



**HAL**  
open science

# Taxonomic and functional trait-based approaches suggest that aerobic and anaerobic soil microorganisms allow the natural attenuation of oil from natural seeps

Aurélie Cébron, Adrien Borreca, Thierry Beguiristain, Coralie Biache, Pierre Faure

## ► To cite this version:

Aurélie Cébron, Adrien Borreca, Thierry Beguiristain, Coralie Biache, Pierre Faure. Taxonomic and functional trait-based approaches suggest that aerobic and anaerobic soil microorganisms allow the natural attenuation of oil from natural seeps. *Scientific Reports*, 2022, 12 (1), pp.7245. 10.1038/s41598-022-10850-4. hal-03805168

**HAL Id: hal-03805168**

**<https://hal.univ-lorraine.fr/hal-03805168>**

Submitted on 7 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1       **Taxonomic and functional trait-based approaches suggest that**  
2       **aerobic and anaerobic soil microorganisms allow the natural**  
3       **attenuation of oil from natural seeps**

4  
5  
6   Aurélie Cébron<sup>1,\*</sup>, Adrien Borreca<sup>1</sup>, Thierry Beguiristain<sup>1</sup>, Coralie Biache<sup>1</sup>, Pierre Faure<sup>1</sup>

7  
8   <sup>1</sup> *Université de Lorraine, CNRS, LIEC, F-54000, Nancy, FRANCE*

9  
10   \* *Corresponding author: Dr. Aurélie Cébron, LIEC UMR7360, Faculté des Sciences et Technologies, Bd des Aiguillettes,*  
11   *BP70239, 54506 Vandoeuvre-les-Nancy; Mail: [aurelie.cebron@univ-lorraine.fr](mailto:aurelie.cebron@univ-lorraine.fr); Phone: +33 3 72 74 52 15.*

12  
13   **Running title:** microbial diversity in oil seep impacted soils

14  
15   **Key-words:** bacteria, fungi, petroleum seep, hydrocarbons, soil, indicator taxa,  
16   aerobic/anaerobic metabolisms

17  
18   **Originality-Significance Statement:** The significance and novelty of our work is the study of  
19   bacterial and fungal diversity and the identification of indicator taxa developing next to  
20   petroleum seeps and most likely involved in the hydrocarbon degradation. The originality is  
21   that these indicator taxa are well-known aerobic and anaerobic bacteria having specific  
22   functional traits making them well-adapted to hydrocarbon-contaminated environment as well  
23   as fluctuating redox conditions.

## 27 **Abstract**

28 Natural attenuation, involving microbial adaptation, helps mitigating the effect of oil  
29 contamination of surface soils. We hypothesized that in soils under fluctuating conditions and  
30 receiving oil from seeps, aerobic and anaerobic bacteria as well as fungi could coexist to  
31 efficiently degrade hydrocarbons and prevent the spread of pollution. Microbial community  
32 diversity was studied in soil longitudinal and depth gradients contaminated with petroleum  
33 seeps for at least a century. Hydrocarbon contamination was high just next to the petroleum  
34 seeps but this level drastically lowered from 2 meters distance and beyond. Fungal abundance  
35 and alpha-diversity indices were constant along the gradients. Bacterial abundance was constant  
36 but alpha-diversity indices were lower next to the oil seeps. Hydrocarbon contamination was  
37 the main driver of microbial community assemblage. 281 bacterial OTUs were identified as  
38 indicator taxa, tolerant to hydrocarbon, potentially involved in hydrocarbon-degradation or  
39 benefiting from the degradation by-products. These taxa belonging to lineages of aerobic and  
40 anaerobic bacteria, have specific functional traits indicating the development of a complex  
41 community adapted to the biodegradation of petroleum hydrocarbons and to fluctuating  
42 conditions. Fungi are less impacted by oil contamination but few taxa should contribute to the  
43 metabolic complementary within the microbial consortia forming an efficient barrier against  
44 petroleum dissemination.

45

46

## 47 **1. Introduction**

48

49 The expansion of the oil industry in the first half of the 20<sup>th</sup> century left behind sites where oil  
50 continues to seep to the surface. Contamination of surface soils by oil from leakage of old  
51 petroleum pumping wells or natural petroleum seeps is not necessarily a major problem because  
52 natural attenuation often limits the extent of the damage. During natural attenuation process,  
53 indigenous microorganisms degrade organic contaminants without external inputs that can  
54 promote bioremediation<sup>1</sup>. These natural phenomena occur in environments contaminated by  
55 many different compounds such as polychlorobiphenyls<sup>2</sup> or hydrocarbons<sup>3,4</sup>. Natural  
56 attenuation reflects the fact that biota can adapt over time to resist contamination and efficiently  
57 degrade it in various environments and conditions. Most of the reported studies on  
58 microbiology of sites contaminated with hydrocarbons are focused on the short-term effect of  
59 the crude oil after a major contamination event (oil spill, oil pipe damage)<sup>4,5,6,7</sup>, but very few  
60 have studied the long-term effect of chronic crude oil seeps on the composition of microbial

61 communities. Studies focusing on petroleum-contaminated soils in oil exploitation area have  
62 mostly tended to isolate microbes with hydrocarbon degradation properties for bioremediation  
63 application<sup>8</sup>. But there has been little in the way of attempts to analyse the composition of  
64 microbial communities in surface soils<sup>9</sup>, such as those contaminated for decades by oil seeps,  
65 to identify who are the best-adapted microbes and tend to found taxa as bioindicator of efficient  
66 petroleum degradation.

67 A wide diversity of microorganisms (fungi, bacteria and algae) are able to degrade and use  
68 petroleum (aliphatic and aromatic hydrocarbons) as sole carbon source using specific  
69 enzymatic machinery<sup>13,14</sup>. Although aerobic hydrocarbon biodegradation<sup>14,15</sup> is the most  
70 energy-efficient respiration process<sup>16</sup>, the contaminated subsurface environments are often  
71 naturally anaerobic<sup>17</sup>, especially within oil plumes where the available oxygen has been  
72 depleted. Anaerobic bacterial degradation of petroleum-related compounds such as benzene<sup>18</sup>,  
73 toluene<sup>19</sup>, polycyclic aromatic hydrocarbons<sup>20</sup> and alkanes<sup>21</sup> has been demonstrated in aquifers,  
74 sediment and soil environments. Local consumption of oxygen<sup>22</sup> or fluctuating conditions<sup>23</sup>  
75 could create aerobic and anaerobic microniches in soil, then a multitude of degradation  
76 processes could coexist at a small scale, involving aerobic and anaerobic microbial functional  
77 populations<sup>24</sup>. Fungi could also be key players in degrading recalcitrant organic contaminants  
78 because of their variety of extracellular enzymes<sup>25</sup>. The fungal mobilization and degradation of  
79 contaminants could contribute to the release of bioavailable intermediates that could be further  
80 degraded by bacteria<sup>26,27</sup>. Hence, catabolic interactions among different microbial groups  
81 during biodegradation are extremely important. In this context, we therefore hypothesized that  
82 in environments under fluctuating conditions and receiving constant input of petroleum, both  
83 bacteria and fungi and, among bacteria, both aerobic and anaerobic taxa could coexist and be  
84 involved in hydrocarbon degradation processes. It is therefore essential to identify the whole  
85 microbial community adapted to chronic soil petroleum contamination and to estimate how far  
86 hydrocarbon contamination and its impact on communities extend around oil seeps.

87 In this study, we present a detailed spatial analysis of the microbial community structure in two  
88 forest soils which have both been impacted by petroleum seeps for at least a century but with  
89 contrasting soil characteristics. For both sites, we studied the hydrocarbon contamination  
90 longitudinal gradients away from oil seeps. Bacterial and fungal community composition was  
91 studied through high throughput sequencing of 16S and 18S rDNA and indicator taxa  
92 developing mainly next to the oil seeps and potential hydrocarbon-degraders were identified  
93 thanks to an original analysis rarely used in microbial ecology (TITAN) but well-adapted to  
94 environmental gradients.

95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128

## 2. Material and Methods

### 2.1. Soil sampling

The first sampling campaign was carried out in May 2016, on two sites with contrasting soil characteristics located in the Haguenau forest (Alsace, France; **Figure 1**). Both sites have been impacted by oil seeps since the end of petroleum extraction which took place in the region between the end of the 18<sup>th</sup> century and the middle of the 20<sup>th</sup> century by Pechelbronn petroleum. The soil of the Gunstett site (GUN; 48°55'9.8"N, 7°48'6.7"E) presented a clay texture (42% clay, 40% silt, 18% sand) while that of Etangs-Verts (ETV; 48°52'21.8"N, 7°46'51.6"E) is sandy loam (15% clay, 23% silt, 62% sand). This difference was also highlighted by the different cationic exchange capacities (CEC) which were of 24.5 and 6.0 cmol.kg<sup>-1</sup> on average in GUN and ETV soils, respectively. Continuous oil seeps created ponds of 5 to 10 meters of diameter (**Figure 1**). On each site, 15 meters longitudinal transects from each oil seep were carried out and 5 sampling zones were chosen at 0, 2, 4, 10 and 15 meters from the oil seeps (named GUN-0, GUN-2, GUN-4, GUN-10, GUN-15 and ETV-0, ETV-2, ETV-4, ETV-10, ETV-15). Surface core soil samples (0 to 5 cm deep) were taken in triplicates (3 subsites 1 meter apart from each other) in the 5 sampling zones, representing 30 samples (2 sites x 5 sampling zones x 3 replicates; **Figure 1**).

The second sampling campaign focused only on the Etangs-Verts (ETV) site and was carried out in May 2017. Only the first zone (0 m, ETV-0), the closest to the oil seep, was sampled (**Figure 1**). Three independent soil cores of 30 cm depth were taken and were separated into 4 samples, to recover samples from 0-2 cm, 2-5 cm, 5-15 cm and 15-30 cm deep (named ETV0\_P0-2, ETV0\_P2-5, ETV0\_P5-15 and ETV0\_P15-30) representing 12 samples. The 0-2 cm depth sample was a solidified petroleum crust.

Soils (except ETV0\_P0-2) were sieved at 4 mm, aliquots of fresh soil sample were kept at 4°C until the return to the laboratory for mineralization activity measurements and estimation of the percentage of humidity. The rest was immediately frozen at -80°C for further analyses (DNA and organic extractions). Texture (NF X 31-107), total nitrogen and organic carbon content (NF ISO 13878 and NF ISO 10694) and pH (in water, NF ISO 10390) were measured at the LAS INRAE at Arras (France).

### 2.2. Mineralization activities under aerobic and anaerobic conditions

129 Two replicates of each soil samples (10 g fresh weight corresponding to 5.8 to 7.8 g dry weight)  
130 collected in May 2016 on longitudinal transects were incubated into 125 ml glass flasks sealed  
131 hermetically with Teflon septum. One replicate was incubated under aerobic condition (ambient  
132 atmosphere) and the other under anaerobic condition by flushing and replacing the atmosphere  
133 by N<sub>2</sub>/H<sub>2</sub> (95:5) gas mix. Empty control flasks without soil (similar atmosphere than for aerobic  
134 and anaerobic tests) were run in the same way to correct CO<sub>2</sub> measurements. Flasks were  
135 incubated at 24°C for 7 days. Carbon dioxide (CO<sub>2</sub>) release was measured after 1, 2, 3, 4 and 7  
136 days on a 2 ml portion of the flask atmosphere (sampled using a syringe) by infrared  
137 spectrophotometer (Binos 1004, absorption at 2325.6 cm<sup>-1</sup>)<sup>28</sup>. The mineralization activities  
138 were represented by the produced CO<sub>2</sub> expressed as carbon mass per gram of dry weight soil  
139 per day (µg C g<sup>-1</sup> d<sup>-1</sup>).

140

### 141 **2.3. Organic extraction and hydrocarbon measurement**

142 Solvent extractions were performed using an automated solvent extractor Dionex® ASE 350.  
143 Before extraction procedure, activated copper powder (1 g) and Na<sub>2</sub>SO<sub>4</sub> (1 g) were added to 2  
144 to 10 g of soil in the extraction cells to remove the molecular sulphur and the residual water,  
145 respectively. Extractions were performed twice at 130°C and 100 bars with HPLC grade  
146 dichloromethane (DCM) using a static time of 5 min. Organic extracts were diluted with DCM  
147 to reach 20 ml. An aliquot (3 ml) was dried in a fume hood (36 hours) to quantify the mass of  
148 the extractable organic matter (EOM)<sup>29</sup>.

149 The Hydrocarbon Index (HI) was measured according to the ISO 16703:2004 procedure using  
150 a GC-FID 7890A Agilent technologies<sup>30</sup>. The integration limits of the chromatogram area,  
151 corresponding to the *n*-decane (C<sub>10</sub>H<sub>22</sub>) and *n*-tetracontane (C<sub>40</sub>H<sub>82</sub>), were defined using a  
152 standard mixture (ASTM D5307). For external calibration an ISO 11046 standardized oil was  
153 used. Twelve solutions of different concentrations (from 1 to 31 mg ml<sup>-1</sup>) were injected into the  
154 GC-FID in order to obtain calibration curves. The HI was estimated from integration of the area  
155 under each chromatogram and corrected by the background noise (by integrating the area  
156 chromatogram obtained with a sample of pure dichloromethane that serves as a control).

157 Compound molecular distribution of soil EOM and crude oil (sampled directly in the middle of  
158 the oil seep) was determined through injection in a Shimadzu gas chromatograph (GC2010)  
159 equipped with a capillary column in silica glass DB-5MS (60 m × 0.25 mm i.d. × 0.1 µm film  
160 thickness) coupled with a mass spectrometer (QP2010) operating at 70 eV in fullscan mode.  
161 The inlet temperature was set at 300°C and injections were run in split mode (split ratio of 1:5).  
162 The oven temperature program was as followed: 70°C for 2 minutes, from 70 to 130°C at

163 15°C/min, then 130–315°C at 4 °C/min followed by an isothermal stage at 315°C for 25 min.  
164 The carrier gas was helium at 1.5 ml.min<sup>-1</sup> constant flow. Compound identification was  
165 confirmed by the mass spectra obtained with the MS detector in comparison with the NIST08  
166 and Wiley database spectra. Carbon preference index (CPI) was calculated according to  
167 Colombo *et al.*<sup>31</sup> with the *n*-alkanes areas obtained on the GC-MS extracted ion chromatogram  
168 (m/z 57) of EOM from the different sampling points, using the formula:  $CPI = 2 \times (n-C_{27} + n-$   
169  $C_{29}) / (n-C_{26} + 2 \times n-C_{28} + n-C_{30})$ . CPI is a ratio classically used for EOM source and contribution  
170 type<sup>31,32</sup>, a ratio close to one is indicator of crude oil signature and higher values (>5) indicates  
171 biogenic contribution.

172

#### 173 ***2.4. DNA extraction, bacterial and fungal abundance and quantification of known*** 174 ***hydrocarbon-degradation genes***

175 Genomic DNA was extracted from *ca.* 0.5 g of the soil samples using a Fast DNA spin kit for  
176 Soil (MP Biomedicals). Genomic DNA extracts were diluted to 5 ng µl<sup>-1</sup> to be used as templates  
177 for further PCR and qPCR analyses.

178 Quantification of bacterial and fungal communities was done using real-time PCR  
179 quantification using primer sets 968F/1401R and Fung5F/FF390R for 16S and 18S rRNA gene  
180 quantification, as described previously by Cébron *et al.*<sup>33</sup> and Thion *et al.*<sup>34</sup>. PAH-RHD $\alpha$  genes  
181 encoding alpha subunits of ring-hydroxylating dioxygenases from Gram negative (GN)  
182 Proteobacteria and Gram-positive (GP) Actinobacteria, *alkB* genes encoding alkane  
183 monooxygenases<sup>35</sup> and *bssA* genes encoding benzylsuccinate synthases were quantified using  
184 PAH-RHD $\alpha$  GN and GP primers, *alkB*-F/R primers and 7772F/8546R primers following  
185 Cébron *et al.*<sup>33</sup>, Gielnik *et al.*<sup>36</sup> and Winderl *et al.*, respectively. Real-time PCR assays were  
186 performed using iQ SybrGreen Supermix (Biorad) and CFX96 apparatus (BioRad). Data were  
187 expressed as gene copy number per gram of dry weight soil thanks to ten-time dilutions (from  
188 10<sup>8</sup> to 10<sup>2</sup> copies/µl) of standard plasmids.

189

#### 190 ***2.5. Bacterial 16S and fungal 18S rDNA sequencing***

191 The V3/V4 region of bacterial 16S rRNA genes (*ca.* 450 bp) was amplified using S-D-Bact-  
192 0341-a-S-17 and S-D-Bact-0787-b-A-20 primers following a previously described dual-index  
193 strategy<sup>37</sup> using PCR primers with an Illumina adaptor, pad and index sequences<sup>38</sup>. PCR  
194 reactions were performed on 2 µl of diluted gDNA using Phusion high-fidelity polymerase  
195 (Thermo Scientific). PCR reactions consisted of 31 cycles with touchdown annealing  
196 temperature for 18 cycles (63 to 54°C with a decrease of 0.5°C/cycle) and 13 cycles at 54°C.

197 Amplification products were checked on 1% agarose gel electrophoresis and purified using the  
198 UltraClean-htp 96Well PCR Clean-Up kit (Qiagen) following the manufacturer's instructions.  
199 After Quant-iT PicoGreen ds-DNA Assay Kit (Invitrogen) quantification, an amplicon library  
200 was prepared (equimolar pool at 10 nM), purified on QIAquick PCR purification Kit Column  
201 (Qiagen) and sent for sequencing to the Genewiz platform (South Plainfield, NJ, USA) which  
202 used an Illumina MiSeq V2 Kit for 2x250 bp paired-end sequencing.

203 The V7-V8 region of fungal 18S rRNA genes was amplified (*ca.* 390 bp) using FR1 and FF390  
204 primers<sup>39</sup> fused with a partial sequencing adapter. This first PCR was performed using Phusion  
205 high-fidelity polymerase (Thermo Scientific). The PCR reactions consisted of 25 to 30 cycles  
206 (depending on the samples) with an annealing temperature of 54°C. Amplification products  
207 were checked by 1% agarose gel electrophoresis before being sent to Microsynth AG (Balgach,  
208 Switzerland) for further steps namely second PCR, library preparation and Illumina MiSeq 2 x  
209 250 bp paired-end sequencing.

210 Illumina MiSeq paired-end reads for 16S and 18S rDNA were deposited in the SRA database  
211 under BioProjects ID: PRJNA685399 and PRJNA686256, respectively.

212

## 213 **2.6. Processing of bacterial 16S and fungal 18S rDNA sequencing data**

214 Sequence data were analysed following the MiSeq SOP procedure available in March 2017 and  
215 described in Kozich *et al.*<sup>37</sup>, using Mothur v.1.36.0<sup>40</sup>. Only 41 out of 42 samples were treated  
216 for 16S rDNA because sequencing failed for GUN-15 third replicate. Paired-end reads were  
217 trimmed to a minimum QScore of 20 (both 16S and 18S rRNA reads) and joined using the  
218 following criteria: no ambiguous bases and 404 bp < length < 430 bp or 348 bp < length < 359  
219 bp for 16S or 18S rDNA paired-end reads, respectively. After, alignment of unique sequence  
220 representatives to Silva V132 fasta data (released in Dec. 2017) and pre-clustering of sequences  
221 (sequences that are within 2 nt of each other), chimeras were detected and removed using  
222 Uchime and singletons (split.abund command removing sequences represented only once  
223 among all reads) were removed. Taxonomy was assigned using the Silva V132 tax data with a  
224 cut-off = 80 and using the Wang method. Sequences affiliated to archaea, chloroplasts,  
225 unknown, mitochondria and eukaryota were removed for 16S rDNA analysis, while sequences  
226 affiliated to archaea, chloroplasts, unknown, mitochondria and bacteria were removed for 18S  
227 rDNA analysis. After uncorrected pairwise distances calculation between aligned sequences,  
228 the bacterial and fungal sequences were clustered in Operational Taxonomic Units (OTUs) at  
229 97% similarity, and consensus taxonomy for each OTU was determined using classify.otu.



230 Finally, datasets were rarefied to the lowest number of sequences per sample (53,422 and  
231 14,260 reads/sample for bacterial and fungal sequences, respectively).

232 The number of reads for each OTU was expressed as a percentage to express the relative  
233 proportion or relative abundance of each OTU within the total community. Alpha diversity was  
234 expressed by calculating Chao1 richness, Pielou's evenness  $J'$  and Shannon diversity  $H'$  indices  
235 using Mothur.

236

### 237 **2.7. Bacterial functional trait inference**

238 Functional traits were compared between increasing indicator taxa found in ETV0 and GUN0  
239 samples (identified through TITAN2 analysis, see below) and the whole bacterial community  
240 inhabiting the control forest samples ETV15 and GUN15.

241 Inference of 19 bacterial traits (e.g. tolerance to pH, NaCl, temperature, preference to oxygen,  
242 Gram staining, width, length, shape, motility, pigment, spore production, trophic type and DNA  
243 GC %) was performed using the BactoTraits database<sup>41</sup>. The relative proportion of each trait  
244 attributes for all bacterial OTUs were calculated and attributed based on their taxonomic  
245 affiliation to get OTUs trait profiles. Based on the relative abundance of each OTU, the relative  
246 abundance of each trait attributes for the 19 traits was calculated and compared for the two  
247 different populations by Wilcoxon rank-sum test (p value <0.05).

248 The potential functional pathways of the two bacterial populations were profiled using Tax4Fun  
249 (v 1.0.1) (Aßhauer et al. 2015) run on Galaxy (X. SIGENAE [<http://www.sigena.org/>]).  
250 Differences in the level 1, level 2 and level 3 KEGG predicted functions of increasing indicator  
251 taxa and of control forest soil communities were compared and analyzed by Wilcoxon rank-  
252 sum test (p value <0.05).

253

254

### 255 **2.8. Statistical analysis**

256 Statistical analyses were performed using R studio version 3.6.0<sup>42</sup>. Values of soil parameters  
257 (extractable organic matter EOM, hydrocarbon index HI, total organic carbon C, nitrogen N  
258 and C/N ratio), mineralization activities, bacterial and fungal quantification (16S and 18S rRNA  
259 gene copy numbers), alpha-diversity estimators (Chao1, Shannon diversity and Pielou's  
260 evenness indices), relative abundance of bacteria phyla and fungal divisions and the relative  
261 abundance of the bacterial functional trait attributes were compared among the 5 different zones  
262 (0, 2, 4, 10 and 15 meters from the petroleum seeps) for both sites independently (GUN and  
263 ETV) and for different sampling depths at the ETV site. This was carried out through one-way

264 analyses of variance (Anova) with  $p < 0.05$ , followed by Tukey HSD *PostHoc* tests.  
265 Using the Vegan R package<sup>43</sup>, non-metric multidimensional scaling (NMDS) analyses were  
266 performed (*'metaMDS'* function) based on the Bray-Curtis dissimilarity matrix generated from  
267 bacterial and fungal relative abundances to estimate the dissimilarity in structure between all  
268 samples. For bacteria, OTUs having at least 0.1% relative abundance ( $\geq 54$  reads) were kept for  
269 this analysis. The environmental gradients of hydrocarbon content were built on NMDS  
270 ordination space using the *'Ordisurf'* function of the Vegan R package which used generalized  
271 additive modelling (GAM) to overlay environmental variables<sup>43,44</sup>. Permutational multivariate  
272 analyses of variance (PERMANOVA) were performed with 999 permutations using the  
273 *'adonis'* function in Vegan R package to determine if bacterial and fungal community  
274 composition differed significantly between sites (next to the oil seeps vs. the others).  
275 Threshold Indicator Taxa ANalysis (TITAN) was performed to identify environmental variable  
276 values (in our case hydrocarbon index values) maximizing taxa frequency and abundance using  
277 bootstrapping to identify reliable indicator taxa<sup>45</sup>. TITAN 2.4 package in R<sup>46</sup> was used.  
278 Association was measured by taxon abundances weighted by their occurrence in each sample  
279 as for Indicator taxa determination<sup>47</sup> and standardized as z-scores to facilitate cross-taxon  
280 comparison via permutation of samples along the predictor. The number of bootstraps and  
281 permutations performed in our TITAN2 analysis was 500 and 250, respectively. Analysis was  
282 performed at the OTU taxonomic level. Only OTUs observed more than 3 times across all sites  
283 were used and thus 3,306 and 861 OTUs were considered for bacterial and fungal analyses  
284 respectively. TITAN2 distinguishes declining and increasing Indicator Taxa along  
285 environmental gradient.

286

287

### 288 **3. Results and discussion**

289

#### 290 ***3.1. Soil characteristics, and hydrocarbon contamination levels***

291 Soil properties were characterised for the two sampling sites and both campaigns (**Table 1**). At  
292 the ETV site, nitrogen content was higher next to the oil seep and its content decreased with  
293 depth, while staying constant all over the transect at GUN site. Due to high concentration of  
294 petroleum carbon, the C/N ratio was high next to the oil seeps (c.a. 40) compared to values  
295 found at a 15 m distance (c.a.15). Finally, the pH values were also higher next to the oil seeps  
296 and decreased with distance. It was previously shown that oil contaminated soils harbored

297 alkaline pH (Wang et al. 2010; Sun et al. 2015). At the first zone next to the petroleum seep, at  
298 both sites, the average quantity of extractable organic matter (119.2 and 118.6 mg EOM g<sup>-1</sup> in  
299 the GUN and ETV sites, respectively), and the hydrocarbon index (HI, 73.7 and 35.3 mg g<sup>-1</sup> in  
300 the GUN and ETV sites, respectively) values were significantly higher than in the samples  
301 farthest from the seeps at a distance of 10 or 15 m which revealed high oil contamination of the  
302 soil at the vicinity of the seeps. The EOM and HI values did not decrease gradually with distance  
303 but dropped by more than 90% from the second sampling zone (a 2 m distance from the oil  
304 seeps). In opposite way, an increase in the carbon preference index (CPI) with increasing  
305 distance from the oil seep (**Table 1**), was shown. CPI is a ratio classically used for EOM source  
306 and contribution type<sup>31,32</sup>. The low ratio (around 1) as measured next to the oil seeps is indicator  
307 of crude oil signature and higher values (>5) indicates a major contribution of biogenic  
308 compounds to the soil organic matter. For both sites, samples closest to the oil seeps were  
309 dominated by *n*-alkanes similar to those of the crude oil (**Figure S1**); this strong crude oil  
310 contribution in the EOM fingerprint decreased as the sampling points get further from the oil  
311 seepage. The EOM molecular distribution became dominated by *n*-alkanols and fatty aldehydes  
312 with even number of carbon atoms, and *n*-alkanes with odd number of carbon atoms (**Figure**  
313 **S1**), characteristic of biogenic samples and higher plant contribution<sup>31,32</sup> commonly found in  
314 natural soils. At the farthest sampling points from the oil seeps (ETV15 and GUN15) petroleum  
315 signature completely disappeared (**Figure S1**). Samples from the second sampling campaign  
316 showed a similar gradient of hydrocarbon contamination with depth but at a smaller scale,  
317 highlighting higher contamination in the surface crust (108.9 and 42.3 mg g<sup>-1</sup> for EOM and HI,  
318 respectively) compared to the samples taken at a depth of 15 and 30 cm (**Table 1**). Same  
319 gradients were observed for the total organic carbon content with distance and depth (**Table 1**).

320

### 321 **3.2. Microbial activity and abundance**

322 To evaluate the functional potential of hydrocarbon degradation, we measured the microbial  
323 activity of organic matter mineralization under aerobic and anaerobic conditions. Both aerobic  
324 and anaerobic microbial mineralization activities were detected. Mineralization activity was 3  
325 to 7 times greater under aerobic than under anaerobic conditions for both sites (**Table 2**). Both  
326 aerobic and anaerobic mineralization activities were significantly higher next to the oil seeps at  
327 the first sampling zone compared to the rest of the transect, and decreased with depth at ETV  
328 site. Significant positive Pearson correlations ( $p < 10^{-5}$ ) were found between organic matter  
329 mineralization activities and HI values ( $r = 0.61$  and  $0.52$  for aerobic and anaerobic  
330 mineralization activities, respectively). As much of the organic matter found next to the oil seep

331 is composed of petroleum hydrocarbons we can assume that microbial activities were mainly  
332 dedicated to hydrocarbon degradation both in aerobic and anaerobic conditions.  
333 The distance to the petroleum seep was not found to significantly influence the number of  
334 bacteria and fungi (**Table 2**). We quantified  $5 \times 10^{10}$  to  $10 \times 10^{10}$  copies of 16S rRNA genes per  
335 gram of soil in both soil transects. Fungal 18S rRNA gene copy number was 100 times lower  
336 than for 16S rRNA one. The quantification of bacteria showed a significant decrease with soil  
337 depth from  $10 \times 10^{10}$  to  $2 \times 10^{10}$  16S rRNA gene copy numbers per gram of soil from the surface  
338 (0-2 cm) to the 15-30 cm depth samples.

339

### 340 **3.3. Quantification of known hydrocarbon-degradation genes**

341 The well-known bacterial genes encoding PAH dioxygenase (PAH-RHD $\alpha$ ), alkane  
342 monooxygenase (*alkB*) and benzylsuccinate synthase-like (*bssA*) were quantified to estimate  
343 the functional population potentially involved in hydrocarbon degradation in aerobic and  
344 anaerobic conditions (**Table S1**). No significant difference in PAH-RHD $\alpha$  and *alkB* gene  
345 abundance was shown across the transects even if these genes involved in aerobic hydrocarbon  
346 degradation tended to be in higher quantity next to the oil seep in samples from ETV site. At  
347 this site, the abundance of PAH-RHD $\alpha$  genes from both Gram-negative Proteobacteria and  
348 Gram-positive Actinobacteria was significantly higher in surface crust (0-2 cm) than in the  
349 deeper samples (5 to 30 cm). However, when expressed as percentage compared to 16S rRNA  
350 genes, PAH-RHD $\alpha$  genes represent less 0.02% while *alkB* genes could represent up to 1.0%  
351 next to the oil seep. Even if well-known Proteobacteria and Actinobacteria possess these  
352 functional genes<sup>13,33,35</sup>, it's more likely that most taxa involved in hydrocarbon degradation *in*  
353 *situ* would possess other suites of genes because they were detected in really low quantity. It  
354 was previously showed that for active taxa belonging to PAH-degrading consortia identified  
355 through stable isotope probing, the ratio of PAH-RHD $\alpha$  relative to 16S rRNA genes only  
356 accounted for 1.2 to 5.9%<sup>48</sup>. Similarly, genes involved in all steps of aromatic compounds  
357 metabolism represented less than 5% of the metagenome of an active PAH-degrading consortia  
358 where most of the taxa were not involved in the first hydroxylating step but in intermediate  
359 metabolite degradation<sup>49</sup>. Targeting PAH-RHD $\alpha$  and *alkB* genes involved in this first step is  
360 not sufficient to evaluate the abundance of bacteria involved in all steps of hydrocarbon  
361 mineralization. When targeting *bssA* genes involved in anaerobic hydrocarbon degradation, a  
362 significant higher abundance of these genes was detected in DNA extracted from sites close to  
363 the oil seeps in both ETV and GUN sites and in the surface petroleum crust in ETV site (Table

364 S2). However, these genes also represent a low percentage compared to 16S rRNA genes (less  
365 than 0.004%). Even if the identification of individual metabolic genes involved in oil  
366 degradation *in situ* sometime showed correlation between functional-population abundance and  
367 contamination level (Cébron et al. 2008<sup>33</sup>; Winderl et al. 2008), here it was not sufficient to  
368 conclude about functional pathways involved in oil degradation. Metagenomic analysis could  
369 be helpful in resolving this question but unknown genes and pathways (not annotated) as well  
370 as the metabolic complementarity of the taxa could prevent the analysis of such complex  
371 metagenomes, especially as there seem to be both anaerobic and aerobic processes taking place  
372 in our study sites.

373

### 374 ***3.4. Bacterial and fungal community : alpha-diversity indices and composition***

375 To help us assess the impact of hydrocarbons on the microbial community structure and  
376 composition, 16S rDNA and 18S rDNA sequencing were performed. After quality filtering, a  
377 total of 5,187,772 and 2,485,385 paired-end sequences were retrieved from the 16S rDNA and  
378 18S rDNA libraries, ranging from 53,422 to 122,937 reads per sample for 16S rDNA and from  
379 4,016 to 79,581 reads per sample for 18S rDNA. After sub-sampling to 53,422 and 14,260 reads  
380 per samples for 16S and 18S rDNA libraries, respectively, both 16S and 18S rDNA data were  
381 analyzed for 41 samples out of 42 (GUN-15 third replicate failed for 16S rDNA sequencing  
382 and ETV-15 second replicate was removed for 18S rDNA analysis because of a too low number  
383 of sequences). No general trend was found for fungal alpha-diversity indicators (**Table 2**),  
384 either according to the distance to the oil seep or in depth. However, on both sites, bacterial  
385 richness (Chao1), Shannon diversity and Evenness indices were lower next to the oil seeps  
386 compared to all other sampling distance and also in the surface crust compared to the other  
387 depth samples. Negative Pearson correlations were observed between EOM, HI and TOC  
388 values and the 3 alpha-diversity indices (Chao1, Shannon and Evenness) describing bacterial  
389 community composition (**Table 3**) for the 2 sites. Similarly, a lower bacterial taxonomic  
390 richness was observed in oil plume from Deepwater horizon spill (Hazen et al. 2010)<sup>6</sup>. This  
391 observation indicate that mostly hydrocarbon-degrading bacteria were enriched next to the oil  
392 seeps and that bacteria sensitive to petroleum contamination disappeared as shown previously  
393 following oil contamination due to pipeline rupture (Yang et al. 2016).

394 Overall no significant bacterial community composition difference (Wilcoxon test, p=0.543)  
395 was detected between GUN and ETV sites at the phylum level. The distance to the petroleum  
396 seep positively or negatively influenced many bacterial phyla (**Figure 2A**). The relative  
397 abundance of Actinobacteria, Chloroflexi, Firmicutes and Kiritimatiellaeota members was

398 significantly higher next to the oil seep and decreased with distance on GUN site ( $p \leq 0.003$ ).  
399 For the two first phyla the same tendency was observed on ETV site along the longitudinal  
400 gradient ( $p \leq 0.016$ ) and with depth with a significantly higher proportion ( $p < 0.0001$ ) of these  
401 phyla in the surface petroleum crust than at 5-30 cm depth. It was previously shown that  
402 Actinobacteria and Firmicutes dominated the bacterial community in soil after oil pipeline  
403 rupture (Yang et al. 2014) and that Chloroflexi and Firmicutes relative abundance correlated  
404 with diesel-contamination level (Sutton et al. 2013). *Chloroflexi* can degrade petroleum  
405 hydrocarbons by anaerobic respiration. The relative abundance of Verrucomicrobia, Alpha-  
406 Proteobacteria, Planctomycetes, Elusimicrobia and Fibrobacteres (the two latter classified in  
407 the group “others”) members (**Figure 2A**) increased significantly when moving away from the  
408 oil seep for both sites ( $p \leq 0.047$ ). Although no trend was found in the literature at the phylum  
409 level, some members of these phyla are most likely sensitive to oil contamination. For both  
410 GUN and ETV sites, the bacterial community structures at the OTU level were found to be  
411 changed according to the petroleum contamination of soil samples as confirmed using NMDS  
412 analysis where the hydrocarbon concentration gradient was overlapped (**Figure 3A**). On the  
413 first component, samples were separated according to the hydrocarbon content and the two sites  
414 were separated on the 2<sup>nd</sup> component. At OTU level, the bacterial community composition was  
415 significantly different among the different sampling areas (Permanova,  $p=0.001$ ). Previous  
416 studies have shown such structuration of the bacterial communities according to hydrocarbon  
417 contamination level (Paissé et al. 2008; Abed et al. 2015<sup>103</sup>).

418 Overall no significant fungal community composition difference (Wilcoxon test,  $p=0.861$ )  
419 was detected between GUN and ETV sites at the division level. Fewer changes were observed  
420 for the fungal community composition at the division and sub-division levels along longitudinal  
421 and depth gradients (**Figure 2B**). Although the relative abundance of members of the  
422 Taphrinomycotina sub-division (Ascomycota) was very low ( $< 0.3\%$ ), a significant increase  
423 was observed next to the oil seep ( $p=0.014$ ) and in the petroleum surface crust ( $p=0.042$ ).  
424 Similarly, the relative abundance of Ustilaginomycotina, Pucciniomycotina (both  
425 Basidiomycota) and Pezizomycotina (Ascomycota) members was significantly higher in the  
426 surface petroleum crust and decreased with depth ( $p \leq 0.032$ ). Pezizomycotina was found to be  
427 dominant (Cury et al. 2015) and were isolated (Ferrari et al. 2011) in diesel-contaminated soil,  
428 maybe because of their hydrocarbon-degradation pathways (Moktali et al. 2012). Modification  
429 of the fungal community structure at the OUT level was less pronounced than for bacteria, but  
430 a similar gradient was observed according to the hydrocarbon content (**Figure 3B**). The lower  
431 impact of on fungal community can suggest that hydrocarbons are less toxic to fungi, that fungi

432 can cope with hydrocarbon contamination or that they can benefit from hydrocarbon  
433 compounds for their growth being well adapted to the degradation of complex carbon  
434 substrates. Fungi with their multiple enzymatic activities (hydrolytic enzymes or oxidases such  
435 as peroxidases or phenol oxidases) are able to degrade organic pollutants such as aliphatic  
436 hydrocarbons and PAHs<sup>86</sup>.

437

### 438 **3.5. Bacterial and fungal indicator taxa**

439 As bacterial and fungal community composition are shaped by petroleum contamination,  
440 indicator taxa (IT) relative to the hydrocarbon level were sought to further understand which  
441 microorganisms are tolerant/sensitive to oil contamination, potentially involved in hydrocarbon  
442 degradation or benefiting of degradation metabolites. Data from both sampling sites were  
443 combined and indicator taxa were sorted using TITAN2, a well-adapted tool for gradient  
444 studies. Bacterial and fungal indicator taxa increasing (**Figure 4**) or decreasing (**Figure S2**)  
445 along the hydrocarbon contamination gradient (HI values) were identified. As both bacterial  
446 and fungal indicator taxa were identified, we confirmed our hypothesis that different groups of  
447 bacteria and fungi with complementary metabolic properties may coexist and potentially  
448 cooperate at the vicinity of the oil seeps for oil degradation<sup>93</sup>. The presence of all these  
449 microorganisms with varied metabolisms should form an efficient barrier against dissemination  
450 of the petroleum contamination under fluctuating environmental conditions.

451 TITAN2 sorted 281 and 664 bacterial OTUs as indicator taxa which were found to increase and  
452 decrease respectively with HI values (the complete list is available in **Table S1**). The relative  
453 abundance of these indicator taxa reached 30-40% when getting closer to or further from the  
454 oil seeps for increasing and decreasing taxa respectively. Among the bacterial IT increasing  
455 next to the oil seeps, we found 110 OTUs belonging to Proteobacteria, 51 to Acidobacteria, 39  
456 to Actinobacteria, 16 to Chloroflexi, 16 to Firmicutes and 14 to Bacteroidetes (**Figure 4A**).  
457 Various IT belonging to lineages of anaerobic or facultative anaerobic bacteria, had their  
458 relative abundance increasing next to the oil seeps. Members of the Acidithiobacillaceae family  
459 were identified. In this family, bacterial strains belonging to the *Acidithiobacillus* genera are  
460 chemolithoautotrophs that can oxidize elemental sulphur or oxidize minerals containing ferrous  
461 iron<sup>52</sup>. Among the Rhodocyclaceae family, OTUs belonging to *Dechloromonas* and  
462 *Sulfuritalea* genera were identified as increasing IT. These genera were previously detected in  
463 hydrocarbon-polluted sites<sup>54</sup> and could degrade a wide variety of aromatic compounds under  
464 anaerobic nitrate-reducing conditions<sup>55-57</sup> through potentially original pathways<sup>58</sup>. Moreover,  
465 *Sulfuritalea* are also well recognized as chemolithoautotrophic sulfur-oxidizing bacteria<sup>59</sup>.

466 Some OTUs affiliated to the Chloroflexi and belonging to the *Anaerolineae* were also identified  
467 as increasing IT. This lineage is composed of bacteria known as strictly anaerobic fermentative  
468 and sulphate-reducing bacteria, previously found in hydrocarbon-rich sludge samples<sup>60</sup> and in  
469 anaerobic alkane-degrading culture<sup>61</sup>. Other IT mostly known as aerobic or facultative aerobic  
470 taxa had their relative abundance increasing next to the oil seeps. OTUs belonging to Gamma-  
471 Proteobacteria were identified, they are affiliated to the Porticoccaceae family and to  
472 *Immundisolibacter*, *Rhodanobacter* and *Pseudoxanthomonas* genera. Members of  
473 Porticoccaceae family have been described as aromatic-hydrocarbon degraders (Gutierrez et al.  
474 2012) and alkane-degraders<sup>50</sup>, but they also possess in their genomes a high number of  
475 secondary alkane-degrading genes<sup>51</sup>. *Immundisolibacter* strain have been isolated from aerobic  
476 bioreactor for its capacity to degrade high-molecular-weight PAHs<sup>62</sup>. Even if *Rhodanobacter*  
477 was previously identified as a member of a PAH-degrading consortium<sup>64</sup> it would more likely  
478 degrade intermediate metabolites<sup>65</sup> because no PAHs degradation gene have been detected in  
479 sequenced genomes<sup>66</sup>. The *Pseudoxanthomonas* species have been found to degrade different  
480 PAHs<sup>67,68</sup> and BTEX compounds<sup>69</sup> and some strains could produce a rhamolipid surfactant that  
481 enhances the hydrocarbon degradation<sup>70</sup>. Two types of Alpha-Proteobacteria were also  
482 identified as aerobe increasing IT. The Methyloligellaceae members that are well-known  
483 methylotrophs<sup>71</sup>, and the *Parvibaculum* genus possessing genes involved in alkane utilisation<sup>72</sup>  
484 that was detected *in situ*<sup>73</sup> and in aerobic enrichment culture degrading crude oil<sup>74</sup>. Well known  
485 aerobic alkanes and PAHs degraders, such as the Actinobacteria belonging to *Williamsia*<sup>75,76</sup>,  
486 *Cellulomonas*<sup>77,78</sup>, *Agromyces*<sup>79</sup> and *Mycobacterium*<sup>80</sup> genera, were almost undetected in the  
487 forest soil but became abundant next to the oil seeps and identified as increasing IT. These  
488 Actinobacteria are equipped with hydroxylase enzymes such as monooxygenases and  
489 dioxygenases, to initiate the aerobic oxidation of hydrocarbon molecules<sup>13,81</sup>. In addition, some  
490 of these genera can also produce exopolysaccharides (EPS) which serve as bioemulsifiers such  
491 as glycolipides or curdlan-like EPS produced by *Cellulomonas spp.*<sup>78,82</sup>. Finally, two types of  
492 Acidobacteria appeared as increasing IT, the members of Solibacterales order with  
493 undemonstrated hydrocarbon-degradation capacities but that were previously found in oil  
494 contaminated soil and sediments<sup>83,84</sup> and the members of the Acidobacteriales order that were  
495 identified as aerobic benzene-degraders using stable isotope probing technique<sup>85</sup>.

496

497 As shown before, the fungal community structure was less impacted by petroleum  
498 contamination and thus a low number of OTUs indicator taxa was found. TITAN2 sorted 61  
499 and 40 fungal OTUs as indicator taxa which were found to increase (**Figure 4B**) and decrease



500 (Figure S2) respectively with hydrocarbon index values (complete list is available in Table  
501 S2). The relative abundance of the indicator taxa which increased with the gradient (Figure  
502 4B) was highly variable from one sample to another (1 to 42%). However, two fungal OTUs  
503 have higher relative abundance close to the oil seep, one belonging to *Lactarius* genus  
504 (Basidiomycota, *Agaricomycetina*) and another belonging to Cryptomycota LKM11. *Lactarius*  
505 genera are well-known forest soil symbiotroph ectomycorrhizal fungi but can also be facultative  
506 saprotroph<sup>87</sup>. *Lactarius* species increased up to 33.8% close to the oil seep indicating an  
507 essential role in the biodegradation of hydrocarbons. Indeed, the role of *Agaricomycetes* in  
508 bioremediation of crude oil-contaminated soil was demonstrated<sup>88</sup> and the *in vitro* degradation  
509 of various hydrocarbons was reported for many Basidiomycetes<sup>89</sup>. Cryptomycota LKM11  
510 represented 1.8 to 12.7% at the sampling sites close to the oil seeps. This fungal sub-division  
511 was recently found in PAH contaminated soils<sup>90</sup>, anoxic and hydrocarbon-enriched oil sand  
512 tailings ponds<sup>91</sup> or creosote polluted site<sup>92</sup>.

513  
514

### 515 **3.6. Functional traits of the increasing bacterial indicator taxa**

516 Functional traits were predicted from BactoTraits<sup>41</sup> (morphological and physiological traits and  
517 affiliation to functional groups) and Tax4Fun (Abhauer et al. 2015) (KEGG functional  
518 pathways) tools to better characterize the bacterial indicator taxa increasing and dominant next  
519 to the oil seeps (Figure 5). We hypothesized that increasing bacterial IT possess specific  
520 functional traits making them tolerant to hydrocarbon contamination and giving them a  
521 functional advantage to survive and grow next the oil seeps.

522 Among the bacterial IT increasing next to the oil seeps, a higher proportion of Gram positive  
523 and non-motile bacteria were found (Figure 5A) and confirmed by predicted functions showing  
524 that functional pathways related to cell motility (bacterial chemotaxis and flagellar assembly)  
525 were detected in lower proportion (Figure 5B). Bacteria characterized as “competitors” and  
526 “colonizers” were more abundant while the “stress-sensitives” and “mesophiles” were in lower  
527 proportion among the bacterial IT next to the oil seeps than in the control forest soil bacterial  
528 community (Figure 5A). We previously described that competitors possessed traits improving  
529 the access to resources even when limited such as in oligotrophic environments and that Gram-  
530 positive and non-motile bacterial traits characterized this functional group. In contrast,  
531 colonizers were bacteria with high phenotypic plasticity with the ability to grow in anaerobic  
532 or facultative anaerobic conditions (Cébron et al. 2021). The higher proportion of bacteria  
533 belonging to these two functional groups among the IT could confirm the presence of two

534 complementary populations for efficient hydrocarbon degradation next to the oil seeps.  
535 There was also a higher proportion of known lithotrophs and of anaerobic or facultative  
536 anaerobic bacteria close to the oil seeps (**Figure 5A**). Chemolithotrophic bacteria close to the  
537 oil seeps could provide Fe(III) as electron acceptor to hydrocarbon-degraders<sup>53</sup>. Moreover,  
538 among bacterial IT, both aerobic and anaerobic/facultative anaerobic taxa were identified  
539 representing 62.5 and 33.1% respectively, while control bacterial community was mainly  
540 composed of aerobic bacteria. This finding agrees with higher rates of both aerobic and  
541 anaerobic hydrocarbon mineralization that were measured next to the oil seep (**Table 2**). These  
542 findings could indicate a zone of mixing between aerobic bacteria from surface soil and  
543 anaerobic bacteria brought to the surface soil with the oil seep as anaerobic hydrocarbon  
544 degraders prevail in sub-surface oil samples and contaminated petroleum groundwater<sup>96,97</sup>. The  
545 co-occurrence of different metabolisms has the advantage of reducing the concentrations of  
546 both toxic contaminants and of degradation metabolites<sup>98</sup>. Aerobic and anaerobic populations  
547 could alternatively be active depending on fluctuating redox conditions in the soil or could  
548 cooperate for hydrocarbon degradation such as in syntrophic metabolisms<sup>24</sup> by substrates  
549 sharing<sup>97,100</sup>. Joint anaerobic and aerobic metabolisms were rarely suggested<sup>101</sup>, but recent  
550 evidence of the co-transcription of genes involved in both aerobic and anaerobic benzene  
551 degradation in continuous culture<sup>102</sup> would explain our findings as well as previous ones of co-  
552 existence of aerobic and anaerobic bacteria in oil contaminated desert soils<sup>103</sup>.

553 At level 1 of KEGG functional categories, functions related to “Metabolism” and “Genetic  
554 information processing GIP” were detected in higher proportion while “Environmental  
555 information processing” and “Cellular processes” were detected in lower proportion for  
556 bacterial IT next to the oil seeps than for the control communities (**Figure 5B**). Within the  
557 Metabolism category, “carbohydrate metabolism” and “nucleotide metabolism” were over-  
558 represented, with higher proportion of functions involved in glycolysis/gluconeogenesis,  
559 pentose phosphate pathway, and galactose, starch, sucrose, propanoate, C5-branched dibasic  
560 acid, and purine metabolisms. Some functions involved in the metabolism of lipids, cofactors  
561 and vitamins and terpenoids and polyketides were also detected in higher proportion. Similarly,  
562 Zhang et al. (2019) showed that in sediment supplemented with pyrene, genes involved into  
563 almost every metabolic pathways showed a dramatic increase. Our finding indicates the  
564 possible availability of simple carbon sources besides hydrocarbons because these functions  
565 are dedicated to optimize the use of available resources and underline the large genetic  
566 investment of bacteria in the processing of environmental information important in these oil  
567 polluted environments (Mukherjee et al. 2017).

568 Moreover, within the “xenobiotic biodegradation and metabolism” category, functions  
569 involved in polycyclic aromatic hydrocarbon degradation as well as drug metabolism were  
570 significantly more present in bacterial IT increasing next to the oil seeps than in the control  
571 communities (**Figure 5B**). All these traits indicate greater and different metabolic capacities of  
572 the bacterial IT identified, as well adapted to hydrocarbon contamination. Finally, within the  
573 GIP category, functions related to “Folding, sorting and degradation”, “Transcription” and  
574 “Translation” were also over-represented. These regulatory functions may be related to the  
575 adaptation of microorganism to this highly hydrocarbon-contaminated environment. Similarly,  
576 through genome analysis of the oil-degrading *Pseudomonas aeruginosa* N002 strain, authors  
577 found among other gene categories, a very high abundance of genes for transcription (Das et  
578 al. 2015).

579

## 580 **Conclusions**

581 At this study site, where petroleum seeps have been contaminating the soil for at least a century,  
582 we were able to demonstrate the natural attenuation phenomenon. Indeed, the presence of  
583 microbial communities, that are well adapted to oil and its degradation, limits the extent of the  
584 contamination to a few meters around the seeps. We identified bacterial taxa that are well  
585 known for their capacity to biodegrade hydrocarbons, which it would be interesting to isolate  
586 to study their metabolism and use them for bioremediation purposes. Amongst the indicator  
587 taxa, which are highly favoured in the vicinity of oil seeps, we have shown the coexistence and  
588 possible cooperation of fungi as well as aerobic and anaerobic bacteria having specific  
589 functional traits that make them well adapted to the hydrocarbon contamination and to  
590 fluctuating redox conditions.

591

592

## 593 **Acknowledgements**

594 We would like to thank the CONPET project consortium for their involvement during sampling  
595 dates, C. Lorgeoux and D. Faure-Cattelain (GeoRessources) for hydrocarbon analyses and M.  
596 E. Baker and R. S. King for their help in using the TITAN2 package they developed. We also  
597 thank R. Dickinson (Inist-CNRS translation unit) for his careful proofreading of the manuscript.  
598 Funding: This work was supported by the French national programme EC2CO (CONPET  
599 Project) and is included in the scientific program of the GISFI research consortium dedicated  
600 to the knowledge and the development on remediation technologies for degraded and polluted  
601 lands (Groupement d'Intérêt Scientifique sur les Friches Industrielles - [18](http://www.gisfi.univ-</a></p></div><div data-bbox=)

602 lorraine.fr).is part of the GISFI research project (scientific interest group on brownfields).

603

#### 604 **Author contributions statement**

605 AC: Conceptualization; Data acquisition; Data curation; Formal analysis; Statistical analysis;  
606 Investigation; Methodology; Supervision; Validation; Writing original draft; Writing – review  
607 & editing.

608 AB: Formal analysis; Writing – review & editing.

609 TB: Conceptualization; Formal analysis; Methodology; Writing – review & editing.

610 CB: Conceptualization; Formal analysis; Methodology; Writing – review & editing

611 PF: Conceptualization; Methodology; Data acquisition; Funding acquisition; Project  
612 administration; Writing – review & editing.

613

614

#### 615 **Competing interests**

616 The authors declare no competing interests.

617

618

#### 619 **References**

620

- 621 1. Sarkar, D., Ferguson, M., Datta, R., Birnbaum, S. Bioremediation of petroleum hydrocarbons in contaminated  
622 soils: comparison of biosolids addition, carbon supplementation, and monitored natural attenuation.  
623 *Environmental pollution*, 136(1), 187-195 (2005).
- 624 2. Vergani, L., *et al.* Bacteria associated to plants naturally selected in a historical PCB polluted soil show  
625 potential to sustain natural attenuation. *Frontiers in microbiology*, 8, 1385 (2017).
- 626 3. Kachienga, L., Jitendra, K., Momba, M. Metagenomic profiling for assessing microbial diversity and  
627 microbial adaptation to degradation of hydrocarbons in two South African petroleum-contaminated water  
628 aquifers. *Scientific Reports*, 8(1), 1-6. (2018).
- 629 4. Ferguson, D. K., *et al.* Natural attenuation of spilled crude oil by cold-adapted soil bacterial communities at  
630 a decommissioned High Arctic oil well site. *Science of The Total Environment*, 722, 137258 (2020).
- 631 5. Kleikemper, J., *et al.* Activity and diversity of sulfate-reducing bacteria in a petroleum hydrocarbon-  
632 contaminated aquifer. *Applied and Environmental Microbiology*, 68(4), 1516-1523 (2002).
- 633 6. Hazen, T. C., *et al.* Deep-sea oil plume enriches indigenous oil-degrading bacteria. *Science*, 330(6001), 204-  
634 208 (2010).
- 635 7. Kostka, J. E., *et al.* Hydrocarbon-degrading bacteria and the bacterial community response in Gulf of Mexico  
636 beach sands impacted by the Deepwater Horizon oil spill. *Applied and environmental microbiology*, 77(22),  
637 7962-7974 (2011).

- 638 8. Spini, G., et al. Molecular and microbiological insights on the enrichment procedures for the isolation of  
639 petroleum degrading bacteria and fungi. *Frontiers in microbiology*, 9, 2543 (2018).
- 640 9. Higashioka, Y., Kojima, H., Sato, S., Fukui, M. Microbial community analysis at crude oil-contaminated soils  
641 targeting the 16S ribosomal RNA, *xyfM*, *C23O*, and *bcr* genes. *Journal of applied microbiology*, 107(1), 126-  
642 135 (2009).
- 643 10. Moubasher, H. A., et al. Phytoremediation of soils polluted with crude petroleum oil using *Bassia scoparia*  
644 and its associated rhizosphere microorganisms. *Int. Biodeter. Biodegr.* 98, 113–120 (2015).
- 645 11. Das, N., & Chandran, P. Microbial degradation of petroleum hydrocarbon contaminants: an overview.  
646 *Biotechnol. Res. Int.* 2011:941810 (2011).
- 647 12. Haritash, A. K., & Kaushik, C. P. Biodegradation aspects of polycyclic aromatic hydrocarbons (PAHs): a  
648 review. *J. Hazard Mater.* 169, 1–15 (2009).
- 649 13. Habe, H., & Omori, T. Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria.  
650 *Bioscience, biotechnology, and biochemistry*, 67(2), 225-243 (2003).
- 651 14. Abbasian, F., Lockington, R., Mallavarapu, M., Naidu, R. A comprehensive review of aliphatic hydrocarbon  
652 biodegradation by bacteria. *Applied biochemistry and biotechnology*, 176(3), 670-699 (2015).
- 653 15. Liu, Q., et al. Aerobic degradation of crude oil by microorganisms in soils from four geographic regions of  
654 China. *Scientific reports*, 7(1), 1-12 (2017).
- 655 16. Ghosal, D., Ghosh, S., Dutta, T. K., Ahn, Y. Current state of knowledge in microbial degradation of polycyclic  
656 aromatic hydrocarbons (PAHs): a review. *Frontiers in microbiology*, 7, 1369 (2016).
- 657 17. Aitken, C. M., Jones, D. M., Larter, S. R. Anaerobic hydrocarbon biodegradation in deep subsurface oil  
658 reservoirs. *Nature*, 431(7006), 291-294 (2004).
- 659 18. Burland, S. M., & Edwards, E. A. Anaerobic benzene biodegradation linked to nitrate reduction. *Applied and*  
660 *environmental microbiology*, 65(2), 529-533 (1999).
- 661 19. Ahad, J. M., et al. Carbon isotope fractionation during anaerobic biodegradation of toluene: implications for  
662 intrinsic bioremediation. *Environmental science & technology*, 34(5), 892-896 (2000).
- 663 20. Chang, B. V., Shiung, L. C., Yuan, S. Y. Anaerobic biodegradation of polycyclic aromatic hydrocarbon in  
664 soil. *Chemosphere*, 48(7), 717-724 (2002).
- 665 21. Coates, J. D., et al. Anaerobic degradation of polycyclic aromatic hydrocarbons and alkanes in petroleum-  
666 contaminated marine harbor sediments. *Applied and Environmental Microbiology*, 63(9), 3589-3593 (1997).
- 667 22. Cozzarelli, I. M., Herman, J. S., Baedeker, M. J. Fate of microbial metabolites of hydrocarbons in a coastal  
668 plain aquifer: the role of electron acceptors. *Environmental science & technology*, 29(2), 458-469 (1995).
- 669 23. Cravo-Laureau, C., et al. Role of environmental fluctuations and microbial diversity in degradation of  
670 hydrocarbons in contaminated sludge. *Research in Microbiology*, 162(9), 888-895 (2011).
- 671 24. Gieg, L. M., Fowler, S. J., Berdugo-Clavijo, C. Syntrophic biodegradation of hydrocarbon contaminants.  
672 *Current opinion in biotechnology*, 27, 21-29 (2014).
- 673 25. Fernandez-Luqueno, F., et al. Microbial communities to mitigate contamination of PAHs in soil-possibilities  
674 and challenge: a review. *Environ. Sci. Pollut. R.* 18, 12–30 (2010).
- 675 26. Cerniglia, C. E., & Sutherland, J. B. (2006). Relative roles of bacteria and fungi in polycyclic aromatic  
676 hydrocarbon biodegradation and bioremediation. *Fungi Biogeochem. Cycles*, 24(182).

- 677 27. Jacques, R. J., *et al.* Microbial consortium bioaugmentation of a polycyclic aromatic hydrocarbons  
678 contaminated soil. *Bioresource Technology*, 99(7), 2637-2643 (2008).
- 679 28. Cébron, A., *et al.* Experimental increase in availability of a PAH complex organic contamination from an aged  
680 contaminated soil: consequences on biodegradation. *Environmental Pollution*. 177:98-105 (2013).
- 681 29. Biache, C., *et al.* Effects of thermal desorption on the composition of two coking plant soils: Impact on solvent  
682 extractable organic compounds and metal bioavailability. *Environmental Pollution* 156, 671–677 (2008).
- 683 30. Usman, M., Hanna, K., Faure, P. Remediation of oil-contaminated harbor sediments by chemical oxidation.  
684 *Science of Total Environment*, 634, 1100-1107 (2018).
- 685 31. Colombo, J.C., Pelletier, E., Brochu, C., Khalil, M. Determination of hydrocarbon sources using n-alkane and  
686 polyaromatic hydrocarbon distribution indexes. Case study: Rio de La Plata estuary, Argentina. *Environ. Sci.*  
687 *Technol.* 23, 888–894 (1989).
- 688 32. Bray, E.E., & Evans, E.D. (1961). Distribution of n-paraffins as a clue to recognition of source beds. *Geochim.*  
689 *Cosmochim. Acta* 22, 2–15.
- 690 33. Cébron, A., Norini, M.P., Beguiristain, T., Leyval, C. Real-Time PCR quantification of PAH-ring  
691 hydroxylating dioxygenase (PAH-RHD $\alpha$ ) genes from Gram positive and Gram negative bacteria in soil  
692 sediment samples. *Journal of Microbiological Methods* 73, 148-159 (2008).
- 693 34. Thion, C., Cébron, A., Beguiristain, T., Leyval, C. Long-term in situ dynamics of the fungal communities in  
694 a multi-contaminated soil are mainly driven by plants. *FEMS microbiology ecology*, 82(1), 169-181 (2012).
- 695 35. Powell, S. M., Ferguson, S. H., Bowman, J. P., Snape, I. Using real-time PCR to assess changes in the  
696 hydrocarbon-degrading microbial community in Antarctic soil during bioremediation. *Microbial ecology*,  
697 52(3), 523-532 (2006).
- 698 36. Gielnik, A., *et al.* Functional potential of sewage sludge digestate microbes to degrade aliphatic hydrocarbons  
699 during bioremediation of a petroleum hydrocarbons contaminated soil. *Journal of Environmental*  
700 *Management*, 280, 111648 (2021).
- 701 37. Kozich, J. J., *et al.* Development of a dual-index sequencing strategy and curation pipeline for analyzing  
702 amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and environmental*  
703 *microbiology*, 79(17), 5112-5120 (2013).
- 704 38. Thomas, F., & Cébron, A. Short-term rhizosphere effect on available carbon sources, phenanthrene  
705 degradation, and active microbiome in an aged-contaminated industrial soil. *Frontiers in microbiology*, 7, 92  
706 (2016).
- 707 39. Vainio, E. J., & Hantula, J. Direct analysis of wood-inhabiting fungi using denaturing gradient gel  
708 electrophoresis of amplified ribosomal DNA. *Mycological research*, 104(8), 927-936 (2000).
- 709 40. Schloss, P. D., *et al.* Introducing mothur: open-source, platform-independent, community-supported software  
710 for describing and comparing microbial communities. *Applied and environmental microbiology*, 75(23),  
711 7537-7541 (2009).
- 712 41. Cébron, A., *et al.* BactoTraits—A functional trait database to evaluate how natural and man-induced changes  
713 influence the assembly of bacterial communities. *Ecological Indicators*, 130, 108047 (2021).
- 714 42. R Core Team R: A language and environment for statistical computing. R Foundation for Statistical  
715 Computing, Vienna, Austria (2019).
- 716 43. Oksanen, J., *et al.* Vegan: Community Ecology Package. R package version 2.5-6. (2019).

- 717 44. Marra, G., & Wood, S. N. (2011). Practical variable selection for generalized additive models. *Computational*  
718 *Statistics & Data Analysis*, 55(7), 2372-2387.
- 719 45. Baker, M. E., & King, R. S. (2010). A new method for detecting and interpreting biodiversity and ecological  
720 community thresholds. *Methods in Ecology and Evolution*, 1(1), 25-37.
- 721 46. Baker, M. E., King, R. S., Kahle, D., Kahle, M. D. Package 'TITAN2'. (2015).
- 722 47. Dufrière, M., & Legendre, P. (1997). Species assemblages and indicator species: the need for a flexible  
723 asymmetrical approach. *Ecological monographs*, 67(3), 345-366.
- 724 48. Cébron, A., et al. Root exudates modify bacterial diversity of phenanthrene degraders in PAH-polluted soil  
725 but not phenanthrene degradation rates. *Environmental Microbiology*, 13(3), 722-736 (2011).
- 726 49. Thomas, F., Corre, E., Cébron, A. Stable isotope probing and metagenomics highlight the effect of plants on  
727 uncultured phenanthrene-degrading bacterial consortium in polluted soil. *The ISME journal*, 13(7), 1814-1830  
728 (2019).
- 729 50. Knapik, K., Bagi, A., Krolicka, A., Baussant, T. Discovery of functional gene markers of bacteria for  
730 monitoring hydrocarbon pollution in the marine environment - a metatranscriptomics approach. *bioRxiv*,  
731 857391 (2019).
- 732 51. Ribicic, D., et al. Microbial community and metagenome dynamics during biodegradation of dispersed oil  
733 reveals potential key-players in cold Norwegian seawater. *Marine pollution bulletin*, 129(1), 370-378 (2018).
- 734 52. Kelly, D. P., & Wood, A. P. The family Acidithiobacillaceae. *The prokaryotes. Gammaproteobacteria*, 4, 15-  
735 25 (2014).
- 736 53. Hamamura, N., Olson, S. H., Ward, D. M., Inskeep, W. P. Diversity and functional analysis of bacterial  
737 communities associated with natural hydrocarbon seeps in acidic soils at Rainbow Springs, Yellowstone  
738 National Park. *Applied and environmental microbiology*, 71(10), 5943-5950 (2005).
- 739 54. Yagi, J. M., et al. Subsurface ecosystem resilience: long-term attenuation of subsurface contaminants supports  
740 a dynamic microbial community. *The ISME journal*, 4(1), 131-143 (2010).
- 741 55. Chakraborty, R., O'Connor, S. M., Chan, E., Coates, J. D. Anaerobic degradation of benzene, toluene,  
742 ethylbenzene, and xylene compounds by *Dechloromonas* strain RCB. *Applied and Environmental*  
743 *Microbiology*, 71(12), 8649-8655 (2005).
- 744 56. Keller, A. H., Kleinstuber, S., Vogt, C. Anaerobic benzene mineralization by nitrate-reducing and sulfate-  
745 reducing microbial consortia enriched from the same site: comparison of community composition and  
746 degradation characteristics. *Microbial ecology*, 75(4), 941-953 (2018).
- 747 57. Sperfeld, M., Diekert, G., Studenik, S. Anaerobic aromatic compound degradation in *Sulfuritalea*  
748 *hydrogenivorans* sk43H. *FEMS microbiology ecology*, 95(1), fyy199 (2019).
- 749 58. Salinero, K. K., et al. Metabolic analysis of the soil microbe *Dechloromonas aromatica* str. RCB: indications  
750 of a surprisingly complex life-style and cryptic anaerobic pathways for aromatic degradation. *BMC genomics*,  
751 10(1), 351. (2009).
- 752 59. Watanabe, T., Kojima, H., Fukui, M. Complete genomes of freshwater sulfur oxidizers *Sulfuricella*  
753 *denitrificans* skB26 and *Sulfuritalea hydrogenivorans* sk43H: genetic insights into the sulfur oxidation  
754 pathway of betaproteobacteria. *Systematic and applied microbiology*, 37(6), 387-395 (2014).

- 755 60. Roy, A., *et al.* Petroleum hydrocarbon rich oil refinery sludge of North-East India harbours anaerobic,  
756 fermentative, sulfate-reducing, syntrophic and methanogenic microbial populations. *BMC microbiology*,  
757 18(1), 151 (2018).
- 758 61. Tan, B., Dong, X., Sensen, C. W., Foght, J. Metagenomic analysis of an anaerobic alkane-degrading microbial  
759 culture: potential hydrocarbon-activating pathways and inferred roles of community members. *Genome*,  
760 56(10), 599-611 (2013).
- 761 62. Corteselli, E. M., Aitken, M. D., Singleton, D. R. Description of *Immundisolibacter cernigliae* gen. nov., sp.  
762 nov., a high-molecular-weight polycyclic aromatic hydrocarbon-degrading bacterium within the class  
763 Gammaproteobacteria, and proposal of Immundisolibacterales ord. nov. and Immundisolibacteraceae fam.  
764 nov. *International journal of systematic and evolutionary microbiology*, 67(4), 925 (2017).
- 765 63. Gutierrez, T. Aerobic Hydrocarbon-Degrading Gammaproteobacteria: Xanthomonadales. *Taxonomy*,  
766 *Genomics and Ecophysiology of Hydrocarbon-Degrading Microbes*, 191-205 (2019).
- 767 64. Kanaly, R. A., Harayama, S., Watanabe, K. *Rhodanobacter* sp. strain BPC1 in a benzo[a]pyrene-mineralizing  
768 bacterial consortium. *Applied and Environmental Microbiology*, 68(12), 5826-5833 (2002).
- 769 65. Dominguez, J. J. A., Inoue, C., Chien, M. F. Hydroponic approach to assess rhizodegradation by sudangrass  
770 (*Sorghum x drummondii*) reveals pH-and plant age-dependent variability in bacterial degradation of  
771 polycyclic aromatic hydrocarbons (PAHs). *Journal of Hazardous Materials*, 387, 121695 (2020).
- 772 66. Roslund, M. I., *et al.* Half-lives of PAHs and temporal microbiota changes in commonly used urban  
773 landscaping materials. *PeerJ* 1–27 (2018).
- 774 67. Nayak, A. S., Sanjeev Kumar, S., Anjaneya, O., Karegoudar, T. B. A catabolic pathway for the degradation  
775 of chrysene by *Pseudoxanthomonas* sp. PNK-04. *FEMS microbiology letters*, 320(2), 128-134 (2011).
- 776 68. Patel, V., Cheturvedula, S., Madamwar, D. Phenanthrene degradation by *Pseudoxanthomonas* sp. DMVP2  
777 isolated from hydrocarbon contaminated sediment of Amlakhadi canal, Gujarat, India. *Journal of hazardous*  
778 *materials*, 201, 43-51 (2012).
- 779 69. Choi, E. J., *et al.* Comparative genomic analysis and benzene, toluene, ethylbenzene, and o-, m-, and p-xylene  
780 (BTEX) degradation pathways of *Pseudoxanthomonas spadix* BD-a59. *Applied and environmental*  
781 *microbiology*, 79(2), 663-671. (2013).
- 782 70. Nayak, A. S., Vijaykumar, M. H., Karegoudar, T. B. Characterization of biosurfactant produced by  
783 *Pseudoxanthomonas* sp. PNK-04 and its application in bioremediation. *International Biodeterioration &*  
784 *Biodegradation*, 63(1), 73-79. (2009).
- 785 71. Knief, C. Diversity and habitat preferences of cultivated and uncultivated aerobic methanotrophic bacteria  
786 evaluated based on *pmoA* as molecular marker. *Frontiers in microbiology*, 6, 1346. (2015).
- 787 72. Schleheck, D., *et al.* Complete genome sequence of *Parvibaculum lavamentivorans* type strain (DS-1 T).  
788 *Standards in genomic sciences*, 5(3), 298-310 (2011).
- 789 73. Alonso-Gutiérrez, J., *et al.* Bacterial communities from shoreline environments (Costa da Morte,  
790 Northwestern Spain) affected by the Prestige oil spill. *Applied and Environmental Microbiology*, 75(11),  
791 3407-3418. (2009).
- 792 74. Lai, Q., *et al.* *Parvibaculum indicum* sp. nov., isolated from deep-sea water. *International journal of*  
793 *systematic and evolutionary microbiology*, 61(2), 271-274 (2011).



- 794 75. Obuekwe, C. O., Al-Jadi, Z. K., Al-Saleh, E. S. Hydrocarbon degradation in relation to cell-surface  
795 hydrophobicity among bacterial hydrocarbon degraders from petroleum-contaminated Kuwait desert  
796 environment. *International Biodeterioration & Biodegradation*, 63(3), 273-279. (2009).
- 797 76. Blanco-Enriquez, E. G., *et al.* Characterization of a microbial consortium for the bioremoval of polycyclic  
798 aromatic hydrocarbons (PAHs) in water. *International journal of environmental research and public health*,  
799 15(5), 975. (2018).
- 800 77. Radwan, S. S., Al-Awadhi, H., Sorkhoh, N. A., El-Nemr, I. M. Rhizospheric hydrocarbon-utilizing  
801 microorganisms as potential contributors to phytoremediation for the oil Kuwaiti desert. *Microbiological*  
802 *Research*, 153(3), 247-251 (1998).
- 803 78. Arli, S. D., Trivedi, U. B., Patel, K. C. Curdlan-like exopolysaccharide production by *Cellulomonas flavigena*  
804 UNP3 during growth on hydrocarbon substrates. *World Journal of Microbiology and Biotechnology*, 27(6),  
805 1415-1422 (2011).
- 806 79. Parab, V., & Phadke, M. Co-biodegradation studies of naphthalene and phenanthrene using bacterial  
807 consortium. *Journal of Environmental Science and Health, Part A*, 1-13 (2020).
- 808 80. Kim, S. J., *et al.* Genomic analysis of polycyclic aromatic hydrocarbon degradation in *Mycobacterium*  
809 *vanbaalenii* PYR-1. *Biodegradation*, 19(6), 859-881 (2008).
- 810 81. Van Beilen, J. B., & Funhoff, E. G. (2007). Alkane hydroxylases involved in microbial alkane degradation.  
811 *Applied microbiology and biotechnology*, 74(1), 13-21.
- 812 82. Arino, S., Marchal, R., Vandecasteele, J. P. Production of new extracellular glycolipids by a strain of  
813 *Cellulomonas cellulans* (*Oerskovia xanthineolytica*) and their structural characterization. *Canadian journal*  
814 *of microbiology*, 44(3), 238-243 (1998).
- 815 83. Abbasian, F., Lockington, R., Megharaj, M., Naidu, R. The biodiversity changes in the microbial population  
816 of soils contaminated with crude oil. *Current microbiology*, 72(6), 663-670 (2016).
- 817 84. Neethu, C. S., *et al.* Oil-spill triggered shift in indigenous microbial structure and functional dynamics in  
818 different marine environmental matrices. *Scientific reports*, 9(1), 1-13 (2019).
- 819 85. Xie, S., Sun, W., Luo, C., Cupples, A. M. Novel aerobic benzene degrading microorganisms identified in  
820 three soils by stable isotope probing. *Biodegradation*, 22(1), 71-81 (2011).
- 821 86. Daccò, C., *et al.* Key fungal degradation patterns, enzymes and their applications for the removal of aliphatic  
822 hydrocarbons in polluted soils: A review. *International Biodeterioration & Biodegradation*, 147, 104866  
823 (2020).
- 824 87. Lindahl, B. D., & Tunlid, A. Ectomycorrhizal fungi—potential organic matter decomposers, yet not  
825 saprotrophs. *New Phytologist*, 205(4), 1443-1447 (2015).
- 826 88. Mohammadi-Sichani, M. M., *et al.* Bioremediation of soil contaminated crude oil by Agaricomycetes. *Journal*  
827 *of Environmental Health Science and Engineering*. 15:8 (2017).
- 828 89. Treu, R., & Falandysz, J. Mycoremediation of hydrocarbons with basidiomycetes—a review. *Journal of*  
829 *Environmental Science and Health, Part B*, 52(3), 148-155 (2017).
- 830 90. Lemmel, F., Maunoury-Danger, F., Leyval, C., Cébron, A. Altered fungal communities in contaminated soils  
831 from French industrial brownfields. *Journal of Hazardous Materials*, 124296 (2020).
- 832 91. Aguilar, M., *et al.* Next-generation sequencing assessment of eukaryotic diversity in oil sands tailings ponds  
833 sediments and surface water. *Journal of Eukaryotic Microbiology*, 63(6), 732-743 (2016).

- 834 92. Czaplicki, L. M., *et al.* Investigating the mycobiome of the Holcomb Creosote Superfund Site. *Chemosphere*,  
835 126208 (2020).
- 836 93. Liu, P. W. G., *et al.* Bioremediation of petroleum hydrocarbon contaminated soil: effects of strategies and  
837 microbial community shift. *International Biodeterioration & Biodegradation*, 65(8), 1119-1127 (2011).
- 838 94. Röling, W. F., Head, I. M., Larter, S. R. The microbiology of hydrocarbon degradation in subsurface  
839 petroleum reservoirs: perspectives and prospects. *Research in Microbiology*, 154(5), 321-328 (2003).
- 840 95. Tischer, K., *et al.* Microbial communities along biogeochemical gradients in a hydrocarbon-contaminated  
841 aquifer. *Environmental microbiology*, 15(9), 2603-2615 (2013).
- 842 96. Mbadinga, S. M., *et al.* Microbial communities involved in anaerobic degradation of alkanes. *International*  
843 *Biodeterioration & Biodegradation*, 65(1), 1-13 (2011).
- 844 97. Kleinsteuber, S., Schleinitz, K. M., Vogt, C. Key players and team play: anaerobic microbial communities in  
845 hydrocarbon-contaminated aquifers. *Applied microbiology and biotechnology*, 94(4), 851-873 (2012).
- 846 98. Bauer, R.D., *et al.* Mixing-controlled biodegradation in a toluene plume—results from two-dimensional  
847 laboratory experiments. *Journal of Contaminant Hydrology*, 96(1-4), 150-168 (2008).
- 848 99. Morris, B.E., Henneberger, R., Huber, H., Moissl-Eichinger, C. Microbial syntrophy: interaction for the  
849 common good. *FEMS microbiology reviews*, 37(3), 384-406 (2013).
- 850 100. Vogt, C., *et al.* Stable isotope probing approaches to study anaerobic hydrocarbon degradation and degraders.  
851 *Journal of Molecular Microbiology and Biotechnology*, 26(1-3), 195-210 (2016).
- 852 101. da Cruz, G. F., *et al.* Could petroleum biodegradation be a joint achievement of aerobic and anaerobic  
853 microorganisms in deep sea reservoirs? *AMB express*, 1(1), 47 (2011).
- 854 102. Atashgahi, S., *et al.* A benzene-degrading nitrate-reducing microbial consortium displays aerobic and  
855 anaerobic benzene degradation pathways. *Scientific reports*, 8(1), 1-12 (2018).
- 856 103. Abed, R. M., Al-Kindi, S., Al-Kharusi, S. Diversity of bacterial communities along a petroleum contamination  
857 gradient in desert soils. *Microbial ecology*, 69(1), 95-105 (2015).
- 858 Wang, X., Feng, J., & Zhao, J. (2010). Effects of crude oil residuals on soil chemical properties in oil sites,  
859 Momoge Wetland, China. *Environmental monitoring and assessment*, 161(1), 271-280.
- 860 Sun, W., Dong, Y., Gao, P., Fu, M., Ta, K., & Li, J. (2015). Microbial communities inhabiting oil-contaminated  
861 soils from two major oilfields in Northern China: Implications for active petroleum-degrading capacity. *Journal*  
862 *of microbiology*, 53(6), 371-378.
- 863 Winderl, C., Anneser, B., Griebler, C., Meckenstock, R. U., & Lueders, T. (2008). Depth-resolved quantification  
864 of anaerobic toluene degraders and aquifer microbial community patterns in distinct redox zones of a tar oil  
865 contaminant plume. *Applied and Environmental Microbiology*, 74(3), 792-801.
- 866 Yang, S., Wen, X., Shi, Y., Liebner, S., Jin, H., & Perfumo, A. (2016). Hydrocarbon degraders establish at the  
867 costs of microbial richness, abundance and keystone taxa after crude oil contamination in permafrost  
868 environments. *Scientific Reports*, 6(1), 1-13.
- 869 Yang, S., Wen, X., Zhao, L., Shi, Y., & Jin, H. (2014). Crude oil treatment leads to shift of bacterial communities  
870 in soils from the deep active layer and upper permafrost along the China-Russia crude oil pipeline route. *PloS one*,  
871 9(5), e96552.
- 872 Sutton, N. B., Maphosa, F., Morillo, J. A., Abu Al-Soud, W., Langenhoff, A. A., Grotenhuis, T., ... & Smidt, H.  
873 (2013). Impact of long-term diesel contamination on soil microbial community structure. *Applied and*  
874 *environmental microbiology*, 79(2), 619-630.
- 875 Ferrari, B.C., Zhang C.D. & Van Dorst J. 2011. Recovering greater fungal diversity from pristine and diesel fuel  
876 contaminated sub-Antarctic soil through cultivation using both a high and a low nutrient media approach. *Frontiers*  
877 *in Microbiology*, 2, 10.3389/fmicb.2011.00217.
- 878 Muktali V., Park J., Fedorova-Abrams N.D., Park B., Choi J., Lee Y.H., & Kang S. 2012. Systematic and  
879 searchable classification of cytochrome P450 proteins encoded by fungal and oomycete  
880 genomes. *BMC Genomics*, 13, 10 1186/1471-2164-13-525.

881 Cury, J. C., Jurelevicius, D. A., Villela, H. D., Jesus, H. E., Peixoto, R. S., Schaefer, C. E., ... & Rosado, A. S.  
882 (2015). Microbial diversity and hydrocarbon depletion in low and high diesel-polluted soil samples from Keller  
883 Peninsula, South Shetland Islands. *Antarctic Science*, 27(3), 263-273.

884 Paissé, S., Coulon, F., Goñi-Urriza, M., Peperzak, L., McGenity, T. J., & Duran, R. (2008). Structure of bacterial  
885 communities along a hydrocarbon contamination gradient in a coastal sediment. *FEMS microbiology ecology*,  
886 66(2), 295-305.

887 Aßhauer KP, Bernd W, Rolf D et al (2015) Tax4Fun: predicting functional proles from metagenomic 16S  
888 rRNA data. *Bioinformatics* 31:2882–2884

889 Zhang, S., Hu, Z., & Wang, H. (2019). Metagenomic analysis exhibited the co-metabolism of polycyclic aromatic  
890 hydrocarbons by bacterial community from estuarine sediment. *Environment international*, 129, 308-319.

891 Mukherjee, A., Chettri, B., Langpoklakpam, J. S., Basak, P., Prasad, A., Mukherjee, A. K., ... & Chattopadhyay,  
892 D. (2017). Bioinformatic approaches including predictive metagenomic profiling reveal characteristics of bacterial  
893 response to petroleum hydrocarbon contamination in diverse environments. *Scientific reports*, 7(1), 1-22.

894 Das, D., Baruah, R., Roy, A. S., Singh, A. K., Boruah, H. P. D., Kalita, J., & Bora, T. C. (2015). Complete genome  
895 sequence analysis of *Pseudomonas aeruginosa* N002 reveals its genetic adaptation for crude oil degradation.  
896 *Genomics*, 105(3), 182-190.

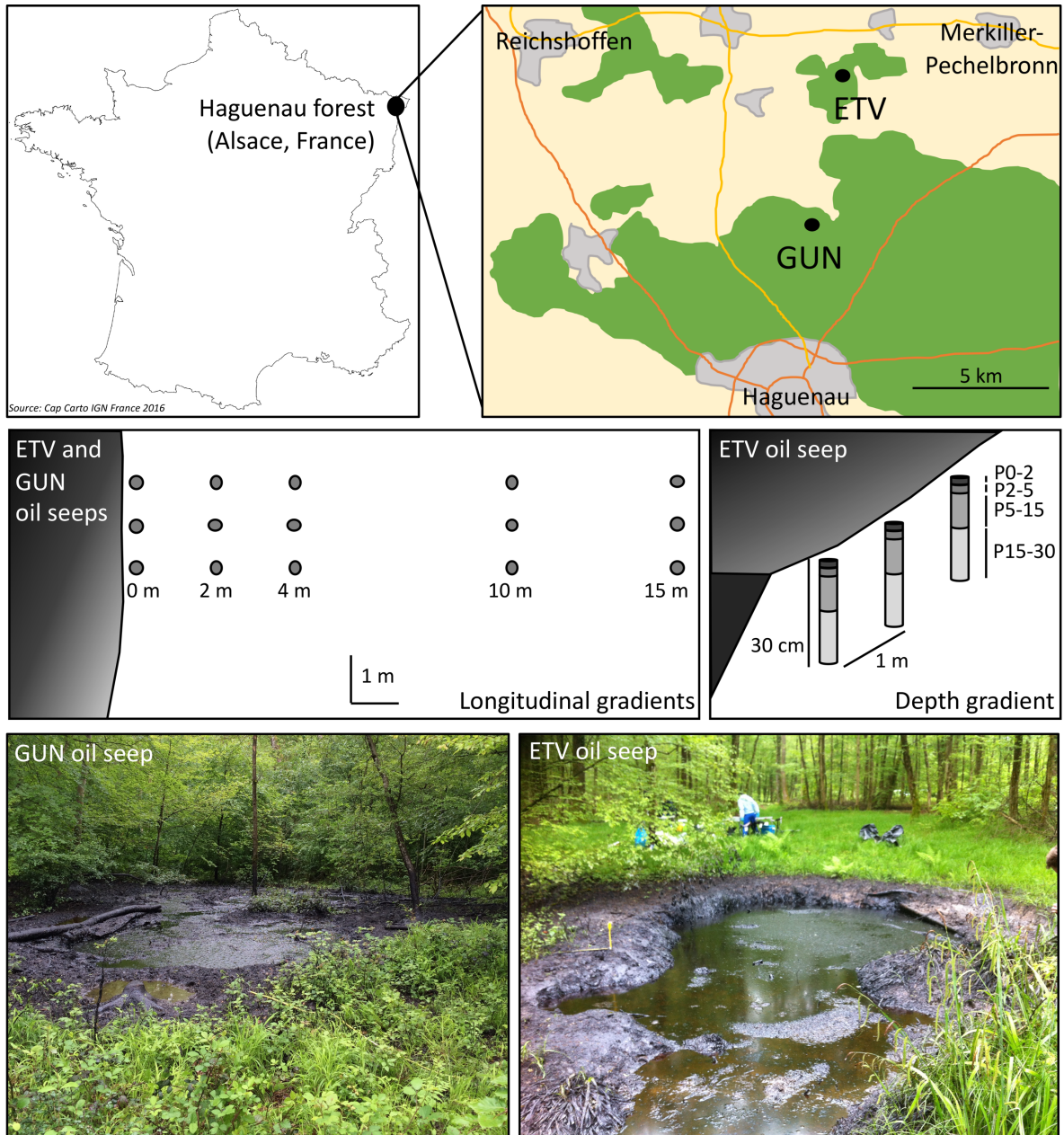
897 Gutierrez, T., Nichols, P. D., Whitman, W. B., & Aitken, M. D. (2012). *Porticoccus hydrocarbonoclasticus* sp.  
898 nov., an aromatic hydrocarbon-degrading bacterium identified in laboratory cultures of marine phytoplankton.  
899 *Applied and environmental microbiology*, 78(3), 628-637.

900  
901  
902

903 **Figure caption.**

904

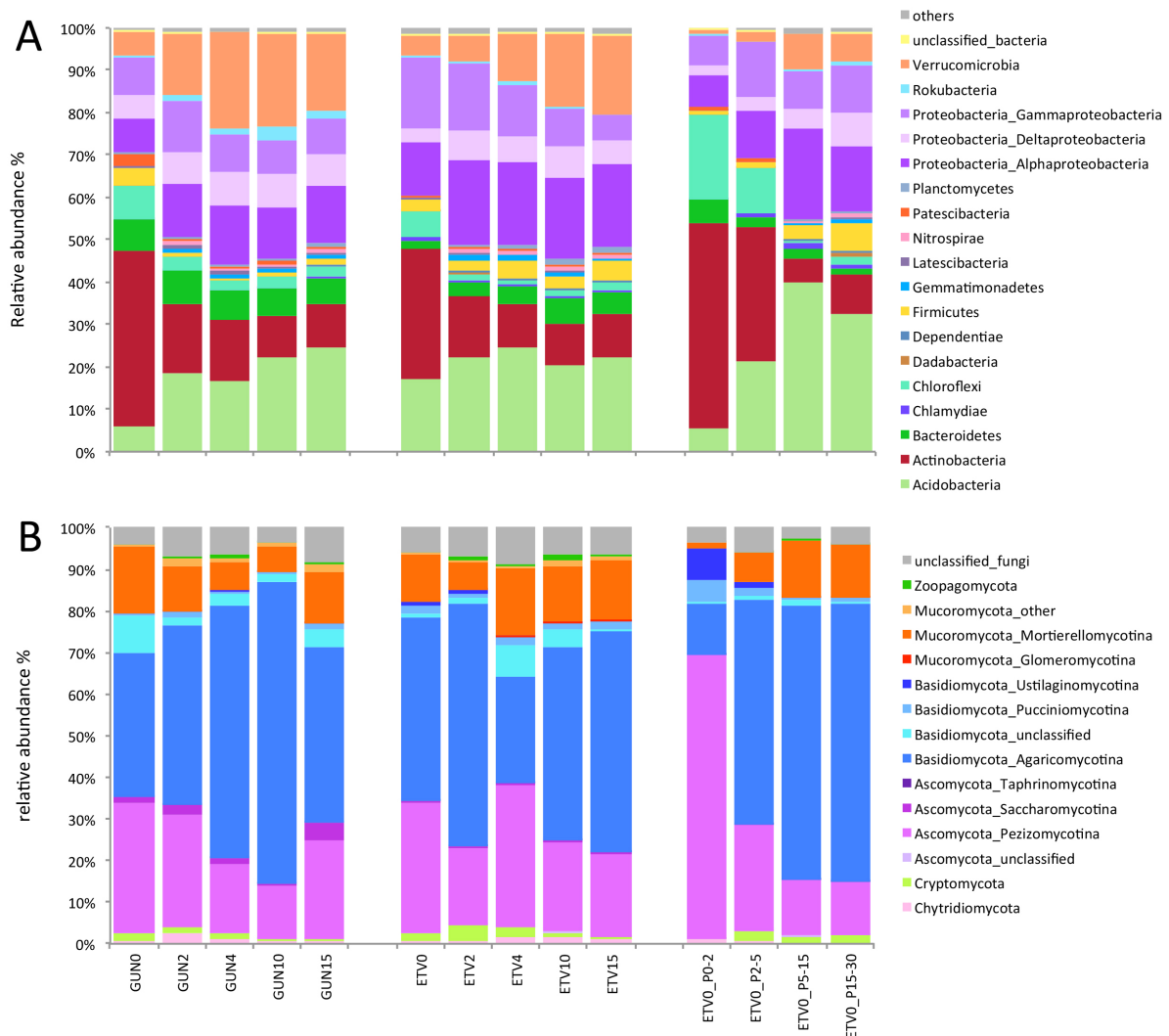
905 **Figure 1. Map showing the location of the ETV and GUN study sites, schematic**  
906 **illustrations of the sampling designs for longitudinal and depth gradients and pictures of**  
907 **the ponds created by oil seeps.**



908

909

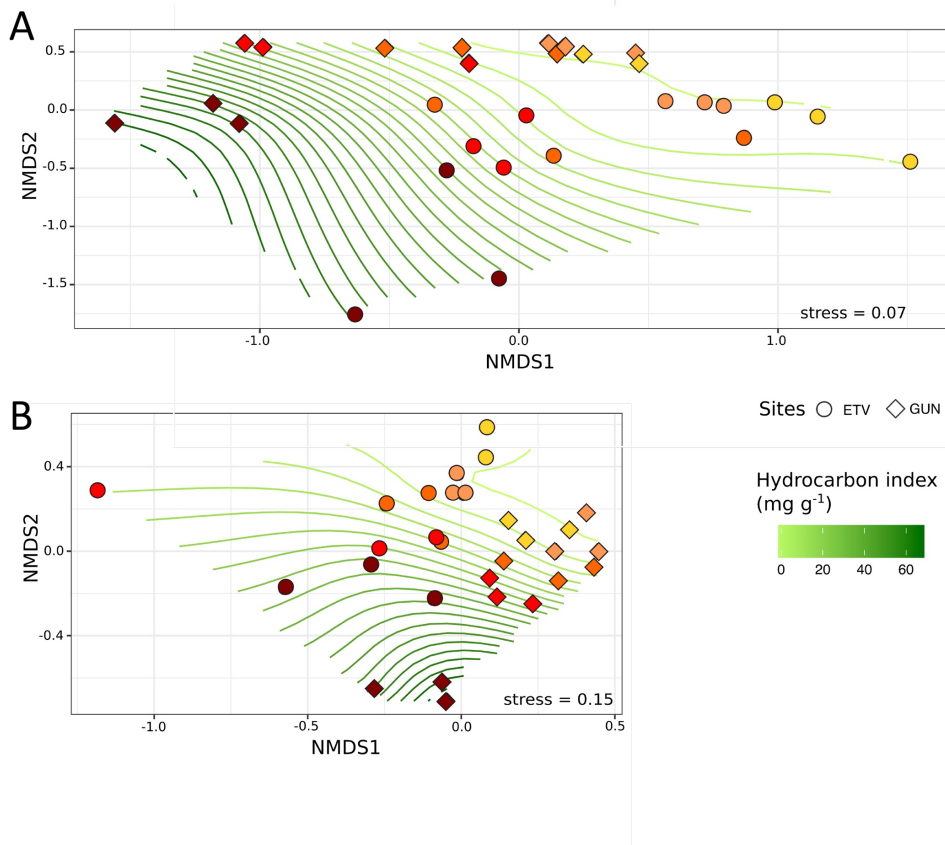
910 **Figure 2. Changes in relative abundance of bacterial phyla (A) and fungal sub-divisions**  
 911 **(B) in GUN and ETV sites according to the distance from the petroleum seeps (0, 2, 4, 10**  
 912 **and 15 m) and depth (P0-2, P0-5, P2-5, P5-15 and P15-30 cm) at ETV site for the 0 m**  
 913 **distance.** Data are means (n=3). Different classes of Proteobacteria (plot A) and sub-divisions  
 914 of Mucoromycota, Basidiomycota and Ascomycota (plot B) were shown. *The group “Others”*  
 915 *(plot A) includes phyla below 0.5% (267 reads) of relative abundance and affiliated to*  
 916 *Armatimonadetes, BCRI, Caldiserica, Calditrichaeota, Cyanobacteria, Deferribacteres,*  
 917 *Deinococcus-Thermus, Elusimicrobia, Entotheonellaeota, FBP, FCPU426, Fibrobacteres,*  
 918 *Firestonebacteria, Fusobacteria, GAL15, Halanaerobiaeota, Hydrogenedentes,*  
 919 *Kiritimatiellaeota, Lentisphaerae, Margulisbacteria, Nitrospinae, Omnitrophicaeota,*  
 920 *Schekmanbacteria, Spirochaetes, TA06, Tenericutes, unclassified\_bacteria, WPS\_2, WS2, WS4*  
 921 *and Zixibacteria.*



922

923

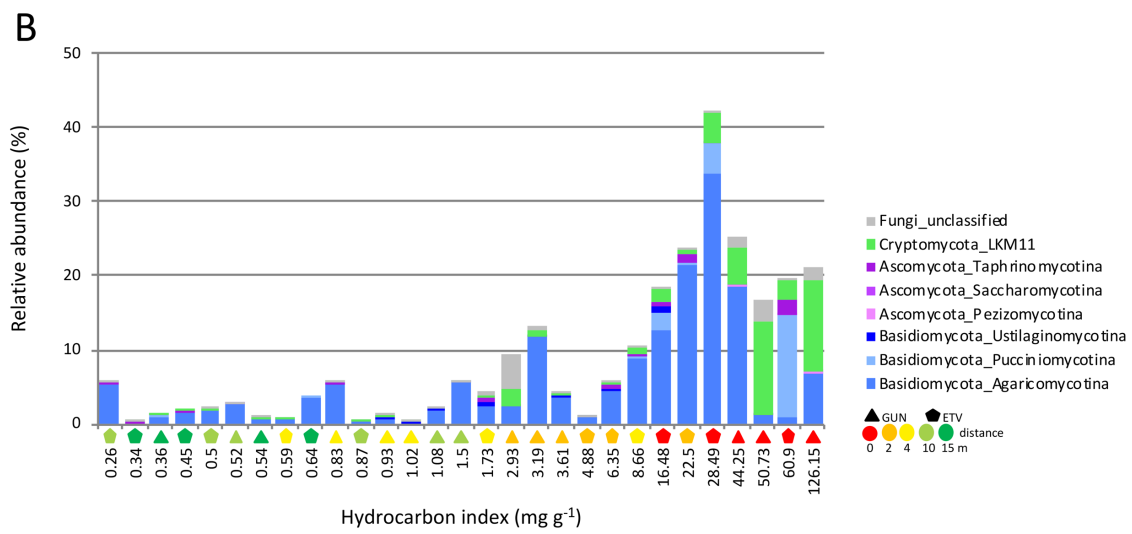
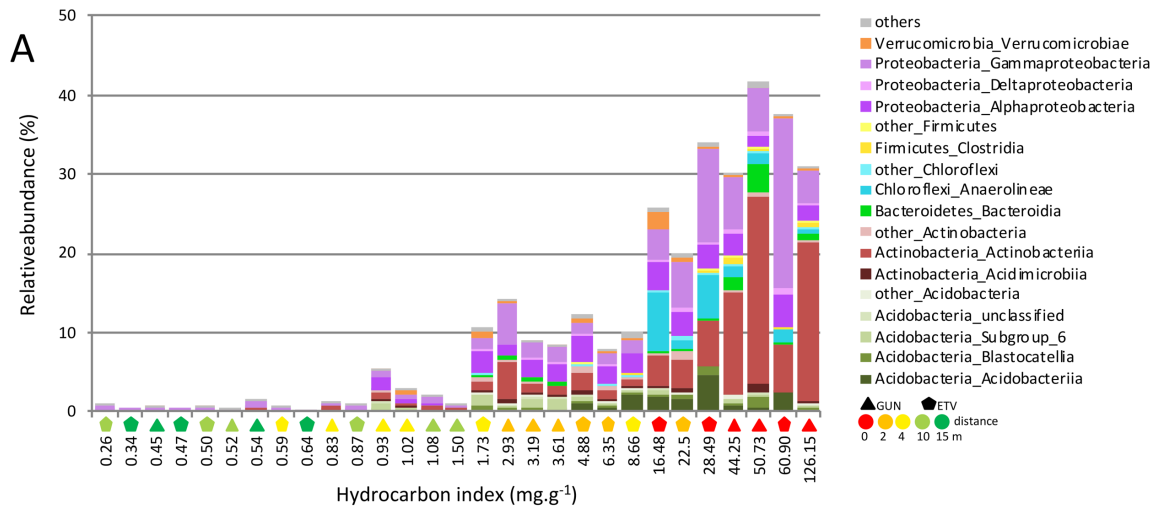
924 **Figure 3. Ordisurf model showing the distribution of hydrocarbon index (HI) level over**  
925 **the NMDS ordination of bacterial (A) and fungal (B) community structures at the OTU**  
926 **level.** The hydrocarbon gradient is visualized through green curves. ETV and GUN sites are  
927 differentiated through circles and diamonds, respectively. Different sampling zones are  
928 visualized thanks to colour gradients from dark red (0m, close to the oil seep), red (2 m), orange  
929 (4 m), light orange (10 m) and yellow (15 m).



930

931

932 **Figure 4. Relative abundance of increasing bacterial (A) and fungal (B) indicator taxa**  
933 **(summed at the class and sub-division levels for bacteria and fungi, respectively)**  
934 **determined using TITAN2 along the gradient of oil contamination for the combination of**  
935 **both ETV and GUN sites.** ETV and GUN sites are symbolised by pentagon and triangle  
936 symbols, respectively, and distance to the oil seep is shown with colour gradient (from red to  
937 green). *At the OTU level, 281 and 61 bacterial and fungal indicator taxa were sorted with*  
938 *purity and reliability parameters both >0.95. For the complete list of bacterial and fungal OTU*  
939 *found as indicator taxa see **Table S1** and **S2**, respectively. The relative abundances of OTU*  
940 *belonging to similar classes were summed. The group “Others” comprise classes having less*  
941 *than 3 representative OTUs and were affiliated to Chlamydiae, Melainanabacteria*  
942 *(Cyanobacteria phylum), Dadabacteriia, Babeliae (Dependentiae phylum),*  
943 *Gemmatimonadetes, Latescibacteria, Nitrospira, Saccharimonadia (Patescibacteria phylum)*  
944 *and Rokubacteria NC10 classe. The group “other\_Acidobacteria” comprise OTU affiliated to*  
945 *Holophagae, subgroup\_5 and subgroup\_22 classes. The group “other\_Actinobacteria”*  
946 *comprise OTU affiliated to Thermoleophilia and Coriobacteriia classes. The group*  
947 *“other\_Chloroflexi” comprise OTU affiliated to Chloroflexia, TK10 and JG30-KF-CM66*  
948 *classes. The group “other\_Firmicutes” comprise OTU affiliated to Bacilli and BRH-c20a*  
949 *classes.*



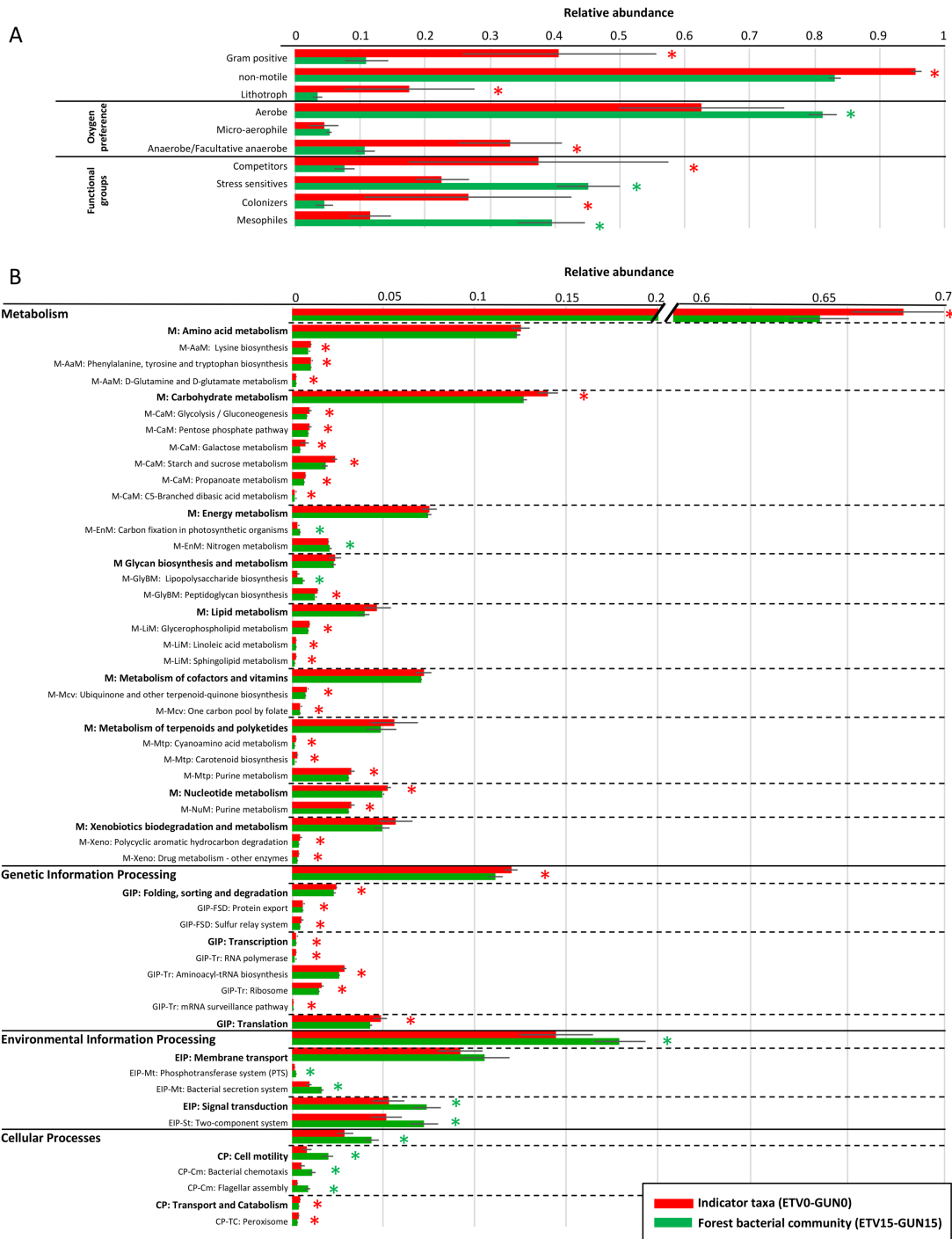
950

951

952



953 **Figure 5. Functional traits of the increasing bacterial indicator taxa (detected in ETV0**  
954 **and GUN0, in red) compared to control forest soil bacteria (detected in ETV15 and**  
955 **GUN15, in green) based on Bactotraits (A) and Tax4Fun (B) analyses.** Only traits with  
956 significant differences between the 2 populations were represented (except for few functional  
957 pathways at the Level2 on plot B) and labelled with a \* meaning that Wilcoxon rank sum tests  
958 had a p value <0.05. Trait attributes are expressed as mean percentage (n=6 for increasing  
959 Indicator taxa, and n=5 for forest bacterial community) and  $\pm$  standard deviation. **A.**  
960 Morphological (Gram staining: the rest are Gram negatives, and motility: the rest are motile  
961 bacteria possessing flagella) and physiological (trophic type as lithotroph, i.e. bacteria that  
962 obtains its energy (electron donor) from inorganic compounds (iron, sulfur...) in opposition to  
963 organotroph, and oxygen preference: all attribute classes are shown) and affiliation to functional  
964 groups (5 functional groups were defined, no difference was found for stress-tolerant group)  
965 were assessed using BactoTraits (Cébron et al. 2021). **B.** Genomic traits, i.e. KEGG functional  
966 pathways shown at Level 1, Level 2 and Level 3. Only functions representing more than 0.1%  
967 were represented. *M: Metabolism, GIP: Genetic information processing, EIP: Environmental*  
968 *Information Processing and CP: Cellular Processes.*



970 **Table 1. Pollution and soil characteristics along longitudinal and depth gradients.** Mean (n=3) and standard error (when no standard error is indicated it  
 971 means that only one value was available, i.e. measurement was performed on a mean sample constituted by a pool of the 3 replicates). Letters indicate significant  
 972 differences (p<0.05) among sampling zones of the same site and sampling date (one-way anova followed by Tukey HSD PostHoc test). GUN: Gunstett site,  
 973 ETV: Etang-vert site. TOC: Total organic carbon, N: total nitrogen, C/N: ratio of carbon to nitrogen.

	Sampling date	Distance from oil seep (m)	Depth (cm)	Extractable organic matter, EOM (mg g <sup>-1</sup> )	Hydrocarbon index, HI (mg g <sup>-1</sup> )	Carbon preference index, CPI	TOC (mg g <sup>-1</sup> )	N (mg g <sup>-1</sup> )	C/N	pH
GUN0		0	0-5	119.2 ± 21.9 <sup>a</sup>	73.7 ± 26.3 <sup>a</sup>	1.0	142	3.5	40.3	6.88
GUN2		2	0-5	11.7 ± 0.9 <sup>b</sup>	3.2 ± 0.2 <sup>b</sup>	6.2	80	4.8	16.6	7.00
GUN4	May 2016	4	0-5	3.0 ± 0.2 <sup>b</sup>	0.9 ± 0.1 <sup>b</sup>	10.0	61	4.1	14.8	6.05
GUN10		10	0-5	3.2 ± 0.8 <sup>b</sup>	1.0 ± 0.3 <sup>b</sup>	10.8	42	2.8	15.1	5.76
GUN15		15	0-5	1.8 ± 0.2 <sup>b</sup>	0.4 ± 0.1 <sup>b</sup>	9.7	43	3.0	14.4	5.51
ETV0		0	0-5	118.6 ± 44.6 <sup>a</sup>	35.3 ± 13.3 <sup>a</sup>	1.0	154	3.5	43.8	6.09
ETV2		2	0-5	48.7 ± 20.2 <sup>ab</sup>	11.2 ± 5.6 <sup>ab</sup>	2.3	103	3.8	27.0	6.14
ETV4	May 2016	4	0-5	20.9 ± 14.4 <sup>ab</sup>	3.7 ± 2.5 <sup>b</sup>	3.1	57	2.4	23.7	6.03
ETV10		10	0-5	6.3 ± 4.4 <sup>b</sup>	0.5 ± 0.2 <sup>b</sup>	6.4	23	1.4	16.9	5.77
ETV15		15	0-5	14.8 ± 9.4 <sup>ab</sup>	0.5 ± 0.1 <sup>b</sup>	6.2	27	1.7	15.7	5.19
ETV0_P0-2		0	0-2	108.9 ± 9.0 <sup>ab</sup>	42.3 ± 1.9 <sup>a</sup>	na	141 ± 2 <sup>a</sup>	3.1 ± 0.2 <sup>a</sup>	46.2 ± 2.8 <sup>ab</sup>	na
ETV0_P2-5	May 2017	0	2-5	113.0 ± 18.2 <sup>a</sup>	24.9 ± 3.1 <sup>b</sup>	na	142 ± 19 <sup>a</sup>	3.8 ± 0.5 <sup>a</sup>	37.6 ± 1.1 <sup>ab</sup>	na
ETV0_P5-15		0	5-15	73.0 ± 27.3 <sup>bc</sup>	11.6 ± 5.1 <sup>c</sup>	na	42 ± 9 <sup>b</sup>	1.7 ± 0.4 <sup>ab</sup>	25.4 ± 0.8 <sup>b</sup>	na
ETV0_P15-30		0	15-30	31.7 ± 19.5 <sup>c</sup>	5.6 ± 3.7 <sup>c</sup>	na	13 ± 3 <sup>b</sup>	0.4 ± 0.2 <sup>b</sup>	51.5 ± 19.6 <sup>a</sup>	na

974  
 975  
 976

977 **Table 2. Microbial mineralization activities, abundances (16S and 18S rRNA gene copy numbers for bacteria and fungi, respectively) and alpha-**  
 978 **diversity indices (Chao1 richness, Shannon diversity and Pielou's evenness) along longitudinal and depth gradients.** Samples from May 2016 were  
 979 sampled in surface (0-5 cm depth) and samples from May 2017 were sampled at 4 depth (0-2, 2-5, 5-15 and 15-30 cm depth). Mean (n=3) and standard deviation.  
 980 Letters indicate significant differences (p<0.05) among sampling zones of the same site and sampling date (one-way anova followed by a Tukey HSD PostHoc  
 981 test). GUN: Gunstett site, ETV: Etang-vert site. Mineralization activities and rRNA gene copies are expressed as gram of dry weight soil.

Sampling date	Distance from oil seep (m)	Aerobic mineralization CO <sub>2</sub> production (µg C-CO <sub>2</sub> g <sup>-1</sup> d <sup>-1</sup> )	Anaerobic mineralization CO <sub>2</sub> production (µg C-CO <sub>2</sub> g <sup>-1</sup> d <sup>-1</sup> )	Bacteria				Fungi				
				16S rRNA gene copies (x 10 <sup>10</sup> g <sup>-1</sup> )	Chao 1 richness	Shannon diversity H'	Pielou's evenness	18S rRNA gene copies (x 10 <sup>8</sup> g <sup>-1</sup> )	Chao 1 richness	Shannon diversity H'	Pielou's evenness	
May 2016	GUN0	0	248 ± 90 <sup>a</sup>	61 ± 5 <sup>a</sup>	8.7 ± 3.3 <sup>ab</sup>	6532 ± 441	6.0 ± 0.4 <sup>b</sup>	0.70 ± 0.04 <sup>b</sup>	10.0 ± 1.6	349 ± 10	3.05 ± 0.16 <sup>b</sup>	0.55 ± 0.03 <sup>b</sup>
	GUN2	2	94 ± 7 <sup>b</sup>	31 ± 1 <sup>b</sup>	9.6 ± 2.2 <sup>ab</sup>	7292 ± 676	7.3 ± 0.2 <sup>a</sup>	0.83 ± 0.02 <sup>a</sup>	12.8 ± 4.6	429 ± 56	3.72 ± 0.22 <sup>a</sup>	0.64 ± 0.02 <sup>a</sup>
	GUN4	4	76 ± 13 <sup>b</sup>	26 ± 5 <sup>b</sup>	10.0 ± 2.1 <sup>a</sup>	7988 ± 435	7.2 ± 0.4 <sup>a</sup>	0.82 ± 0.04 <sup>a</sup>	26.7 ± 16.1	448 ± 79	3.29 ± 0.39 <sup>ab</sup>	0.57 ± 0.05 <sup>ab</sup>
	GUN10	10	61 ± 25 <sup>b</sup>	26 ± 9 <sup>b</sup>	5.1 ± 0.3 <sup>b</sup>	7156 ± 573	7.2 ± 0.2 <sup>a</sup>	0.82 ± 0.02 <sup>a</sup>	14.7 ± 3.9	426 ± 6	2.55 ± 0.18 <sup>c</sup>	0.45 ± 0.03 <sup>c</sup>
	GUN15	15	47 ± 11 <sup>b</sup>	18 ± 4 <sup>b</sup>	6.0 ± 1.6 <sup>ab</sup>	7697 ± 158	7.3 ± 0.0 <sup>a</sup>	0.84 ± 0.01 <sup>a</sup>	14.4 ± 6.7	460 ± 65	3.45 ± 0.14 <sup>ab</sup>	0.59 ± 0.02 <sup>ab</sup>
May 2016	ETV0	0	192 ± 39 <sup>a</sup>	32 ± 12 <sup>a</sup>	6.9 ± 1.7	4683 ± 2150 <sup>b</sup>	5.1 ± 1.3 <sup>b</sup>	0.62 ± 0.13 <sup>b</sup>	8.2 ± 3.8	450 ± 44	3.28 ± 0.38	0.56 ± 0.06
	ETV2	2	134 ± 72 <sup>ab</sup>	38 ± 7 <sup>a</sup>	8.4 ± 1.7	8630 ± 711 <sup>a</sup>	7.5 ± 0.1 <sup>a</sup>	0.84 ± 0.02 <sup>a</sup>	48.1 ± 61.3	421 ± 154	2.79 ± 1.87	0.47 ± 0.30
	ETV4	4	56 ± 46 <sup>b</sup>	18 ± 8 <sup>ab</sup>	8.5 ± 3.0	7908 ± 944 <sup>a</sup>	7.3 ± 0.3 <sup>a</sup>	0.83 ± 0.03 <sup>a</sup>	12.2 ± 0.1	498 ± 69	3.59 ± 0.55	0.60 ± 0.08
	ETV10	10	44 ± 9 <sup>b</sup>	13 ± 7 <sup>b</sup>	5.6 ± 1.0	8670 ± 882 <sup>a</sup>	7.5 ± 0.3 <sup>a</sup>	0.84 ± 0.02 <sup>a</sup>	7.9 ± 2.3	562 ± 57	3.78 ± 0.06	0.62 ± 0.00
	ETV15	15	42 ± 7 <sup>b</sup>	7 ± 5 <sup>b</sup>	6.1 ± 0.7	7198 ± 1361 <sup>a</sup>	7.1 ± 0.2 <sup>a</sup>	0.81 ± 0.01 <sup>a</sup>	11.4 ± 1.4	492 ± 56	3.50 ± 0.21	0.59 ± 0.02
May 2017	ETV0_P0-2	0	207 ± 29 <sup>a</sup>	31 ± 3 <sup>a</sup>	10.6 ± 2.0 <sup>a</sup>	3539 ± 732 <sup>b</sup>	4.2 ± 0.6 <sup>c</sup>	0.53 ± 0.06 <sup>c</sup>	6.5 ± 1.9	355 ± 5	3.03 ± 0.23	0.54 ± 0.05
	ETV0_P2-5	0	88 ± 16 <sup>b</sup>	24 ± 6 <sup>ab</sup>	6.1 ± 0.9 <sup>ab</sup>	4988 ± 479 <sup>a</sup>	5.1 ± 0.3 <sup>b</sup>	0.61 ± 0.03 <sup>b</sup>	3.0 ± 0.5	395 ± 35	2.88 ± 0.62	0.50 ± 0.10
	ETV0_P5-15	0	51 ± 35 <sup>bc</sup>	16 ± 8 <sup>b</sup>	5.1 ± 2.2 <sup>b</sup>	5204 ± 921 <sup>a</sup>	6.4 ± 0.4 <sup>a</sup>	0.76 ± 0.04 <sup>a</sup>	5.2 ± 1.6	303 ± 56	2.15 ± 0.70	0.40 ± 0.11
	ETV0_P15-30	0	15 ± 5 <sup>c</sup>	4 ± 1 <sup>c</sup>	2.3 ± 0.6 <sup>c</sup>	5823 ± 160 <sup>a</sup>	6.8 ± 0.0 <sup>a</sup>	0.80 ± 0.00 <sup>a</sup>	1.9 ± 0.3	347 ± 74	2.46 ± 0.44	0.44 ± 0.06

982

983 **Table 3. Pearson correlation coefficients between the 3 bacterial alpha-diversity**  
 984 **indicators and the extractable organic matter, hydrocarbon index and total organic**  
 985 **carbon values for longitudinal and depth gradients.** P values are given by \* for <0.05, \*\*  
 986 for p<0.01 and \*\*\* for p<0.001.

		Chao1	Shannon H'	Evenness
GUN	Extractable organic matter	-0.65*	-0.91***	-0.91***
(longitudinal)	Hydrocarbon index	-0.59*	-0.85***	-0.85***
	TOC	-0.57*	-0.83***	-0.83***
ETV	Extractable organic matter	-0.68**	-0.85***	-0.88***
(longitudinal)	Hydrocarbon index	-0.71**	-0.87***	-0.89***
	TOC	-0.55*	-0.67**	-0.69**
ETV	Extractable organic matter	-0.72**	-0.74**	-0.72**
(depth)	Hydrocarbon index	-0.86***	-0.91***	-0.90***
	TOC	-0.70*	-0.93***	-0.93***

987

988