviewed by Crawford *et al*. (2007) as well as clones of Beaulieu *et al*. (2000) and

211 Crawford *et al*. (2007), and aligned in ARB 5.2 (Ludwig *et al*. 2004). The alignment of the

212 *pcpB* gene homologs detected by Saboo and Gealt (1998) in *Proteobacteria* not able to

213 degrade PCP (*S. chlorophenolicum* L-1 *pcpB* gene positions 365-784) did not overlap with the

214 majority of the other sequences and could thus not be used to exclude these non-

215 sphingomonad non-degrader *pcpB* gene variants. As the forward primer pcpB-G by Beaulieu

216 *et al*. (2000) (GGSTTCACSTTCAAYTTCGA, *S. chlorophenolicum* L-1 pos. 250-269)

217 covered the full known diversity, only reverse primers for qPCR (short amplicon) and

218 diversity analyses (longer amplicon) were designed (Table 1), taking into account both

219 coverage as well as thermodynamic properties, and complementarity of the primer pair using

220 NetPrimer (Premier Biosoft, 2013). PrimerBLAST (Ye *et al*. 2012) was used to test *in silico*

221 the specificity of both primer pairs. Specificity of the *pcpB* gene primers was tested also *in*

222 *vitro* with non-CP contaminated reference samples.

223

224 **2.4 Gene enumeration by quantitative PCR: *pcpB*, sphingomonads, *Bacteria***

225 The *pcpB* gene was quantified with pcpB-G and pcpB_356r. qPCR was tested also with the

226 longer amplicon, which showed separation of the samples comparable to the shorter amplicon
227 (copy number Spearman correlation 0.87, $p=6\times10^{-5}$, n=14 Kärkölä DNA extracts), but lower
228 amplification efficiency, as expected. Sphingomonads were quantified with the family-
229 specific 16S rRNA gene primers of Zhou *et al*. (2012) SA429f (5'
230 TAAAGCTCTTTTACCCG3') and SA933r (5'AAACCACATGCTCCACC3'). (Note that in
231 the original reference, the reverse primer is given as in forward strand). Bacterial 16S rRNA
232 genes were quantified with primers pE (5'AAACTCAAAGGAATTGACGG3') and pF'
233 (5'ACGAGCTGACGACAGCCATG3') (Sinkkonen *et al*. 2014) with product length of
234 approx. 170 bp, NetPrimer ratings of 85 and 84 with no cross-dimers, and coverage of 94% of
235 *Bacteria* according to SILVA TestPrime (SILVA 117 RefNR database, one mismatch
236 allowed).
237
238 All qPCR reactions were run with LightCycler 96 (Roche) in white LightCycler 8-Tube Strips
239 (Roche) for increased sensitivity. Triplicate reactions of 20 µL consisted of 1× FastStart
240 Essential DNA Green Master (Roche) supplemented with 0.02% BSA (Fermentas, Thermo
241 Fisher Scientific), 0.5 µM of both primers (Oligomer, Helsinki, Finland) and 2 µL of template
242 (0.4-16 ng). The program consisted of preincubation at 95 °C for 10 min; cycling of melting
243 at 95 °C for 10 s, annealing for 20 s and elongation at 72 °C for 20 s (30 s for the longer
244 sphingomonad-16S rRNA amplicon); melting to 97 °C at 0.1 s^{-1} with 5 readings s^{-1}. The
245 optimised annealing temperatures were 53 °C, 55 °C and 57 °C, and cycle numbers 40, 40 and
246 30, for the *pcpB* gene, sphingomonad-16S rRNA gene and bacterial 16S rRNA gene primers,
247 respectively. Genomic DNA of fully sequenced strains purchased from DSMZ (Leibniz
248 Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig,
249 Germany), extracted with GeneJET Genomic DNA Purification Kit (Fermentas, Thermo
250 Fisher Scientific) and quantified with Quant-IT PicoGreen dsDNA Assay Kit (Invitrogen,

251 Thermo Fisher Scientific), were used as standards: *S. chlorophenolicum* L-1 (DSM 6824) for
252 the *pcpB* gene (3.2×10^0-3.2×10^5 copies) and sphingomonad-16S rRNA gene (9.6×10^0-9.6×10^5
253 copies), *Cupriavidus necator* JMP134 (DSM 4058) for bacterial 16S rRNA gene (8.0×10^1-
254 8.0×10^5 copies). Triplicate no-template-controls (with required threshold cycle >6.7 cycles
255 higher than for the most dilute sample) and duplicate standard series (R^2>0.985) were
256 included in each run using dilution aliquots of each standard, which was re-melted a
257 maximum of three times to minimize degradation. Results were analysed with LightCycler 96
258 SW v.1.1 (Roche). The amplification efficiencies were ~90% for the *pcpB* gene and
259 sphingomonad-16S rRNA gene and nearly 100% for bacterial 16S rRNA gene, with y-
260 intercepts of the standard curves at 34, 38 and 34 cycles, respectively. Bacterial 16S rRNA
261 gene copy numbers were found to correlate with extract DNA concentration (Spearman
262 correlation 0.95, $p=4\times10^{-9}$, n=18 DNA extracts), supporting the reliability of this qPCR assay.
263

264 **2.5 High throughput sequencing by Ion PGM: *pcpB gene* and bacterial 16S rRNA gene**

265 The *pcpB* gene was amplified with pcpB-G and pcpB_512r in DNA Engine DYAD (MJ
266 Research, St. Bruno, Canada) at 50 µL volume of 1× Biotools buffer, 1 U Biotools
267 Polymerase (Biotools Ultratools for 16S rRNA gene; Biotools, Spain), 0.2 mM of each dNTP
268 (Fermentas, Thermo Fisher Scientific), 0.04% BSA (Fermentas, Thermo Fisher Scientific),
269 and 1.0 µM of both primers (Oligomer, Helsinki, Finland). The PCR program consisted of
270 preincubation at 94 °C for 5 min; 40 cycles (*pcpB* gene) or 35 cycles (sphingomonads) of
271 melting at 94 °C for 45 s, annealing at 55 °C for 45 s and elongation at 72 °C for 30 s (40 s for
272 longer sphingomonad-16S rRNA amplicon); final elongation at 72 °C for 5 min. The
273 amplicons were cleaned up with High Pure PCR Product Purification Kit (Roche) for Ion
274 Torrent PGM (Life Technologies, Thermo Fisher Scientific) sequencing. Barcoded
275 sequencing adapters were ligated to the PCR products using the Ion Xpress Plus gDNA

ACCEPTED MANUSCRIPT

276 fragment library kit with Ion Xpress barcode adapters (Life Technologies, Thermo Fisher
277 Scientific). Ion PGM Template OT2 400 (Life Technologies, Thermo Fisher Scientific) was
278 used for the emulsion PCR and Ion PGM 400 Sequencing Kit (Life Technologies, Thermo
279 Fisher Scientific) for final sequencing, which was done on the Ion 314 Chip (Life
280 Technologies, Thermo Fisher Scientific) using the Ion Torrent PGM.
281
282 Bacterial 16S rRNA gene was amplified as in Mikkonen *et al*. (2011) with primers fD1
283 (5´AGAGTTTGATCCTGGCTCAG3´) and PRUN518r (5'ATTACCGCGGCTGCTGG3´) in
284 final volume of 50 µL of 1× Biotools buffer, 1 U Biotools Ultratools Polymerase (Biotools,
285 Spain), 0.2 mM of each dNTP (Fermentas, Thermo Fisher Scientific), 0.05% BSA
286 (Fermentas, Thermo Fisher Scientific), 0.6 µM of both primers (Oligomer, Helsinki, Finland)
287 and 0.5-2 ng of template DNA. The cycling conditions were as for the *pcpB* gene except that
288 the annealing and elongation steps were both 1 min and cycle number was limited to 28. To
289 shorten the fragment size suitable for the sequencing chemistry, the products were re-
290 amplified using barcoded adapter primers A_nn_fD1
291 (5'CCATCTCATCCCTGCGTGTCTCCGACTCAGnnAGAGTTTGATCMTGGCTCAG3'),
292 where nn refers to a 10-12 bp long barcode. Shearing the product, Pippin prep purification of
293 the 460-540 bp long constructs (Sage Science, Beverly, MA, USA) and ligation of the
294 adapter P1 on the other side of the construct was done for pooled samples as previously
295 described (Mäki *et al*. 2016) utilizing the chemistry of the Ion Xpress Plus gDNA fragment
296 library kit (Life Technologies, Thermo Fisher Scientific). Downstream reactions were
297 performed as described for the *pcpB gene* amplicons.
298

299 **2.6 Ion Torrent amplicon sequence data analysis**

300 High-throughput sequence data was analysed in mothur (Schloss *et al*. 2009). The *pcpB* gene

ACCEPTED MANUSCRIPT

301 sequence data were filtered based on quality and length (minimum quality window
302 average=20 at window size=10; maximum primer mismatch=3 and barcode mismatch=2;
303 maximum homopolymer length=6, no ambiguous bases, minimum length=200), the primers
304 and barcodes were trimmed out, and the data were rarefied to 175 sequences per sample.
305 Unique sequences were aligned in ARB with the reference *pcpB* gene sequences used in
306 primer design.

307

308 Raw bacterial 16S rRNA gene amplicon sequences were processed in mothur with the default
309 methods (Schloss *et al*. 2011): filtered based on quality and length (minimum quality window
310 average=15 at window size=10; maximum primer and barcode mismatch=1, maximum
311 homopolymer length=8, no ambiguous bases, minimum length=150) and screened to leave
312 only well-aligned sequences (132-249 bp). These were preclustered with maximum
313 difference=2 (approx. 99% identity), chimeras were detected and removed, remaining
314 sequences were classified against Greengenes taxonomy, based on which sequences of
315 chloroplasts, mitochondria, *Archaea*, *Eukaryota* and unknown kingdom were removed. The
316 remaining 60 000 sequences were classified as *Bacteria* and 94% could be identified to
317 phylum level. ClearCut was used to calculate Neighbour Joining tree based on uncorrected
318 pairwise distances between aligned sequences. Data was rarefied to 969 sequences per sample
319 prior to calculation of phylogenetic α-diversity (command: phylo.diversity). Bacterial 16S
320 rRNA gene sequences from *in situ* samples with MIMARKS details have been submitted to
321 NCBI Sequence Read Archive under BioProject PRJNA349270.

322

323 **2.7 Microcosm incubations for Pursiala groundwate**r

324 Three replicate incubation glass bottles of 162 mL, closed airtight with butyl septum, were
325 filled to the top with partly or fully aerated well P5 groundwater, inoculated with 1% v/v

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

326 settled sediment slurry of 14% dry matter aerated fully (vigorous bubbling and stirring at 300
327 rpm for 2 h to minimize chemical oxygen demand upon incubations), and amended with 10%
328 ignited sand (w/v, grainsize 0.5-1.2 mm, Saint-Gobain Weber, Helsinki, Finland) to facilitate
329 mixing. Preparations and incubations were carried out at 15 °C. The two initial aeration levels
330 used were low, with 2 mg L^{-1} O_2 (16-20% air saturation at 15 °C), and high, with 10 mg L^{-1}
331 O_2 (101-104% air saturation). Bottles were incubated for 21 d in the dark, mixing every other
332 day by gently rotating each bottle 7 times. Oxygen concentration was monitored with non-
333 invasive sensor spots and Microx transmitter (PreSens, Germany). After the last day mixing,
334 sediment was allowed to settle for 1 h before decanting 60 mL water for chlorophenol
335 analysis. The remaining contents of the bottle were shaken horizontally at 200 rpm for 5 min
336 to gently detach microbes from the surfaces and homogenize flocs before sampling the water
337 and suspended sediment for DNA-based analyses.
338
339 Concentrations of chlorophenols were analysed from the groundwater before and after
340 incubations by SGS Inspection Services (Kotka, Finland) according to CSN EN 12673. DNA
341 was extracted from the initial sediment slurry (1.62 mL, matching the inoculation volume)
342 with MO BIO PowerLyzer PowerSoil DNA Isolation Kit (MO BIO, Carlsbad, CA, USA), and
343 from the initial groundwater (150 mL) and incubated water (40 mL) with MO BIO
344 PowerWater DNA Isolation Kit (MO BIO, Carlsbad, CA, USA). The *pcpB* gene was
345 quantified as described above, except for using Maxima SYBR Green/Fluorescein qPCR
346 Master Mix (Thermo Scientific) and Bio-Rad CFX96 Real-Time System (Bio-Rad
347 Laboratories, Inc.). Bacterial communities were characterized by Ion PGM high throughput
348 sequencing of 16S rRNA gene V1-V2 region, amplified from 2.5 ng of template DNA using
349 Maxima SYBR Green/Fluorescein qPCR Master Mix (Thermo Scientific) to amplify first
350 with general bacterial primers 27F (5'AGAGTTTGATCMTGGCTCAG3') and 338R

351 (5'TGCTGCCTCCCGTAGGAGT3') for 30 cycles (anneal at 53 °C) and to append the
352 sequencing primers and barcodes in additional 8 cycles. Products purified with Agencourt
353 AMPure XP purification system (Beckman Coulter Life Sciences, Indianapolis, IN, USA)
354 were pooled for 400 bp library sequencing, template prepared with Ion PGM Hi-Q View OT2
355 Kit and sequenced with Ion PGM Hi-Q View Sequencing Kit on Ion 316 Chip v2 chip (all
356 Life Sciences, Thermo Fisher Scientific). The data was analysed in mothur as described
357 above.
358

359 **3 RESULTS**

360 **3.1 Potential aerobic HCP degraders in chlorophenol plume**

361 The qPCR amplification efficiency with the newly developed primers for pentachlorophenol
362 hydroxylase gene (*pcpB* gene) was ~90% using the standard curve of the positive control
363 strain (*Sphingobium chlorophenolicum* L-1). Both the novel primers produced amplicons of
364 expected size with pcpB-G (Table 1). Product melting temperatures were the same for all the
365 samples, and within 0.5 °C from the melting temperature of the positive control (i.e.
366 standard). Specificity of the primers was successfully verified both *in silico* (no false positives
367 from sequence databases) and *in vitro* (no PCR product detected from non-PCP contaminated
368 reference samples: sandy and humus soils with and without hydrocarbon contamination).
369

370 The *pcpB* gene, however, was successfully detected and quantified by direct qPCR in the CP
371 plume at the historically polluted Pursiala site. Relative abundance of the gene in the bacterial
372 community was highest in the second-most contaminated sampling well (2‰ of bacterial 16S
373 rRNA gene copy number), and approx. tenfold lower in the hotspot, possibly related to lower
374 oxygen concentration (Table 2). The *pcpB* gene copy numbers were at the detection limit of
375 the assay at the midpoint of the studied 350-metre plume (slightly above the limit of 20 copies

ACCEPTED MANUSCRIPT

376 µL^{-1} in one duplicate DNA extract, below it in the other), and below detection limit at the two
377 sampling wells furthest from the hotspot (HCP <30 µg L^{-1} and O_2 <1 mg L^{-1}).

378

379 As expected, the number of bacterial 16S rRNA genes per quantity of extracted DNA was
380 unchanged in the plume, but the abundance of sphingomonads varied hundredfold (Table 2).
381 Significant positive correlation was observed between sphingomonad and *pcpB* gene assay
382 results (Spearman rho=0.73, p=0.016, n=10 extracts). *Proteobacteria* represented over 50% of
383 the bacterial community characterised by high-throughput sequencing (Figure S2).
384 *Betaproteobacteria* was the most abundant class (especially *Burkholderiales*), whereas the
385 relative abundance of *Alphaproteobacteria* (especially *Sphingomonadales*) increased up the
386 plume. The abundant *Sphingomonadaceae* OTUs that could be classified down to the genus
387 level were identified as either *Sphingomonas* or *Novosphingobium*.

388

389

390 **3.2 HCP degradation in aerated microcosms**

391 Regardless of detection of the *pcpB* gene, there was no prior knowledge of the viability and
392 HCP degradation potential of the *in situ* bacterial community at the Pursiala site, where no
393 bioremediation had been studied or attempted. We returned to the site three years after the
394 initial analysis and resampled the hotspot well P5 to test biotic effects of groundwater aeration
395 in microcosms. PCP concentration was unchanged (440 µg L^{-1}), whereas TeCPs were reduced
396 by more than half, still remaining at extreme concentration of 4900 µg L^{-1}. With proportions
397 of 8% and 86% of the total chlorophenol concentration (5700 µg L^{-1}), respectively, this
398 isomer distribution resembles relatively closely that of the original KY5 impregnant.

399

400 Vigorous 2 h aeration in the laboratory *per se*, without incubation, did not cause groundwater

ACCEPTED MANUSCRIPT

401 HCP loss (PCP concentration of fully aerated water was 450 μg L^{-1} and TeCPs 4800 μg L^{-1}),
402 whereas following 3-week incubation at 15 °C in the dark decreased both total PC and HCP
403 concentrations by up to 96% (Figure 2). The original groundwater PCP concentration
404 decreased on average by 13% in the bottles of low initial aeration level (final range 330-440
405 μg L^{-1}) and TeCPs concentration by 83% (final range 220-1500 μg L^{-1}). In the bottles of high
406 initial aeration level, PCP decreased on average by 51% (final range 200-240 μg L^{-1}) and
407 TeCPs on average by 100% (final range 5-6 μg L^{-1}). Trichlorophenols decreased on average
408 by 82% and 99% in the microcosms of high and low initial aeration level, respectively.
409 Proportion of PCP enriched during incubation, being 17-61% of total CPs in the less and 95%
410 in the fully aerated bottles.

411

412 Fully aerated bottles remained aerobic throughout the incubation (endpoint air saturation
413 64%), whereas oxygen was depleted from the less aerated bottles in the middle of the
414 experiment. However, known anaerobic PCP degradation metabolites 2,3,4,5-TeCP and 2,4,5-
415 TriCP (Field and Sierra-Alvarez 2008) did not accumulate but reduced on average by 12%
416 and 49%, respectively. Moreover, also total trichlorophenols decreased on average by 82%,
417 dichlorophenols by 37% and monochlorophenols by 90% in the less aerated bottles, showing
418 no indication anaerobic dehalogenation after oxygen depletion. pH varied from 7.8 to 8.1, at
419 which chlorophenols are highly soluble, and losses due to sorption were likely minimal as no
420 foreign organic matter was added to the glass bottles. 0.5% $AgNO_3$ used to kill abiotic control
421 bottles, on the other hand, had dropped pH to 4-5, rendering PCP and TeCPS in poorly
422 soluble protonated forms. This unfortunately prevented accurate quantification of possible
423 abiotic losses.

424

425 The relative abundance of *Sphingomonadaceae* in both initial sediment and water upon start

ACCEPTED MANUSCRIPT

426 of the incubation was 4%, increasing on average to 50% and 27% in the microaerophilic and
427 fully aerated bottles, respectively (each based on >15 000 bacterial 16S rDNA sequences)
428 (Figure 2). The *pcpB* gene copy number was 1.7×10^{3} copies (ng extracted DNA)$^{-1}$ in the
429 initial sediment and similar 1.5×10^{3} copies (ng extracted DNA)$^{-1}$ (or 6.7×10^{2} mL^{-1}) in the
430 initial water, increasing on average by 2400% (to 5.2×10^{5} copies mL^{-1}) and by 360% (to
431 7.8×10^{4} copies mL^{-1}) in the microaerophilic and fully aerated bottles, respectively.

432

433 **3.3 HCP degradation upon *in situ* aeration**

434 Small-scale aeration tests similar to the one described above had earlier been performed at the
435 Kärkölä site (data not shown), encouraging initiation of full-scale aquifer aeration by closed
436 circle groundwater circulation in June 2012. After 47 months of *in situ* biostimulation,
437 groundwater total CP concentration at the sampling well in the middle of the Infiltration and
438 Recharge wells had decreased from 6000 to <100 µg L^{-1}, TeCPs from 4000 to <10 µg L^{-1},
439 and PCP from 400 to <1 µg L^{-1} (Figure 3). Before aeration the percentage of PCP out of total
440 CPs had been stably 4.5% for three years. After the start of the remediation its relative
441 abundance increased slightly (8-month average 5.6%), but also this most chlorinated isomer
442 was degraded rapidly with a dissipation curve resembling that of TeCPs. Also total
443 chlorophenol concentration decreased (fourth year average in well K2 56 µg L^{-1}), indicating
444 that chlorophenols were not only dechlorinated but mineralised. No major changes were
445 observed in groundwater chloride concentrations (approx. 20 mg L^{-1} close to infiltration well
446 and at recharge well, and 25 mg L^{-1} at K2 between them) were observed, possibly due to the
447 comparatively low concentration of contaminants. However, a decreasing trend over time was
448 seen in groundwater alkalinity in all the three monitoring wells, likely caused by HCl
449 produced upon chlorophenol mineralisation (Figure S3).

450

451 Oxygen concentration in well K1 increased from 0.5 mg L^{-1} to 9.7 mg L^{-1} during the first
452 month of remediation, and groundwater close to infiltration well remained well-aerated (>5
453 mg L^{-1}) throughout the 4 years. In well K2 oxygen concentrations were mostly above
454 detection limit (0.2 mg L^{-1}, up to 2.7 mg L^{-1}) during the first 15 months of remediation, but
455 undetectable after that. Differences in groundwater aeration from the infiltration well to the
456 recharge well were evidenced by differences in groundwater iron content, which was <1 mg
457 L^{-1} in well K1 but approximately 7 mg L^{-1} in the recharge well (Figure S3). Iron precipitation
458 in and around the infiltration well gradually decreases infiltration capacity, and regular
459 removal of the iron sludge and occasional acid wash of the well screen were required to
460 maintain infiltration capacity stable for four years.
461
462 Microbial communities in Kärkölä were sampled 4 and 10 months after the start of the *in situ*
463 aeration. The *pcpB* gene was detected in high abundance of up to several percent of bacteria
464 both next to the infiltration well (K1, aerobic zone) and halfway between infiltration and
465 recharge wells (K2, transition zone), tenfold higher abundance than in the Pursiala well with
466 the highest *pcpB* gene abundance (Table 3). Again significant positive correlation was
467 observed between the sphingomonad and *pcpB* gene copy numbers (Spearman rho=0.76,
468 p=0.031, n=8 extracts). *Alphaproteobacteria* (especially *Sphingomonadales*) were on average
469 more abundant than in Pursiala (Figure S2), corresponding with higher sphingomonad
470 abundance (Table 3). Well K1 October sample bacterial community resembled those of
471 Pursiala, whereas in the next April the candidate phylum TM7 had become curiously
472 abundant. In the transition zone well K2 Epsilonproteobacteria (specifically the sulfur-
473 oxidizing genera *Sulfuricurvum* and *Sulfurimonas*) dominated the bacterial community
474 (Figure S2).
475

476 A longer *pcpB* gene amplicon was successfully amplified from all the Kärkölä samples, as
477 well as from well P4 in Pursiala (the one with the highest *pcpB* gene abundance), enabling
478 investigation of diversity of the gene with Ion PGM high-throughput sequencing. Except for
479 obvious sequencing errors related to difficulties in IonTorrent homopolymer calling (indels
480 resulting in frameshift that would make the enzyme dysfunctional), sequences from all the
481 samples represented a single variant matching the *pcpB* gene of *Novosphingobium lentum*
482 MT1. The same variant was recovered by direct Sanger sequencing of the purified *pcpB* gene
483 amplicons from spring samples (without cloning; successful sequencing with both forward
484 and reverse primers done at the DNA Sequencing and Genomics Laboratory core facility,
485 University of Helsinki, Finland; Data S1).
486

487 **4 DISCUSSION**

488 The use of chlorinated phenols as impregnants in the sawmill industry has caused wide-scale
489 pollution of soil and groundwater. Decades after banning, these pollutants are still found in
490 high concentrations at such sites. As the chlorophenol degradation potential of indigenous
491 microbial community has been shown by cultivated groundwater isolates from the
492 chlorophenol plume (Männistö *et al.* 1999, 2001), the explanation for the persistence of
493 chlorophenols, at least at this site, must lie in unfavourable environmental conditions.
494 Successful degradation of both higher chlorinated phenols (HCPs, i.e. pentachlorophenol and
495 tetrachlorophenol) was observed in an aerated and fertilized bioreactor fed with groundwater
496 from the historically polluted Kärkölä aquifer in the 1990s (reviewed by Langwaldt *et al.*
497 2007), but the efficiency of this pump-and-treat technology was insufficient to reduce aquifer
498 chlorophenol content. Therefore full-scale *in situ* aeration of Kärkölä site by re-infiltration of
499 aerated polluted groundwater upstream of the pumping (recharge) well was started in 2012.
500 Neither additives such as fertilisers nor bioaugmentation was used. Impressive decrease,

ACCEPTED MANUSCRIPT

501 following first or second-order degradation kinetics, was observed in HCP concentrations
502 during the four years of observation (2012-2016). PCP/TeCP ratio increased slightly in the
503 beginning, but biostimulation by aeration brought down the concentration of also the fully
504 chlorinated phenol remarkably efficiently: after few-months lag period, PCP concentration at
505 the long-term monitoring well (halfway the 200 m line between infiltration and recharge
506 wells) came down by >99.5% in four years. To our knowledge, this is the first report of
507 successful aerobic full-scale *in situ* bioremediation of HCP-contaminated aquifer. There is at
508 least one documented case of successful anaerobic (dechlorinative) PCP remediation *in situ*
509 with permeable barrier technology, but this required both bioaugmentation as well as
510 supplemental electron donor (Cole 2000).

511

512 The key to successful aerobic *in situ* remediation in Kärkölä was assumedly the proliferation
513 and competitiveness of native bacteria capable of HCP degradation once oxygen became
514 available. Verified HCP degraders have been isolated from the Kärkölä groundwater and on-
515 site fluidised-bed bioreactor before (Männistö *et al.* 2001, Tiirola *et al.* 2002b), but there was
516 no prior cultivation-independent knowledge of their abundance or competitiveness in the
517 aquifer *in situ* conditions. Iron oxidising bacteria compete for the newly available oxygen, but
518 are according to Langwaldt and Puhakka (2003) less competitive than chlorophenol oxidisers
519 in the contaminated groundwater. Since little readily utilisable substrates other than
520 chlorophenols are present in the Kärkölä groundwater (Tiirola *et al.* 2002a), heterotrophs able
521 to utilise them as carbon and energy source are likely to get competitive advantage.

522

523 We developed novel primers for the *pcpB* gene, encoding the first and rate-limiting enzyme in
524 the only well-known HCP mineralization pathway, which enabled first ever cultivation-
525 independent detection and quantification of HCP degradation potential *in situ*. Notably high

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

526 relative abundance of the *pcpB* gene in the bacterial community was observed in the Kärkölä
527 monitoring well-bottom samples taken during the fastest phase of the degradation, with 3-
528 10% of bacterial cells possessing the pathway (assuming equal qPCR efficiency for the 16S
529 rRNA gene and *pcpB* gene, and that each cell has two copies of the former and one of the
530 latter). In groundwater samples the relative abundance might have been either lower (due to
531 preference for growth attached to surfaces), higher (due to preference for assumedly higher
532 oxygen level in the water), or similar. Anyhow, such a percentage of potential degraders of
533 any persistent pollutant in an *in situ* community is exceptionally high. Considering that the
534 gene was not found in non-chlorophenol polluted reference samples, and even laboratory pure
535 cultures are known to readily discard it in the absence of chlorophenol selection pressure,
536 such high abundance can only result from proliferation of the degraders at the *in situ*
537 conditions, matching well with the first or second-order decrease in HCPs upon aeration.
538 Unfortunately the final proof of *in situ* HCP decrease by PcpB-initiated biodegradation
539 remains to lack, as we cannot conclusively verify growth or increased metabolic activity of
540 the degraders during the remediation; no samples were taken for microbial analyses before the
541 aeration started, and no RNA was extracted for transcript analyses. Moreover, the PcpB-
542 initiated aerobic mineralisation pathway, where the first reaction is the rate-limiting step, is
543 not known to accumulate any metabolites that could be observed in the groundwater.
544
545 Another aquifer historically polluted with chlorophenol timber preservatives, Pursiala, was
546 sampled to investigate the presence and abundance of the *pcpB* gene in HCP plume at a site
547 with no remediation history. The *pcpB* gene was detected in three sampling wells closest to
548 the hotspot. No product was amplified from the two wells further down the plume
549 (groundwater PCP ≤ 3 µg L^{-1}, TeCPs ≤ 25 µg L^{-1}), supporting the specificity of the PCR assay
550 for chlorophenol-contaminated environments. However, this detection of the *pcpB* gene *in*

ACCEPTED MANUSCRIPT

551 *situ*, even if only in contaminated soils, does not mean the organisms present produce
552 functional PcpB at the *in situ* chemical and physical conditions. Moreover, detection of the
553 only well-characterised pathway does not mean that the bacteria carrying it account for any
554 HCP decrease; in fact, in Pursiala the persistence of the pollution suggests the opposite. We
555 thus aerated the Pursiala hotspot groundwater, inoculated it with sediment from the same well
556 (where the *pcpB* gene was earlier detected), and observed changes in chlorophenols, the *pcpB*
557 gene and bacterial community in microcosms. After 21-day incubation at 15 °C, HCP
558 concentration decreased by 96% and the *pcpB* gene increased by 360% in the fully aerated
559 microcosms, whereas HCP concentration decreased by 77% and the *pcpB* gene increased by
560 2400% in the lightly aerated microcosms. These results provide strong evidence that
561 catabolically capable degraders (carrying the *pcpB* gene) were present at the contaminated
562 aquifer and could be activated by aeration to efficiently degrade both TeCPs and PCP. Lower
563 HCP degradation rate in the less-aerated microcosms was likely caused by lack of measurable
564 oxygen after day 10, ceasing aerobic biodegradation - however, no increase in anaerobic
565 dehalogenation products was observed either. Lower counts of the *pcpB* gene (and lower
566 relative and absolute abundance of sphingomonads, the assumed host taxon) in the fully than
567 the lightly aerated microcosms is more curious. This may be explained by the fact that PcpB
568 itself causes oxidative stress by releasing H_2O_2 in uncoupled, futile cycles - the enzyme is
569 notably inefficient and apparently represents "evolution in progress" (Hlouchova *et al*. 2012).
570 Indeed, *Novosphingobium lentum* MT1, the dominant strain from the Kärkölä chlorophenol-
571 treating bioreactor ran in the late 1990s, was found to be microaerophilic and to degrade
572 HCPs more efficiently at 8 °C than at room temperature (Tiirola *et al.* 2002a).
573

574 Other suggested routes for bacterial PCP degradation include the actinobacterial (formerly
575 rhodococcal but renamed mycobacterial) pathway initiated by cytochrome P-450 (Uotila *et al.*

576 1992) and strictly anaerobic dehalogenases (Field and Sierra-Alvarez 2008). Unfortunately,
577 the genes are not well-characterized and specific primers to study either of these pathways
578 are, to our knowledge, not available; shotgun metagenomics or metaproteomics on controlled
579 study systems could shed light on their relevance *in situ*. However, *Mycobacterium* was not
580 abundant in the bacterial communities of the aerated chlorophenol-degrading systems,
581 *Mycobacteriaceaea* relative abundances being <0.5% in the Kärkölä wells and <0.03% in the
582 Pursiala microcosms. In poorly aerated contaminated environments especially the anaerobic
583 degradation pathways would be highly beneficial, and indeed some of the lower-chlorinated
584 phenols observed in the aquifer may originate from reductive dechlorination. Unfortunately
585 there is no indication of such natural metabolism having markedly decreased contaminant
586 concentrations in Pursiala (Rautio 2011); one reason for this may be that anaerobic
587 biodegradation of chlorophenols generally requires addition of organic electron donors (Field
588 and Sierra-Alvarez 2008). Former suggested anaerobic degraders did not seem to be major
589 members of the indigenous communities; for example *Clostridium* and *Desulfitobacterium*
590 were absent or formed max 1‰ of the bacterial communities in the Kärkölä and Pursiala *in*
591 *situ* samples. Accumulation of lower-chlorinated phenols, typical of anaerobic dechlorination,
592 was not observed upon bioremediation treatments either, even after running out of oxygen
593 (Kärkölä Recharge well or Pursiala less-aerated microcosms). Also fungi are known to have
594 several pathways for PCP modification and at least partial degradation, but the associated
595 enzymes do not seem to be chlorophenol-specific but cometabolic (Field and Sierra-Alvarez
596 2008). Our findings suggest that the curious aerobic HCP mineralization pathway initiated by
597 pentachlorophenol hydroxylase PcpB is present and provides some competitive advantage
598 even in Pursiala non-aerated plume. Also in Pursiala groundwaters with little other carbon
599 sources (Rautio 2011), HCPs can serve as an abundant source of carbon and energy, if any
600 oxygen becomes available. PcpB-initiated pathway grants the host cell access to this rich

601 resource, but causes marked metabolic stress to the host itself (Hlouchova *et al*. 2012). The
602 *pcpB* gene was both transferred horizontally and discarded from the genomes of previously
603 studied isolates surprisingly readily, especially considering that the gene is likely present as a
604 single copy situated in a chromosome, not in a plasmid (Tiirola *et al*. 2002a; Tiirola *et al*.
605 2002b; Copley *et al*. 2012). The presence of the *pcpB* gene might thus serve as a "biosensor",
606 indicating the current situation - bioavailability and/or ecological effect of HCPs - in the *in*
607 *situ* bacterial community.
608
609 Earlier studies on culturable isolates have established that the *pcpB* gene is found exclusively
610 in sphingomonads (Tiirola *et al*. 2002b; Copley *et al*. 2012). Amplification of fragments
611 somewhat homologous to the *pcpB* gene has been reported also from non-degrader
612 *Gammaproteobacteria* and *Betaproteobacteria* (Saboo and Gealt 1998), but there are no
613 reports of non-sphingomonads being able to utilize the gene for HCP degradation. Our
614 cultivation-independent results are in agreement with the *pcpB* gene being present exclusively
615 in sphingomonads, as positive correlation was observed between the two at both sites.
616 Slightly higher (max. twofold) counts for the *pcpB* gene than sphingomonad 16S rRNA genes
617 were detected in some extracts, but this could be explained by the slight variability in the
618 amplification efficiencies of the differently sized amplicons (107 vs. 504 bp, respectively).
619 *Betaproteobacteria*, especially *Burkholderiales*, were even more abundant in the bacterial
620 communities than *Sphingomonadales*. This order contains many well-characterized degrader
621 taxa, for example genus *Cupriavidus* known for trichlorophenol degradation potential
622 (Sánchez and González 2007). As no correlation was observed between their and the *pcpB*
623 gene relative abundance (Spearman rho=-0.25, p=0.31, n=18 extracts), *Betaproteobacteria*
624 were unlikely the class carrying this gene.
625

626 Sequencing (both Sanger and next generation) revealed the presence of only one variant of the
627 *pcpB* gene in Kärkölä and Pursiala. The Sanger sequences of the amplified 263 bp fragment
628 showed 99% identity to the *pcpB* gene of *Novosphingobium lentum* MT1 (two mismatches,
629 synonymous, other of which at the degenerated base of the reverse primer). The same variant
630 of the gene has been earlier detected in multiple sphingomonad isolates, representing several
631 genera, cultivated from the on-site aerobic bioreactor that processed contaminated Kärkölä
632 groundwaters (Tiirola *et a*l. 2002b), as well as in some of the Canadian mixed slurry
633 bioreactor clones reported by Beaulieu *et al*. (2000). The complete lack of diversity in such an
634 abundant degradation gene at two long-term contaminated sites 150 km apart seems
635 surprising. Three possible explanations include limited distribution of the other *pcpB* gene
636 variants, high positive selection for the observed variant and/or limited divergence/evolution.
637 We never detected the *pcpB* gene in soils that were pristine or oil-polluted, and both *S.*
638 *chlorophenolicum* L-1 and *N. lentum* MT1 are notorious for their readiness to discard the
639 gene in cultivation (Tiirola *et al*. 2000a; Copley *et al*. 2012). PCP-polluted sites may thus
640 represent the only "islands" where the pathway is maintained and multiplied (via reproduction
641 and horizontal gene transfer). Out of the several *pcpB* gene variants detected in North
642 America - or other thus far unknown center of origin and diversity - perhaps only one has
643 made its way to the studied sites in Finland. It is also possible that the MT1-type *pcpB* gene is
644 selected for in aquatic environments, such as Kärkölä and Pursiala groundwaters and the
645 Canadian mixed slurry bioreactors of Beaulieu *et al*. (2000). However, we detected the same
646 variant of the gene, and only that, also in chlorophenol-contaminated surface soils at a third
647 site in Finland over 200 km away from the sites of this study (unpublished data).

648

649 **5 CONCLUSIONS**

650 • Novel PCR-based assays were developed that enable investigation of the presence,

ACCEPTED MANUSCRIPT

651 abundance and diversity of all known variants of the gene encoding pentachlorophenol
652 hydroxylase (*pcpB*).

- 653 Applying these assays, we were for the first time able to detect and quantify a gene
654 specific for degradation of higher chlorinated phenols (HCPs: pentachlorophenol and
655 tetrachlorophenol) *in situ* at polluted sites.
- 656 Detection of *pcpB* in chlorophenol plume and its high abundance during aerobic HCP
657 degradation both *ex situ* and *in situ* indicate that the specific degraders are present and
658 can be activated by groundwater aeration.
- 659 Cultivation-independent analysis supported the earlier cultivation-dependent results of
660 *pcpB* being present only in sphingomonads, and of only one of the *pcpB* variants being
661 found in Finland.
- 662 Bacteria capable of aerobic mineralisation of xenobiotic HCPs present at long-term
663 polluted groundwater sites make bioremediation by simple aeration a viable and
664 economically attractive alternative.

665

666 **Acknowledgements**: The three anonymous reviewers are acknowledged for their insightful
667 and constructive criticism that helped to significantly improve the manuscript. We thank the
668 consultants working at the Pursiala site, Iiro Kiukas/Ramboll and Arto Itkonen/FCG, for
669 helpful discussions, sampling and providing the groundwater data. Anita Mäki is thanked for
670 help with Ion PGM sequencing and Ivonne De La Garcia/Pöyry for drawing the Kärkölä
671 remediation schematic graph.

672

673 **Funding:** This work was supported by the Academy of Finland [grant numbers 139847,
674 260797]; the European Research Council (ERC) under the European Union's Seventh
675 Framework Programme [FP/2007-450 2013, grant number 615146]; and Koskinen Oy

ACCEPTED MANUSCRIPT

676 (Kärkölä remediation and groundwater analyses, including consultancy by Jukka Ikäheimo).
677 The other authors declare no conflict of interest. Funding sources had no involvement in the
678 interpretation of data, writing of the report or the decision to submit the article for publication.

679

680 **Supplements**

681 Figure S1. Map of the Kärkölä site

682 Table S1. Original microbiological data (duplicate DNA extracts)

683 Figure S2. Bacterial community composition in Pursiala and Kärkölä wells

684 Figure S3. Differences and changes in groundwater alkalinity and iron concentration at the
685 Kärkölä site upon 4-year closed-circle aerobic *in situ* biostimulation

686 Data S1. The *pcpB* gene sequences from Pursiala and Kärkölä

687

688 **REFERENCES**

689 Beaulieu, M., Becaert, V., Deschenes, L., and Villemur, R. (2000) Evolution of bacterial
690 diversity during enrichment of PCP-degrading activated soils. Microb Ecol 40: 345-355.

691 Bécaert, V., Deschênes, L., and Samson, R. (2000) A simple method to evaluate the
692 concentration of pentachlorophenol degraders in contaminated soils. FEMS Microbiol Lett
693 184: 261–264.

694 Bouchard, B., Beaudet, R., Villemur, R., McSween, G., Lépine, F., and Bisaillon, J.G. (1996)
695 Isolation and characterization of *Desulfitobacterium frappieri* sp. nov., an anaerobic
696 bacterium which reductively dechlorinates pentachlorophenol to 3-chlorophenol. Int J Syst
697 Bacteriol 46: 1010–1015.

698 Bryant, S.E., and Schultz, T.W. (1994) Toxicological assessment of biotransformation
699 products of pentachlorophenol: Tetrahymena population growth impairment. Arch Environ
700 Contam Toxicol 26: 299–303.

701 Cole, J.D. (2000) Demonstration of a permeable barrier technology for the in-situ
702 bioremediation of pentachlorophenol contaminated groundwater. Dissertation for the degree
703 of PhD (Civil Engineering). Oregon State University, Oregon, United States.

704 Copley, S.D., Rokicki, J., Turner, P., Daligault, H., Nolan, M., and Land, M. (2012) The
705 whole genome sequence of *Sphingobium chlorophenolicum* L-1: insights into the evolution of
706 the pentachlorophenol degradation pathway. Genome Biol Evol 4: 184-198.

707 Crawford, R.L., Jung, C.M., and Strap, J.L. (2007) The recent evolution of pentachlorophenol
708 (PCP)-4-monooxygenase (PcpB) and associated pathways for bacterial degradation of PCP.
709 Biodegradation 18: 525-539.

ACCEPTED MANUSCRIPT

710 D'Angelo, E.M.D., and Reddy K.R. (2000) Aerobic and anaerobic transformations of
711 pentachlorophenol in wetland soils. Soil Sci Soc Am J 64: 933-943.

712 Field, J.A., and Sierra-Alvarez, R. (2008) Microbial degradation of chlorinated phenols. Rev
713 Environ Sci Biotechnol 7: 211-241.

714 Hlouchova, K., Rudolph, J., Pietari, J.M.H., Behlen, L.S., and Copley, S.D. (2012)
715 Pentachlorophenol hydroxylase, a poorly functioning enzyme required for degradation of
716 pentachlorophenol by *Sphingobium chlorophenolicum*. Biochemistry 51: 3848-3860.

717 Järvinen, K.T., Melin, E.S., and Puhakka, J.A. (1994) High-rate bioremediation of
718 chlorophenol-contaminated groundwater at low temperatures. Environ Sci Technol 28: 2387-
719 2392

720 Langwaldt, J.H., and Puhakka, J.A. (2003) Competition for oxygen by iron and 2,4,6-
721 trichlorophenol oxidizing bacteria in boreal groundwater. Water Res 37: 1378-1384.

722 Langwaldt, J.H., Tiirola, M., and Puhakka, J.A. (2007) Microbial adaptation to boreal
723 subsurface environment: implications in bioremediation of polychiorophenol contaminated
724 groundwater. In Psychrophiles: From Biodiversity to Biotechnology. Eds. Margesin R.,
725 Schinner, F., Marx, J.-C. & Gerday, C. Springer Verlag, pp. 409-428.

726 Ludwig, W., Strunk, O., Westram, R., Richter L., Meier, H., Yadhukumar, *et al*. (2004) ARB:
727 a software environment for sequence data. Nucl Acids Res 32:1363-1371.

728 Mahmood, S., Paton, G.I., and Prosser, J.I. (2005) Cultivation-independent *in situ* molecular
729 analysis of bacteria involved in degradation of pentachlorophenol in soil. Environ Microbiol
730 7: 1349-1360.

731 Mikkonen, A., Lappi, K., Wallenius, K., Lindström, K., and Suominen L. (2011) Ecological
732 inference on bacterial succession using curve-based community fingerprint data analysis,
733 demonstrated with rhizoremediation experiment. FEMS Microbiol Ecol 78: 604-616.

734 Mäki, A., Rissanen, A. J., and Tiirola, M. 2016. A practical method for barcoding and size-
735 trimming PCR templates for amplicon sequencing. BioTechniques 60: 88-90.

736 Männistö, M.K, Salkinoja-Salonen, M.S., and Puhakka, J.A. (2001) *In situ* polychlorophenol
737 bioremediation potential of the indigenous bacterial community of boreal groundwater. Water
738 Res 35: 2496-2504.

739 Männistö, M.K., Tiirola, M.A., Salkinoja-Salonen, M.S. Kulomaa, M.S., and Puhakka, J.A.
740 (1999) Diversity of chlorophenol-degrading bacteria isolated from contaminated boreal
741 groundwater. Arch Microbiol 171: 189-197.

742 Olaniran, A.O., and Igbinosa, E.O. (2011) Chlorophenols and other related derivatives of
743 environmental concern: properties, distribution and microbial degradation processes.
744 Chemosphere 83: 1297-1306.

745 Pollock, T.J., and Armentrout, R.W. (1999) Planktonic/sessile dimorphism of polysaccharide-
746 encapsulated sphingomonads. J Ind Microbiol Biotech 23: 436-441.

747 Rautio, J. (2011) Detrimental substances in the Pursiala groundwater area, Mikkeli. Thesis for
748 the degree of M.Sc.(Tech.). 113 pages. Lappeenranta University of Technology, Faculty of
749 Environmental Technology. (in Finnish)

750 Ritalahti, K.M., Amos, B.K., Sung, Y., Wu, Q., Koenigsberg, S.S., and Löffler, F.E. (2006)
751 Quantitative PCR targeting 16S rRNA and reductive dehalogenase genes simultaneously
752 monitors multiple *Dehalococcoides* strains. Appl Environ Microbiol 72: 2765-2774.

ACCEPTED MANUSCRIPT

753 Saboo, V.M., and Gealt, M.A. (1998) Gene sequences of the *pcpB* gene of pentachlorophenol-
754 degrading *Sphingomonas chlorophenolica* found in nondegrading bacteria. Can J Microbiol
755 44:667-675.

756 Sánchez, M.A., and González, B. (2007) Genetic characterization of 2,4,6-trichlorophenol
757 degradation in *Cupriavidus necator* JMP134. Appl Environ Microbiol 73: 2769-2776.

758 Schloss, P.D., Gevers, D., and Westcott, S.L. (2011). Reducing the effects of PCR
759 amplification and sequencing artifacts on 16S rRNA-based studies. PloS ONE 6: e27310.

760 Schloss, P.D., Westcott, S.L., Ryabin T., Hall, J.R., Hartmann, M., Hollister, E.B., *et al*.
761 (2009). Introducing mothur: open-source, platform-independent, community-supported
762 software for describing and comparing microbial communities. Appl Environ Microbiol 75:
763 7537-7541.

764 Sinkkonen, A., Kauppi, S., Simpanen, S., Rantalainen, A.-L., Strömmer, R., and
765 Romantschuk, R. (2013) Layer of organic pine forest soil on top of chlorophenol
766 contaminated mineral soil enhances contaminant degradation. Environ Sci Pollut Res 20:
767 1737-1745.

768 Sinkkonen, A., Ollila, S., Romantschuk, M. (2014) Changes in *TcpA* gene frequency explain
769 2,4,6-trichlorophenol degradation in mesocosms. J Environ Sci Health, Part B 49: 756–759.

770 Tiirola, M.A., Männistö, M.K., Puhakka, J.A., and Kulomaa, M.S. (2002a). Isolation and
771 characterization of *Novosphingobium* sp. strain MT1, a dominant polychlorophenol-degrading
772 strain in a groundwater bioremediation system. Appl Environ Microbiol 2002: 173-180.

773 Tiirola, M.A., Wang, H., Paulin, L., and Kulomaa, M.S. (2002b). Evidence for natural
774 horizontal transfer of the *pcpB* gene in the evolution of polychlorophenol-degrading
775 sphingomonads. Appl Environ Microbiol 68: 4495-4501.

776 Uotila, J.S., Kitunen, V.H., Saastamoinen, T., Coote, T., Häggblom, M.M., and Salkinoja-
777 Salonen, M.S. (1992) Characterization of aromatic dehalogenases of *Mycobacterium*
778 *fortuitum* CG-2. J Bacteriol 174: 5669-5675.

779 Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T. (2012) Primer-
780 BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC
781 Bioinformatics. 13: 134.

782 Zhou, L., Li, H., Zhang, Y., Wang, Y., Han, S., and Xu, H. (2012) Abundance and diversity
783 of *Sphingomonas* in Shenfu petroleum-wastewater irrigation zone, China. Environ Sci Pollut
784 Res 19: 282-294.

785

ACCEPTED MANUSCRIPT

Table 1. Novel *pcpB* reverse primers. The degenerate positions are underlined. Positions refer to *S. chlorophenolicum* L-1 *pcpB* (total size 1617 bp).

Primer name	Application	Sequence (5'=>3')	Positions	Amplicon length with pcpB-G	Netprimer rating
pcpB_356r[1)]	qPCR	TCGGTCTCATTCTGGTTGTAG TCGGTCTCATTCTGGTTATAG TCGGTCTCGTTCTGATTGTAG	336-356	107 bp	92-100
pcpB_512r	Sequencing	CCGATCACCCAGCGYGG	496-512	263 bp	75-88

[1)]Equimolar mix of three primers - degeneracy forms not detected are not included

ACCEPTED MANUSCRIPT

1 **Table 2.** Chlorophenol and oxygen concentrations, as well as bacterial, sphingomonad and *pcpB* gene copy numbers and bacterial α-diversity, up
2 the chlorophenol plume in Pursiala site (no active remediation). Groundwater samples showing contamination level were taken as for the regular
3 monitoring, typically 1 m above the well bottom (the wells were perforated from the bottom to a height of 10-30 meters). DNA-based results are
4 averages of duplicate DNA extractions from well-bottom sediment (both presented in Table S1).

Well code	Well depth (m)	Sampling time	Water O_2 (mg L^{-1})	Water total CPs (µg L^{-1})[1)]	Water PCP (µg L^{-1})	Water TeCPs (µg L^{-1})	Water TriCPs (µg L^{-1})	Bacterial 16S rRNA gene copies (/ng DNA)	Sphingomonad 16S rRNA gene copies (/ng DNA)	*pcpB* copies (/ng DNA)	*pcpB*/bacterial 16S rRNA gene copies (%)	Bacterial Phylogenetic diversity
P1	41	Apr-2013	0.7	49	3	25.2	11	4.1×10^5	1.2×10^2	$<5.1 \times 10^0$	<0.001	47
P2	23	Apr-2013	0.6	1	<0.05	0.53	0.19	5.2×10^5	6.0×10^2	$<6.3 \times 10^1$	<0.010	31
P3	22	Apr-2013	0.5	841	74	591	140	4.1×10^5	1.2×10^3	$\leq 1.1 \times 10^1$	≤0.003	33
P4	32	Apr-2013	1.2	7380	390	5870	108	6.4×10^5	1.0×10^4	1.4×10^3	0.23	19
P5	20	Apr-2013	0.6	14000	450	13000	496	7.4×10^5	3.5×10^3	2.4×10^2	0.03	27

[1)]For comparison, total chlorophenol concentration in drinking water must be <10 µg L^{-1}.

5

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

1 **Table 3.** Chlorophenol and oxygen concentrations, as well as bacterial, sphingomonad and *pcpB* gene copy numbers and bacterial α-diversity,
2 after 4 and 10 months of *in situ* aeration of Kärkölä site. Groundwater samples were taken as for the regular monitoring, 1.5-2.6 m above the well
3 bottom. DNA-based results are averages of duplicate DNA extractions from well-bottom sediment (both presented in Table S1).

Well code	Well depth (m)	Sampling time	Water O_2 (mg L^{-1})	Water total CPs (µg L^{-1})[1)]	Water PCP (µg L^{-1})	Water TeCPs (µg L^{-1})	Water TriCPs (µg L^{-1})	Bacterial 16S rRNA gene copies (/ng DNA)	Sphingomonad 16S rRNA gene copies (/ng DNA)	*pcpB* copies (/ng DNA)	*pcpB*/bacterial 16S rRNA gene copies (%)	Bacterial Phylogenetic diversity
K1	12	Oct-2012	7.2	470	24	210	112	4.5×10^5	1.3×10^4	1.2×10^4	2.7	20
		Apr-2013	8.9	657	16	232	301	3.2×10^5	8.1×10^3	1.4×10^4	4.2	32
K2	16	Oct-2012	0.4	4650	240	3130	1050	3.0×10^5	6.1×10^3	1.4×10^4	4.6	18
		Apr-2013	<0.2	1020	48	648	273	4.4×10^5	2.1×10^3	6.0×10^3	1.3	23

[1)] For comparison, total chlorophenol concentration in drinking water must be <10 µg L^{-1}.

4

ACCEPTED MANUSCRIPT

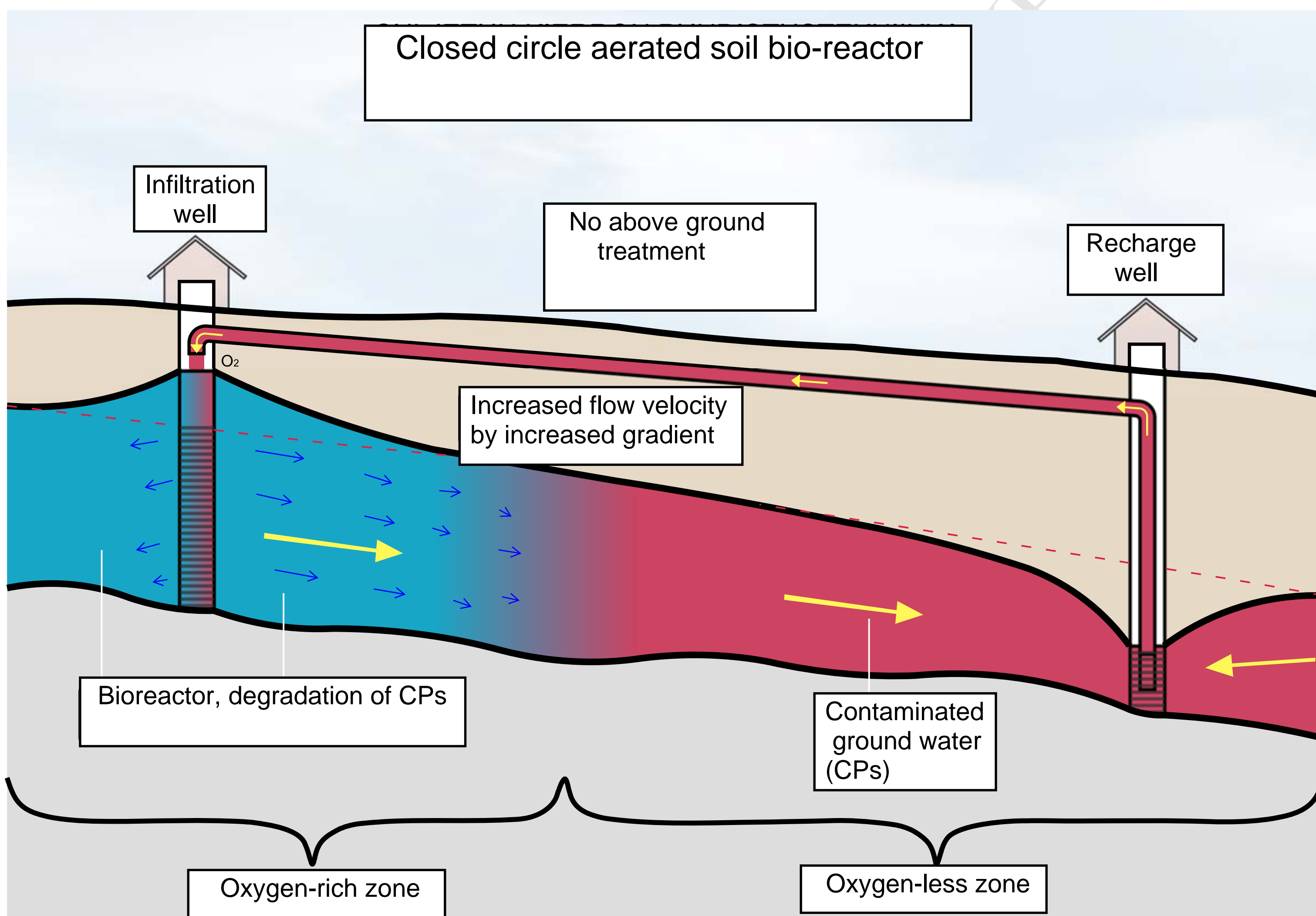


Figure 1. Principle of Kärkölä *in situ* bioremediation of chlorophenols (CPs).

ACCEPTED MANUSCRIPT

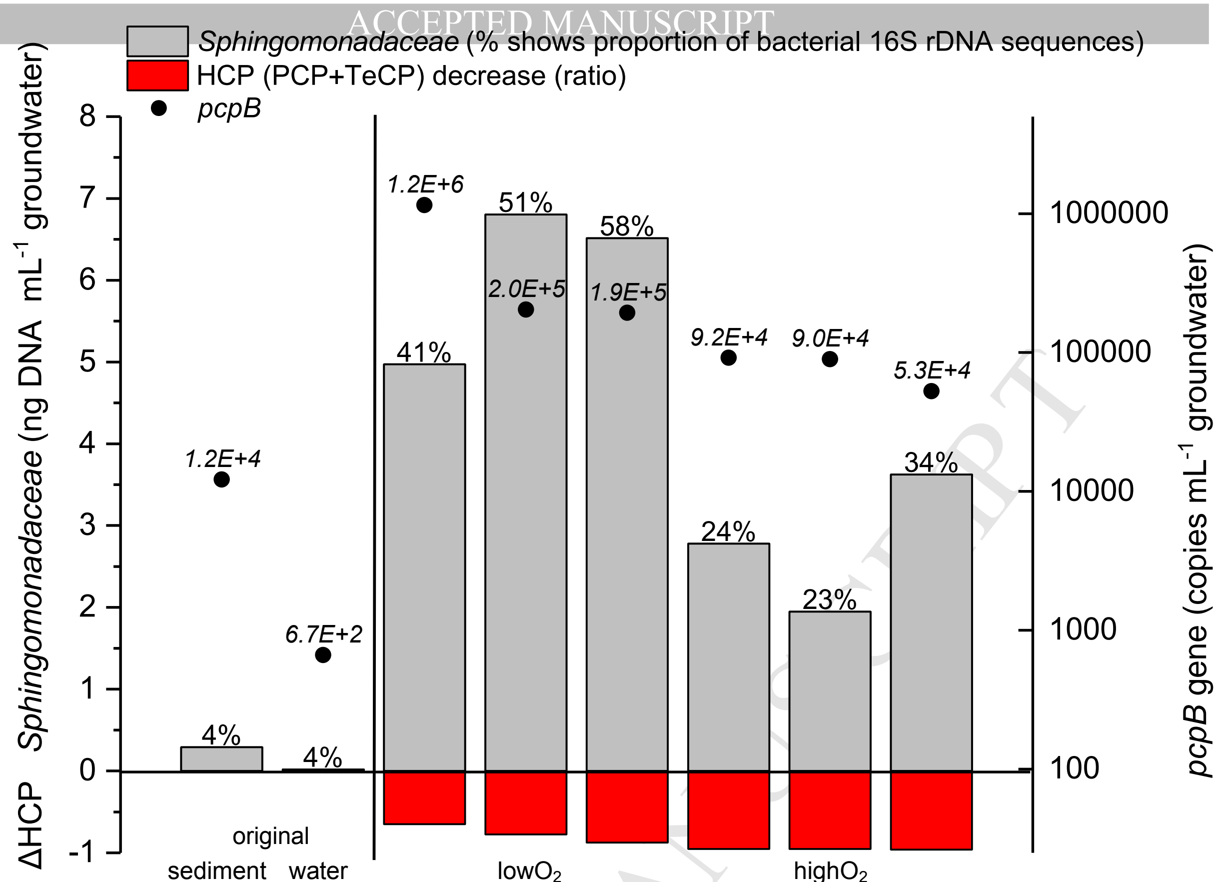


Figure 2. Decrease in Pursiala groundwater HCPs coupled with increase in *Sphingomonadaceae* and the *pcpB* gene upon aerobic incubation in microcosms. Groundwater was aerated to low (2 mg L^{-1}) or high (10 mg L^{-1}) O_2 concentration, amended with 1% sediment slurry from the same well and incubated in tightly sealed triplicate bottles for 21 d at 15 °C in the dark. The abundance of sphingomonads was calculated by multiplying the total microbial biomass (DNA yield) by the relative abundance of sphingomonads (proportion *Sphingomonadaceae* in bacterial communities analysed by 16S rRNA gene high throughput sequencing).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

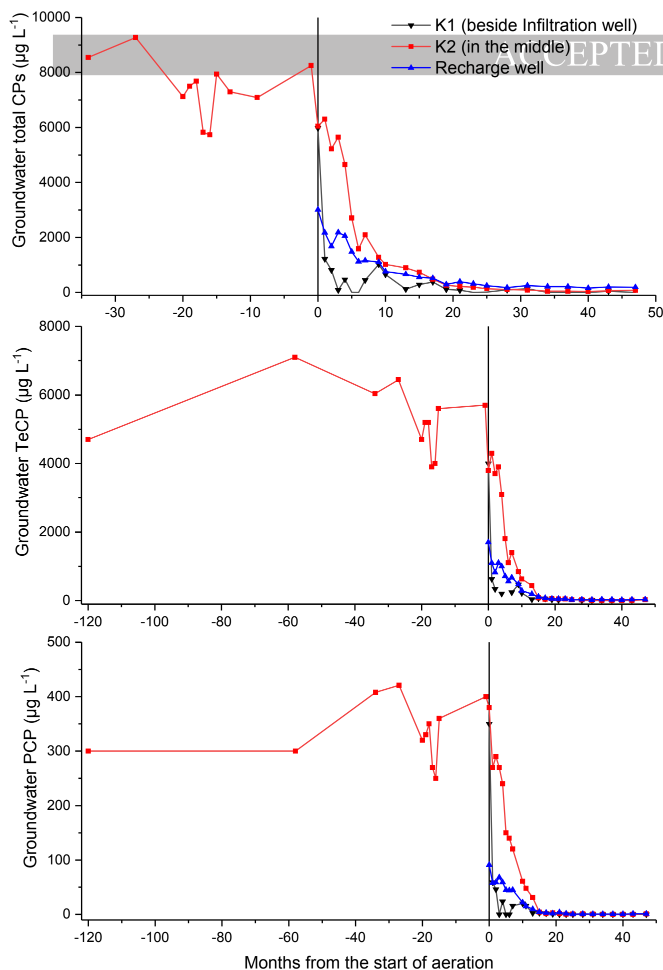


Figure 3. Decrease of chlorophenol (CP) concentrations in groundwater at the Kärkölä site upon 4-year closed-circle aerobic *in situ* biostimulation. Note the different time scales on x-axes; before year 2009 only the main fractions of the polluting KY5 impregnant were quantified, after that all the 18 chlorophenol isomers.

Highlights

- Simple groundwater aeration successfully remediated chlorophenol-polluted aquifer
- New PCR-assay showed up to 10% abundance of pentachlorophenol hydroxylase gene *pcpB*
- *pcpB* was found *in situ* also at other aquifer with historical chlorophenol pollution
- *pcpB* and sphingomonads multiplied upon aerobic incubation in microcosms
- 51% of pentachlorophenol and 100% of tetrachlorophenol was depleted in 3 weeks

ACCEPTED MANUSCRIPT