

# Stable isotope probing and metagenomics highlight the effect of plants on uncultured phenanthrene-degrading bacterial consortium in polluted soil

François Thomas, Erwan Corre, Aurélie Cebron

#### ▶ To cite this version:

François Thomas, Erwan Corre, Aurélie Cebron. Stable isotope probing and metagenomics highlight the effect of plants on uncultured phenanthrene-degrading bacterial consortium in polluted soil. The International Society of Microbiologial Ecology Journal, 2019, 13 (7), pp.1814-1830. 10.1038/s41396-019-0394-z. hal-02322773

### HAL Id: hal-02322773 https://hal.science/hal-02322773v1

Submitted on 11 Dec 2019

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 2 Stable isotope probing and metagenomics highlight the effect of plants on uncultured 3 phenanthrene-degrading bacteria in polluted soil 4 Running title: Phenanthrene degraders in bare vs. planted soil 5 6 François THOMAS<sup>1,2</sup>, Erwan CORRE<sup>3</sup>, Aurélie CEBRON<sup>1†</sup> 7 8 <sup>1</sup> Université de Lorraine, CNRS, LIEC 54500 Nancy, France 9 <sup>2</sup> Sorbonne Université, CNRS, Integrative Biology of Marine Models (LBI2M), 10 Station Biologique de Roscoff (SBR), 29680 Roscoff, France 11 <sup>3</sup> CNRS, Sorbonne Université, FR2424, ABiMS, Station Biologique de Roscoff, 29680, 12 13 Roscoff, France 14 15 † Corresponding author: aurelie.cebron@univ-lorraine.fr; Full postal address: LIEC 16 UMR7360, Faculté des Sciences et technologie, Bd des Aiguillettes, BP70239, 54506 17 Vandoeuvre-les-Nancy, France. Phone: (+33) 372745215. 18 19 ISME J sections: Microbial ecology and functional diversity of natural habitats 20 21 **Abstract** 22 Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous soil pollutants. The discovery that 23 plants can stimulate microbial degradation of PAHs promoted research on rhizoremediation 24 strategies. We combined DNA-SIP with metagenomics to assess the influence of plants on the 25 identity and metabolic functions of active PAH-degrading bacteria in contaminated soil, using phenanthrene (PHE) as a model hydrocarbon. <sup>13</sup>C-PHE dissipation was 2.5-fold lower in 26 27 ryegrass-planted conditions than in bare soil. Metabarcoding of 16S rDNA from bare and planted conditions revealed respectively 130 and 73 OTUs significantly enriched in <sup>13</sup>C-SIP 28 29 incubations compared to <sup>12</sup>C-controls. Active PHE-degraders were taxonomically diverse 30 (Proteobacteria, Actinobacteria and Firmicutes), with Sphingomonas and Sphingobium dominating in bare and planted soil, respectively. Metagenomes of <sup>13</sup>C-enriched DNA 31

fractions contained more genes involved in aromatic compounds metabolism in bare soil,

whereas carbohydrate catabolism genes were more abundant in planted soil. Functional gene

3233

annotation allowed reconstructing complete pathways for PHE catabolism. *Sphingomonadales* were the major taxa performing the first steps of PHE degradation in both conditions, suggesting their critical role to initiate *in-situ* PAH remediation. Analysis of metagenomic bins from <sup>13</sup>C-enriched *Sphingomonas* and *Sphingobium* highlighted different metabolic strategies, the latter containing less PAH-degrading genes but a more diverse repertoire of carbohydrate active enzymes potentially targeting plant root material.

4041

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

34

35

36

37

38

39

#### Introduction

Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous soil pollutants with environmental and public health concerns [1]. PAHs released into the environment by industrial activities can be removed through physical, chemical and biological processes [2]. Microorganisms that use PAHs as carbon sources play essential roles in natural attenuation of pollutants in contaminated ecosystems [3]. Numerous PAH catabolic pathways are described in bacteria, archaea and fungi, including both aerobic and anaerobic processes [1,3]. Several studies have evidenced an increased microbial degradation of PAHs in the presence of plants [4,5]. The discovery of this "rhizosphere effect" promoted the development of rhizoremediation strategies to treat PAH-contaminated soils [6,7]. However, literature shows contrasting results with no rhizosphere effect or even a negative influence on PAH-degradation efficiency [8,9]. These discrepancies may reflect the use of different model plants, soil types, indigenous microorganisms [10] and types of PAHs [11]. Such variations may also result from temporal variations in rhizospheric processes during plant development. We previously showed that phenanthrene degradation was slowed down during the early development of ryegrass [12], a plant commonly used in PAH rhizoremediation studies [13]. Differences in PAH-degradation with or without plants could be explained by the selection of two distinct PAH-degrading bacterial populations. Indeed, active bacterial populations were different in soil supplemented or not with root exudates [14] and in planted or bare soil microcosms [12]. Plant roots exude labile carbon sources that can be preferentially degraded by rhizospheric microbial communities to the detriment of PAH dissipation [9]. Therefore, investigation of PAHdegraders in the rhizosphere of contaminated soils is needed to decipher the influence of plants and ultimately improve PAH biodegradation efficiency. One challenge is to specifically target microorganisms actively degrading PAHs in polluted soils, without the cultivation bias. DNA-Stable Isotope Probing (DNA-SIP) is a powerful technique for linking functions with identity of uncultured microorganisms in complex microbiota. Approaches combining DNA-SIP and metagenomics are increasingly used to investigate microbial communities involved in anthropogenic compound biodegradation [15,16]. To date DNA-SIP has been successfully applied to identify soil bacteria that metabolize PAHs such as naphthalene [17,18,19,20], phenanthrene [14,18,21,22,23,24,25,26], anthracene [24,27,28], fluoranthene [24], pyrene [29,30] and benzo[a]pyrene [31]. DNA-SIP has also been used to explore PAH-degraders in the presence of purified root exudates [14] and to identify degraders of root exudates in the rhizosphere [32]. However, we still lack a comprehensive view on the major microbial actors of PAH degradation and their catabolic pathways in contaminated soils, and how it changes in the presence of plants.

In this study we investigated PAH-degrading bacteria of a historically polluted soil, using <sup>13</sup>C-labelled phenanthrene (PHE) as a metabolic tracer. We combined DNA-SIP with metagenomics to assess for the first time the influence of plant rhizosphere on the diversity, identity and metabolic functions of bacteria actively involved in phenanthrene degradation.

#### **Material and Methods**

#### Soil sample and phenanthrene spiking

Soil was collected from a former coking plant site (NM soil, Neuves-Maisons, France), dried at room temperature, sieved to 2 mm and stored in the dark at room temperature until experimental set-up. PAH contamination dates back ca. 100 years and reaches 1 260 mg Σ16PAHs.kg<sup>-1</sup>. Other soil characteristics were detailed elsewhere [33]. For spiking, two batches of soil were recontaminated at 2 500 mg.kg<sup>-1</sup> respectively with [U-<sup>13</sup>C<sub>14</sub>]-labeled phenanthrene (Sigma-Aldrich Isotec, St Louis, USA) and unlabeled <sup>12</sup>C-PHE (Fluka). Ten grams of dry soil were mixed with 2.5 ml of PHE stock solutions (10 mg.ml<sup>-1</sup> in hexane). After complete solvent evaporation, aliquots of 1.1 g recontaminated soil were mixed with 9.9 g of non-recontaminated soil (1:10 dilution) to obtain a final concentration of fresh PHE of 250 mg.kg<sup>-1</sup>.

#### SIP incubations

97 SIP incubations were performed using previously described two-compartment microcosms 98 [12]. Eight ryegrass seedlings (*Lolium multiflorum*, Italian ryegrass, Podium variety, LG 99 seeds, France) were allowed to grow for 21 days until roots reached the bottom of the first 100 compartment containing 30 g of non-recontaminated NM soil and maintained at 80% of the soil water-holding capacity in a growth chamber (22°C /18°C day/night, 80% relative humidity, c.a. 250 µmol photons m<sup>-2</sup> s<sup>-1</sup>, 16 hours photoperiod). A second compartment containing 10 g of <sup>12</sup>C or <sup>13</sup>C PHE-spiked soil was then appended below the first one to allow root colonization in the growth chamber. After 10 days, soil from the second compartment was retrieved in a glass Petri dish. Roots were removed using a brush and tweezers. Aliquots of soil were collected in glass vials for organic extraction and isotopic analysis; the rest was stored in plastic vial and immediately frozen in liquid nitrogen for further nucleic acid extraction. All samples were stored at -80°C until analysis. Non-vegetated ("bare") microcosms were handled in the same way, except no ryegrass was planted in the first compartment. Independent triplicates were performed for the four conditions (i.e. <sup>12</sup>C-bare, <sup>12</sup>C-planted, <sup>13</sup>C-bare and <sup>13</sup>C-planted), for a total of 12 microcosms. Aerial plant biomass was harvested, ground in liquid nitrogen, dried at 60°C and stored at room temperature. To capture the initial PHE concentration, an additional set of second compartments was prepared and sacrificed within 2 hours.

114115116

101

102

103104

105

106

107

108

109

110

111

112

113

#### Organic extraction and phenanthrene quantification

- 117 Lyophilized soil was pulverized in a mixer mill (MM200, Retsch). Soil samples (250 mg)
- were mixed with 1 g activated copper and 2 g anhydrous Na<sub>2</sub>SO<sub>4</sub> and extracted with
- dichloromethane (DCM) using an high pressure and temperature automated extractor Dionex
- ASE350 as described elsewhere [34]. Post-extraction soil residues were recovered for isotopic
- analysis. DCM extracts were analyzed by gas chromatography coupled to mass spectrometry
- 122 (GC-MS) [35].
- The  $^{13}$ C content coming from the added  $^{13}$ C-PHE was expressed as  $^{13}$ C<sub>sample</sub> =  $\alpha \times$  PHE<sub>sample</sub>,
- where PHE<sub>sample</sub> is the PHE content (mg.kg<sup>-1</sup>) and  $\alpha$  the massic proportion of carbon atoms in
- 125 PHE molecule (0.9434). The proportion of <sup>13</sup>C remaining compared to the amount added
- initially was expressed as  $100 \times {}^{13}C_{sample D10}/{}^{13}C_{sample D0}$ . D0 was the initial measure and D10
- after 10 days.

128

129

#### Isotopic analysis

- 130 Isotopic composition of soil samples and post-extraction soil residues, dry DCM extracts and
- aerial plant biomass was determined at the INRA PTEF (Champenoux, France) using
- 132 Elemental analyzer (vario ISOTOPE cube, Elementar, Hanau, Germany) interfaced in line
- with a gas isotope ratio mass spectrometer (IsoPrime 100, Isoprime Ltd, Cheadle, UK).
- Carbon isotopic composition was expressed as  $\delta^{13}$ C (‰) versus Vienna PeeDee Belemnite

- 135 (V-PDB). The ratio of  $^{13}$ C versus  $^{12}$ C in samples was expressed as  $R_{\text{sample}} = R_{\text{V-PDB}} \times (1 + 1)$
- 136  $\delta^{13}C_{sample}/1000$ ), where  $R_{V-PDB} = {}^{13}C_{V-PDB}/{}^{12}C_{V-PDB} = 0.0112375$ . The proportion of  ${}^{13}C$  in a
- sample compared to the amount initially added from <sup>13</sup>C-PHE was calculated as (R<sub>sample D10</sub> –
- $R_{\text{matrix}}$ /( $R_{\text{sample D0}} R_{\text{matrix}}$ ), where  $R_{\text{matrix}}$  is the natural  $^{13}\text{C}/^{12}\text{C}$  ratio of the NM soil, DCM
- extracts or soil residues (average 0.01085).

140141

#### DNA extraction and isopycnic separation

- DNA was extracted from 0.5 g of soil using the Fast DNA Spin Kit for Soil (MP Biomedicals,
- 143 France). Five replicate extractions were pooled for each sample. DNA was quantified using
- the Quant-iT Picogreen dsDNA assay kit (Invitrogen). Isopycnic separation of <sup>12</sup>C- and <sup>13</sup>C-
- labeled DNA ("light" and "heavy" DNA, respectively) was performed as described previously
- 146 [36]. DNA (3.9 μg) was mixed in 5.1 ml tubes with CsCl solution and gradient buffer to a
- final density of 1.725 g.ml<sup>-1</sup>. Ultracentrifugation was performed in a vertical rotor (VTi 65.2,
- Beckman), at 15°C, 176 985g for 40 h (INRA, Champenoux). Thirteen fractions (ca. 400 μl)
- were collected per tube and weighed on a digital balance (precision 10<sup>-4</sup> g) to confirm
- gradient formation. DNA was precipitated with 800 µl polyethylene glycol 6000 (1.6 M) and
- 151 2 μl polyacryl carrier (Euromedex) overnight at room temperature, recovered by
- 152 centrifugation 45 min at 13 000g and washed once with 500 μl 70% (v/v) ethanol. Pellets
- were dried for 2 min in a vacuum concentrator (Centrivap Jouan RC1010, ThermoScientific),
- resuspended in 30 µl molecular-biology grade water (Gibco, Life Technologies) and stored at
- 155 -20°C.

156157

#### **Real-time quantitative PCR (qPCR)**

- Bacterial 16S, Archaeal 16S and fungal 18S rRNA gene copies were quantified using the
- primer pairs 968F/1401R [37], 571F/910R [38] and FF390R/Fung5F [39,40]. qPCR reactions
- 160 (20µl) were performed as described previously [41] on a CFX96 real-time system (BioRad)
- and contained 1X iQ SybrGreen Super Mix (BioRad), 12 µg bovine serum albumin, 0.2 µl
- dimethyl sulfoxide, 40 µg of T4 bacteriophage gene 32 product (MP Biomedicals, France), 1
- 163 µl of template (DNA or 10-fold linearized standard plasmid dilution series from 10<sup>8</sup> to 10<sup>2</sup>
- gene copies.ul<sup>-1</sup>) and 8 pmol of each primer. Reactions were heated at 95°C for 5 min,
- followed by 45 cycles of 30 s at 95°C, 30 s at 56°C for bacterial 16S rRNA, 60°C for archaeal
- 166 16S rRNA or 50°C for 18S rRNA, 30 s at 72°C and 10 s at 82°C to capture the fluorescence
- signal while dissociating primer dimers. Dissociation curves were obtained by heating

reactions from 50 to 95 °C. Fractions identified as containing "heavy" DNA were pooled and quantified by Picogreen assay.

170171

168

169

#### 16S rRNA gene amplicon sequencing and analysis

172 Fragments of 430 bp covering the V3/V4 region of bacterial 16S rRNA genes were amplified from 12 samples of "heavy" DNA recovered from <sup>13</sup>C-SIP and <sup>12</sup>C-control samples and 173 sequenced as described previously [12] with a dual-index paired-end strategy [42]. Amplicons 174 175 were obtained by 28 cycles of PCR using Accuprime Super Mix (Invitrogen) on 1 µl template 176 DNA, purified using the UltraClean-htp 96 Well PCR Clean-Up kit (MOBIO) and quantified 177 by Picogreen assay. An equimolar pool at 10 nM was purified using Nucleospin PCR Clean-178 Up kit (Macherey-Nagel) and sequenced on a single lane of Illumina Miseq PE250 at the 179 Georgia Genomics Facility (Athens, GA, USA). Paired-end reads were trimmed to a 180 minimum Oscore of 20, joined with Pandaseq [43] and filtered for length in the 400-450 bp 181 range with no ambiguous bases. Sequence data were analyzed as described previously [12] in 182 QIIME v1.9 [44] with chimera detection and clustering in Operational Taxonomic Units 183 (OTUs) at 97% using UCHIME and USEARCH v6.1, respectively [45, 46], followed by 184 taxonomy assignment using the RDP classifier [47] with the Greengenes database v13 8 [48]. 185 After removal of chloroplasts and mitochondria OTUs, datasets were rarefied to the lowest number of sequences per sample (9 532 sequences). To identify OTUs significantly enriched 186 in <sup>13</sup>C-SIP incubations compared to <sup>12</sup>C-controls, a subset of data was produced to keep only 187 OTUs (i) represented by at least 5 sequences in each triplicate <sup>13</sup>C-SIP sample and (ii) with an 188 average abundance higher in <sup>13</sup>C-SIP samples compared to <sup>12</sup>C-controls. This subset was then 189 190 log-transformed and compared using Welch's test with Benjamini-Hochberg correction of the 191 p-value performed in R v3.1.3 [49], separately for bare and planted samples.

192193

194195

196

197

198

199

200

201

#### Shotgun metagenomic sequencing and analysis

"Heavy" DNA recovered from the six <sup>13</sup>C-SIP samples was sequenced on 3 lanes of Illumina MiSeq PE300 at the Georgia Genomics Facility (Athens, GA, USA). Adapters removal and quality filtering was performed on raw reads using Trimmomatic [50] with the following parameters: remove adapters (ILLUMINACLIP:adapters.fa:2:30:10), trim 5'- or 3'-bases if phred Qscore <25 (LEADING:25 TRAILING:25), trim read when average quality < 25 (SLIDINGWINDOW:4:25) and discard reads shorter than 100 bp (MINLEN:100). Based on FastQC report, paired and unpaired reads 2 obtained after Trimmomatic were further cropped at 250 and 230 bp, respectively. Unassembled DNA sequences were uploaded to MG-RAST

[51] and compared to the Subsystems database (November 2017) with default parameters. Proportions in functional profiles were compared using Welch's two-sided test with Benjamini-Hochberg correction. Raw metagenomic reads were assembled separately for each sample using SPAdes v3.7.0 [52], with the -meta option and testing kmer sizes of 21, 33, 55, 77,99 and 127. Taxonomic assignment of assembled contigs longer than 5 kb was performed using PhyloPythiaS [53]. Genomic features on assembled contigs were predicted and annotated using prokka v1.12 [54]. Predicted genomic features were also screened for functional genes encoding enzymes potentially involved in aerobic degradation of aromatic compounds using AromaDeg [55], with minimum BLAST homology of 50% and minimum alignment length of 150 bp. Sequences of selected AromaDeg enzyme families were further aligned using MAFFT [56] with the L-INS-i method. Alignments were manually edited in JalView [57] and used for maximum-likelihood phylogenetic tree construction in MEGA v6 [58] after selection of the best protein model. Genes encoding putative carbohydrate active enzymes (CAZymes) were annotated using dbcan [59]. The GhostKOALA annotation server [60] was used to assign KEGG orthologies to genes in the metagenomes and reconstruct metabolic pathways.

217218

219

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

#### Data access

16S rRNA gene amplicon sequences are available at NCBI under BioProject ID PRJNA485442 (BioSamples: SAMN09791490-SAMN09791501). Raw and assembled shotgun metagenomics data are available on MG-RAST under study name RHIZORG\_WGS and RHIZORG ASSEMBLED, respectively.

224

225

#### **Results**

226227

228

229230

231

232

233

234

#### Fate of phenanthrene

The fate of spiked PHE was followed in total soil, organic extracts and soil residues after extraction (**Table 1**). No decrease in <sup>13</sup>C content was observed in total soil over the 10 days period. The proportion of <sup>13</sup>C remaining in DCM extracts (containing PAHs) decreased of ca. 25% in bare soil and only 10% in planted soil (Welch t-test P=0.06). Consistency between AE-IRMS and GC-MS results indicates that most of the DCM-extractable <sup>13</sup>C was <sup>13</sup>C-PHE, without major contributions of <sup>13</sup>C-labeled degradation products. At Day 0, 3.5% of the spiked <sup>13</sup>C was already non-extractable with DCM and recovered in soil residues, suggesting

a rapid sequestration of PHE on soil particles. After 10 days, this non-extractable fraction reached 13% and 5.8% in bare and planted microcosms, respectively (Student t-test, P<0.001). This increase likely reflects the incorporation of <sup>13</sup>C in microbial biomass, and the production of <sup>13</sup>C-labeled hydrophilic intermediates that were not extracted with DCM. Aerial plant biomass was only weakly enriched in <sup>13</sup>C (0.048% <sup>13</sup>C content increase compared to <sup>12</sup>C-controls).

241242

243244

245

246

247248

249250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

235

236

237

238

239240

#### Taxonomic characterization of active PHE degraders

Community gDNA extracted from <sup>12</sup>C-controls and <sup>13</sup>C-SIP incubations was separated by isopycnic centrifugation and fractionated. Quantification of bacterial 16S rRNA genes showed a 6-fold increase in fractions 6-8 (buoyant density 1.713-1.727 g.ml<sup>-1</sup>) of <sup>13</sup>C-SIP bare microcosms compared to <sup>12</sup>C-controls (**Figure S1A**). This increase was only 2-fold in planted microcosms (**Figure S1B**), suggesting a lower incorporation of <sup>13</sup>C within bacterial biomass in the presence of ryegrass. Archaeal and fungal rRNA genes were not enriched in any fractions from <sup>13</sup>C-SIP compared to <sup>12</sup>C-controls (not shown), suggesting that Archaea and Fungi did not metabolize a significant proportion of PHE in the tested conditions. Fractions 6-8 from each sample were pooled (hereafter "heavy DNA") and analyzed by high-throughput sequencing of the 16S rRNA genes. We obtained a total of 405 345 quality-filtered paired-end 16S rDNA sequences from the 12 samples, ranging from 9 532 to 52 544 sequences per sample. Based on rarefied dataset, sequences were clustered in 5 690 OTUs at 97% identity. The taxonomic affiliation of OTUs differed between heavy DNA fractions of <sup>13</sup>C-SIP incubations and <sup>12</sup>C-controls in both bare and planted conditions (**Figure 1A**). We detected 130 and 73 OTUs significantly enriched in <sup>13</sup>C-SIP incubations compared to <sup>12</sup>C-controls in bare and planted conditions, respectively, with 40 OTUs shared between the two conditions (Figure 1B). These active PHE-degrading OTUs were affiliated to Actinobacteria, Alpha-, Beta- and Gammaproteobacteria in both conditions, while Firmicutes were only enriched in heavy DNA fractions from bare soil. Arthrobacter, unknown Sphingomonadaceae and Gammaproteobacteria "PYR10d3" were present at a similar level in both conditions. PHEdegrading Sphingomonas and Alcaligenaceae OTUs were favored in bare soil compared to planted condition. Conversely, Sphingobium and unknown Micrococcaceae were more represented in planted condition. In each condition, one genus dominated the PHE-degrading population. Namely, Sphingomonas dominated in bare soil (43% of total sequences and 61% of total OTUs), while Sphingobium prevailed in planted soil (28% of total sequences, 41% of total OTUs).

269

270

302

Functional profiling of <sup>13</sup>C-labeled metagenomes

271 A total of ~90 million paired-end reads representing 17.6 Gb sequence data were obtained for the six heavy DNA samples from <sup>13</sup>C-SIP incubations (**Table S1**). MG-RAST analysis 272 273 showed that 14 of the 28 functional categories were differentially represented in the two 274 conditions (Figure 2A). The greatest differences between bare and planted conditions were 275 found for the two categories: "Carbohydrates" (8.63% and 10.61% in bare and planted soil 276 <sup>13</sup>C-metagenomes, respectively; q-value=0.016) and "Metabolism of aromatic compounds" (4.66% and 2.45% in bare and planted soil <sup>13</sup>C-metagenomes, respectively; q-value=0.015). 277 278 Within the category "Carbohydrates" (Figure 2B), the abundance of gene sequences affiliated 279 to the metabolism of polysaccharides, monosaccharides, fermentation, di- and 280 oligosaccharides, central carbohydrates and aminosugars, were all significantly overrepresented in the planted soil <sup>13</sup>C-metagenomes. Within the category "Metabolism of 281 282 aromatic compounds" (Figure 2C), sequences matching the anaerobic degradation genes 283 were detected in identical relative abundance between the two conditions. The genes 284 belonging to the three other sub-categories, i.e. peripheral pathways for catabolism of 285 aromatic compounds, metabolism of central aromatic intermediates, and other, were significantly over-represented in bare soil <sup>13</sup>C-metagenomes. 286 287 Genes encoding enzymes potentially involved in aerobic degradation of aromatic compounds 288 were detected in assembled metagenomes using AromaDeg, and their prevalence was 289 calculated relative to the single-copy gene recA (Table 2). A significantly lower prevalence of 290 genes encoding benzoate oxygenases, biphenyl oxygenases, extradiol dioxygenases of the 291 vicinal oxygen chelate superfamily, homoprotocatechuate oxygenases and salicylate oxygenases was found in <sup>13</sup>C-enriched metagenomes from planted condition compared to bare 292 293 soil. We further analyzed selected members of two enzyme families involved in the first steps 294 of PAH degradation, namely the biphenyl/naphthalene family of Rieske non-heme iron 295 oxygenases and extradiol dioxygenases of the vicinal oxygen chelate family (EXDO I). The 296 biphenyl/naphthalene family comprises most of the dioxygenases reported to date to activate 297 PAHs for further aerobic degradation, whereas the EXDO I family comprises enzymes that 298 fission the ring of pre-activated mono- or polyaromatic derivatives [55]. We notably detected 299 66 and 20 ORFs encoding oxygenases of the biphenyl/naphthalene family from 300 Proteobacteria (Clusters XXIV and XXVI) or Actinobacteria (Clusters I, II and V), respectively, as well as 46 ORFS encoding proteobacterial EXDOs preferring bicyclic 301 substrates related to dihydroxynaphthalene dioxygenases (Cluster XII). Phylogenetic analysis

of biphenyl/naphthalene oxygenases from *Proteobacteria* (**Figure 3A**) revealed that the vast majority (56/66) was closely related to known sequences from *Sphingomonas*, *Sphingobium* and *Novosphingobium* with equal distributions between bare and planted soil. Few additional sequences were related to *Cycloclasticus*, *Burkholderia* or *Acidovorax* spp. No sequences were affiliated to *Pseudomonas*. Within *Actinobacteria* (**Figure 3B**), 12 biphenyl/naphthalene dioxygenases were closely related to sequences from *Arthrobacter phenanthrenivorans* (Cluster II) and *Arthrobacter keyseri* (Cluster V). In planted conditions only, 4 additional sequences related to *Mycobacterium* and *Terrabacter* were detected in Cluster V, as well as 2 more divergent sequences. In Cluster XII of the EXDO I family (**Figure 4**), most detected sequences grouped with known proteins from *Sphingomonas*, *Sphingobium* and *Novosphingobium*. Among these, a relatively divergent clade of 20 sequences emerged, with more detected members in bare soil compared to planted conditions. Finally, few additional EXDO I Cluster XII sequences were related to *Sphingopyxis*, *Pseudomonas*, *Acidovorax* or *Burkholderia*.

#### Reconstruction of phenanthrene degradation pathways

Combining results of the above AromaDeg analysis and the GhostKOALA annotation pipeline, we reconstructed both the O-phthalate/protocatechuate pathway leading to the 3oxoadipate, as well the naphthalene pathway leading to salicylate (Figure 5). GhostKOALA failed to recognize genes involved in the first steps of PHE degradation that were detected with AromaDeg, likely because the KEGG database only contains nahAc and nidA genes from *Pseudomonas* and *Mycobacterium*, respectively. Salicylate could be converted to gentisate or catechol, being further degraded through ortho- and meta-cleavage pathways leading to intermediates of the TCA cycle. All genes involved in these pathways were present in both conditions. Sequences affiliated to Sphingomonadales dominated the early steps of degradation leading to 1-hydroxy-2-naphthaldehyde and pyruvate, as well as later reactions converting naphthalene-1,2-diol to salicylaldehyde and pyruvate. Genes involved in the downstream conversion to gentisate were mainly detected from Betaproteobacteria (including Alcaligenaceae) and Sphingomonadales. Genes annotated in the meta- and ortho-cleavage pathways for catechol utilization were taxonomically more diverse, with sequences affiliated to Alpha-, Beta- and Gammaproteobacteria, Actinobacteria and Firmicutes. Overall, Actinobacteria and Firmicutes had higher contributions to the phthalate and protocatechuate pathway than the naphthalene and salicylate pathway.

#### Focus on Sphingomonas and Sphingobium metagenomes

337

338 We further focused on the two dominant PHE-degrading populations of Sphingomonas and 339 Sphingobium identified in bare and planted soil through SIP (Figure 1B). After taxonomic 340 affiliation of contigs larger than 5 kb, the best metagenome assemblies were obtained for 341 Sphingomonas population in bare soil and Sphingobium population in planted soil, 342 respectively, with maximum contig size larger than 1 Mb (Supplementary Table S2). The 343 Sphingobium metagenome from bare soil was more fragmented (average size 71 kb, maximum size 278 kb). This likely reflects a lower coverage of the Sphingobium-affiliated 344 sequences in <sup>13</sup>C-enriched metagenomes and corroborates the lower success of this PHE-345 346 degrading genus in bare soil compared to planted conditions. The single-copy RecA protein 347 of <sup>13</sup>C-enriched metagenomes of *Sphingomonas* was identical in the 3 bare soil microcosms 348 and had its best blastp hit (91% identity) against Sphingomonas sp. MM-1 [61], while that of 349 Sphingobium was identical in the 3 planted microcosms and had its best blastp hit (100%) 350 identity) against Sphingobium herbicidovorans NBRC 16415 [62] and Sphingobium sp. 351 MI1205 [63]. Functional gene annotation revealed a large arsenal of dioxygenases and 352 monooxygenases with some that could have potential activity on aromatic compounds in 353 Sphingomonas metagenomes from bare soil (Supplementary Table S3), often grouped in 354 genomic regions. The potential for aromatic compounds degradation was much more 355 restricted in Sphingobium metagenomes from planted conditions. We then hypothesized that 356 the greater success of Sphingobium PHE-degraders compared to Sphingomonas in planted 357 conditions might not be directly due to aromatic compound catabolism, but rather to a more 358 efficient use of plant-derived carbon sources. To test this hypothesis, we screened the 359 metagenomes for genes encoding carbohydrate active enzymes (CAZymes) in the groups 360 carbohydrate esterases (CE), glycoside hydrolases (GH) and polysaccharide lyases (PL) 361 (**Table 3**). Similar numbers of CAZymes were detected in *Sphingomonas* (36 to 47 total 362 CAZymes) and Sphingobium (46 to 48 total CAZymes). However, a larger diversity of CAZy 363 families was found in Sphingobium (27 different families) than in Sphingomonas (19 to 21) 364 metagenomes. Families detected only in Sphingobium metagenomes include enzymes 365 potentially involved in plant cell wall breakdown, including the degradation of xylan (CE6, 366 CE7, GH10, GH115, GH43 12, GH67), pectin (PL1 2) and other cell wall compounds 367 (GH16), as well as in the use of disaccharides (maltose, trehalose) that are found in root 368 exudates (GH65). Furthermore, the prevalence of three additional families also potentially 369 involved in plant cell wall breakdown was higher in Sphingobium metagenomes compared to 370 Sphingomonas (GH13, GH3, PL22).

#### 371

372

373

374

## 375376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395396

#### **Discussion**

Phenanthrene is often highly concentrated in PAH-contaminated environments and is a model for research on PAH catabolism [1]. We used DNA-SIP to investigate the diversity and metabolic potential of microorganisms involved in phenanthrene degradation in historically contaminated soil. We further assessed the influence of ryegrass, a plant commonly used in phytoremediation studies on PAH-contaminated soils [2,11,64]. To our knowledge, this study is the first to use DNA-SIP combined to metagenomics to assess the influence of plants on bacteria actively involved in phenanthrene degradation in polluted soils. Here, we showed a decreased dissipation rate of phenanthrene in the ryegrass rhizosphere, corroborating previous results [12,14]. The rhizosphere environment contains higher nutrient concentration than the surrounding bulk soil due to root exudates, comprising an array of organic compounds including carbohydrates, amino acids, proteins, flavonoids, aliphatic acids, organic acids and fatty acids [10,32]. Excessive nutrient availability can inhibit the biodegradation of pollutants [65,66,67]. Thus, rhizospheric bacteria may preferentially use labile carbon sources from exudates, leading to slower phenanthrene degradation. The phenanthrene degradation rates we observed (20% and 10% in bare and planted soil in 10 days, respectively) were lower than previous reports, e.g. 60 to 70% PHE degradation in 9 days [21,24] or 68% in 12 days [14]. This discrepancy might be linked to the more realistic conditions of the present experimental setup within a genuine plant rhizosphere, allowing constant input of rhizodeposits during the 10-days time course. The limited dissipation of <sup>13</sup>C-PHE after 10 days minimized the risk of cross-feeding, ensuring that <sup>13</sup>C-labeled microorganisms are the primary degraders of phenanthrene in the NM soil.

398399

400401

402

403

397

The active <sup>13</sup>C-PHE-degrading OTUs belonged to similar taxa in bare and planted soil but the presence of plants modified their relative abundance. <sup>13</sup>C-enriched OTUs affiliated to *Arthrobacter*, unclassified *Sphingomonadaceae* and *Gammaproteobacteria* "PYR10d3" were present at a similar level in both conditions. Members of the *Actinobacteria* and *Sphingomonadaceae* are considered potent PAH-degraders in soil and sediments

[23,68,69,70]. Representatives of the Arthrobacter genus are able to degrade many organic pollutants [71]. Using DNA-SIP, some species were previously shown as the dominant phenanthrene degraders in soil supplemented with root exudates [14] and in activated sludge [26]. Interestingly, an Arthrobacter oxydans strain was isolated from the same soil used in the present study, for its ability to degrade phenanthrene [72]. The Pyr10d3 candidate order is a separate branch in Gammaproteobacteria first identified in a SIP experiment using <sup>13</sup>Cpyrene as substrate [29] and its abundance was increased considerably with the increase of petroleum-hydrocarbon contamination of a soil located close to petrochemical plant [73]. Firmicutes affiliated to Paenibacillaceae were only active in bare soil while in a previous study it was found as phenanthrene degraders when the soil was supplemented with root exudates [14]. Paenibacillus spp. were previously enriched from petroleum-hydrocarboncontaminated sediment and salt marsh rhizosphere using either naphthalene or phenanthrene as the sole carbon source [74]. The greatest influence of ryegrass on active PHE-degraders was found for Sphingomonas and Sphingobium-related <sup>13</sup>C-enriched OTUs, which dominated in bare and planted soil, respectively. Sphingomonads, members of the family Sphingomonadaceae within the Alphaproteobacteria, utilize both substituted and unsubstituted mono- and poly-aromatic hydrocarbons up to 5-rings [27] and have been widely identified as phenanthrene degraders [24,75]. PAH-degrading sphingomonads are common Gram-negative, aerobic, chemoheterotroph bacteria adapted to oligotrophic environments [75]. They evolved original strategies to enhance PAH bioavailability, e.g. hydrophobic and negatively charged cell surface, production of a specific sphingoglycolipid, formation of biofilms due to sphingans exopolysaccharide production, presence of high-affinity uptake system, and chemotactic response towards PAH [75]. Sphingomonas spp. are commonly encountered PAH-contaminated environments as phenanthrene degraders [11,22,23,68,75,76,77,78,79] and the amount of phenanthrene available in soils influences the diversity of Sphingomonas [18]. Similarly, representatives of Sphingobium can biodegrade environmental pollutants such as PAHs [80,81,82,83]. In bare soil, <sup>13</sup>C-enriched Sphingomonas OTUs coincided with a greater proportion of active Alcaligenaceae OTUs, while in rhizospheric soil Sphingobium OTUs were associated with a greater proportion of active unknown Micrococcaceae. Alcaligenes representatives were previously identified during creosote-contaminated soil bioremediation [84], and many studies showed PAH degradation by various Alcaligenes isolates [85,86,87]. Micrococcaceae are well-known PAH degraders [88,89] and their abundance was favoured in planted compared to bare aged-PAH contaminated soil [33].

404

405

406

407

408

409

410

411

412

413

414

415

416

417418

419

420

421

422

423

424

425

426

427

428

429

430431

432

433

434

435

436

437

439 To date, the impact of plant rhizosphere on the functional diversity was mostly assessed in 440 pristine soils [90,91]. Our study is one of the first to highlight the differences in metagenomes 441 of PAH-degrading bacteria from bulk and rhizospheric soil. We recovered enough <sup>13</sup>C-labeled 442 DNA for direct shotgun metagenomic sequencing, overcoming the potential bias of multiple 443 displacement amplification often used in DNA-SIP approaches [92]. The <sup>13</sup>C-enriched 444 metagenomes from planted soil showed that ryegrass rhizosphere selected for an active 445 population with specific functions compared to bare soil, i.e. a lower proportion of genes 446 involved in aromatic compounds utilization together with a higher and diversified capability 447 for carbohydrate degradation. This is reminiscent of the higher transcription of genes related 448 to carbon and amino-acid uptake and utilization shown in the willow rhizosphere [93]. 449 Moreover, the AromaDeg analysis confirmed the lower proportion of genes potentially 450 involved in the first steps of aerobic degradation of aromatic compounds in planted soil 451 metagenomes. Reconstruction of complete catabolic pathways (Figure 5) showed that several 452 routes are used to mineralize phenanthrene both in bare and planted soil, since genes assigned 453 to the o-phthalate/protocatechuate, gentisate and catechol ortho and meta-cleavage pathways 454 were present. Sphingomonadales were the major taxa performing the first steps of 455 phenanthrene degradation. They appear to degrade phenanthrene preferentially through lower 456 meta-cleavage pathway, as it was shown for Sphingobium chungbukense [94]. These results 457 suggest that autochthonous Sphingomonadales could have a critical role to initiate in-situ 458 PAH remediation in historically polluted soils, by increasing PAH bioavailability and opening

463464465

466

467

468

469

470

471

459

460

461

462

438

The dominant PHE-degraders *Sphingomonas* and *Sphingobium* OTUs were active both in bare and planted soils but in various proportions, leading to a lower phenanthrene degradation rate in the rhizosphere where *Sphingobium* dominated. Previous studies did not evidence a specificity of *Sphingomonas* and *Sphingobium* spp. for environments with low and high nutrient levels, respectively. For example, Vinas et al. [84] detected *Sphingomonas* in soil bioremediation treatments both with and without nutrient amendment. Recently, comparative genomics of *Novosphingobium* strains showed that phylogenetic relationships were less likely

new substrate niches for other degraders. Our data further show that active PHE-degrading

communities act as consortia, whereby the complete mineralization of phenanthrene is

achieved through combined activity of taxonomically diverse co-occurring bacteria using

various degradation pathways. This strategy might limit the competition for substrates

between different degrading populations and overall increase PAH dissipation efficiency.

to describe functional genotype similarity, in term of metabolic traits, than the habitat from which they were isolated [95]. Thus, catabolic differences among Sphingomonads appear strain-specific. Some Sphingomonad strains can simultaneously take up mixed substrates rather than being specialists in the degradation of aromatic compounds, because they can assimilate various mono-, di-, and polysaccharides [75,96]. The greater success of <sup>13</sup>C-enriched *Sphingobium* OTUs in planted soil, together with a larger diversity of CAZymes than the *Sphingomonas* metagenome from bare soil, suggests a similar behaviour. In the conditions tested, PHE-degrading *Sphingobium* representatives likely took advantage of labile carbon compounds from root exudates, outcompeting the less versatile PHE-degrading *Sphingomonas*.

#### Conclusion

In summary, this study identified the metabolically active phenanthrene degraders in an aged-polluted soil, and showed their taxonomic diversity among Alpha-, Beta- and Gammaproteobacteria, Actinobacteria and Firmicutes that could act as a consortium. Members of the Sphingomonadales were the dominant PHE-degraders identified through DNA-SIP, and the main actors of the first steps in the degradation pathways. Hence, they likely play a crucial role to initiate in-situ PAH remediation. Results further indicated that the presence of plants impacted the active PHE-degrading community and induced a slower PHE dissipation. In particular, plants induced a drastic shift in the taxonomic composition of PHE-degrading *Sphingomonadales*, favouring the growth of *Sphingobium* populations with a more diverse repertoire of carbohydrate-active enzymes potentially targeting plant root material, to the detriment of less versatile *Sphingomonas* representatives that prevailed in bare soil.

#### Acknowledgement.

- 498 This study was part of the RhizOrg project funded by the ANR (Agence Nationale de la
- Recherche, ANR-13-JSV7-0007 01 project allocate to A. Cébron). The authors thank Dr. S.
- 500 Uroz (INRA Champenoux) for giving them access to the ultracentrifuge equipment, Dr. E.
- Morin (INRA Champenoux) for initial discussions on metagenome assembly, and the ABiMS
- 502 platform (Roscoff) where metagenomic analyses were performed.

#### Conflict of interest

The authors declare that they have no conflict of interet.

506

507508

#### References

- 509 1. Ghosal D, Ghosh S, Dutta TK, Ahn Y. Current state of knowledge in microbial
- degradation of polycyclic aromatic hydrocarbons (PAHs): a review. Front Microbiol.
- 511 2016; **7**:1369.
- 512 2. Khan S, Afzal M, Iqbal S, Khan QM. Plant-bacteria partnerships for the remediation of
- 513 hydrocarbon contaminated soils. Chemosphere. 2013; **90**:1317-1332.
- 3. Doyle E, Muckian L, Hickey AM, Clipson N. Microbial PAH degradation. Adv Appl
- 515 Microbiol. 2008; **65**:27-66.
- 4. Reilley KA, Banks MK, Schwab AP.Dissipation of polycyclic aromatic hydrocarbons in
- the rhizosphere. J Environ Qual. 1996; **25**:212-219.
- 5. Binet P, Portal JM, Leyval C. Dissipation of 3–6-ring polycyclic aromatic hydrocarbons in
- the rhizosphere of ryegrass. Soil Biol Biochem. 2000; **32**:2011-2017.
- 6. Chaudhry Q, Blom-Zandstra M, Gupta SK, Joner E. Utilising the synergy between plants
- and rhizosphere microorganisms to enhance breakdown of organic pollutants in the
- environment. Environ Sci Pollut Res. 2005; 12:34-48.
- 7. El Amrani A, Dumas AS, Wick LY, Yergeau E, Berthomé R. "Omics" insights into PAH
- degradation toward improved green remediation biotechnologies. Environ Sci Technol.
- 525 2015; **49**:11281-11291.
- 526 8. Rentz JA, Alvarez PJ, Schnoor JL. Repression of Pseudomonas putida phenanthrene
- degrading activity by plant root extracts and exudates. Environ Microbiol. 2004; 6:574-
- 528 583.
- 9. Phillips LA, Greer CW, Farrell RE, Germida JJ. Plant root exudates impact the
- 530 hydrocarbon degradation potential of a weathered-hydrocarbon contaminated soil. Appl
- 531 Soil Ecol. 2012; **52**:56-64.
- 532 10. Berg G, Smalla K. Plant species and soil type cooperatively shape the structure and
- function of microbial communities in the rhizosphere. FEMS Microbiol Ecol. 2009; **68**:1-
- 534 13.
- 535 11. Guo M, Gong Z, Miao R, Rookes J, Cahill D, Zhuang J. Microbial mechanisms
- controlling the rhizosphere effect of ryegrass on degradation of polycyclic aromatic
- 537 hydrocarbons in an aged-contaminated agricultural soil. Soil Biol Biochem. 2017;
- **113**:130-142.

- 539 12. Thomas F, Cébron A. Short-term rhizosphere effect on available carbon sources
- phenanthrene degradation and active microbiome in an aged-contaminated industrial soil.
- 541 Front Microbiol. 2016; **7**:92.
- 13. Olson PE, Castro A, Joern M, DuTeau NM, Pilon-Smits EA, Reardon KF, Comparison of
- plant families in a greenhouse phytoremediation study on an aged polycyclic aromatic
- hydrocarbon–contaminated soil. J Environ Qual. 2007; **36**:1461-1469.
- 545 14. Cébron A, Louvel B, Faure P, France Lanord C, Chen Y, Murrell JC, Leyval C. Root
- exudates modify bacterial diversity of phenanthrene degraders in PAH polluted soil but not
- phenanthrene degradation rates. Environ Microbiol. 2011; **13**:722-736.
- 548 15. Sul WJ, Park J, Quensen JF, Rodrigues JL, Seliger L, Tsoi TV et al. DNA-stable isotope
- probing integrated with metagenomics for retrieval of biphenyl dioxygenase genes from
- polychlorinated biphenyl-contaminated river sediment. Appl Environ Microbiol. 2009;
- **75**:5501-5506.
- 16. Kim SJ, Park SJ, Cha IT, Min D, Kim JS, Chung WH et al. Metabolic versatility of
- toluene degrading iron reducing bacteria in tidal flat sediment characterized by stable
- isotope probing based metagenomic analysis. Environ Microbiol. 2014; **16**:189-204.
- 17. Jeon CO, Park W, Padmanabhan P, DeRito C, Snape JR, Madsen EL. Discovery of a
- bacterium with distinctive dioxygenase that is responsible for in situ biodegradation in
- 557 contaminated sediment. PNAS. 2003; **100**:13591-13596.
- 18. Singleton DR, Powell SN, Sangaiah R, Gold A, Ball LM, Aitken MD. Stable-isotope
- probing of bacteria capable of degrading salicylate naphthalene or phenanthrene in a
- bioreactor treating contaminated soil. Appl Environ Microbiol. 2005; 71:1202-1209.
- 19. Padmanabhan P, Padmanabhan S, DeRito C, Gray A, Gannon D, Snape JR et al.
- Respiration of 13C-labeled substrates added to soil in the field and subsequent 16S rRNA
- gene analysis of 13C-labeled soil DNA. Appl Environ Microbiol. 2003; **69**:1614-1622.
- 20. Uhlik O, Wald J, Strejcek M, Musilova L, Ridl J, Hroudova M et al. Identification of
- bacteria utilizing biphenyl benzoate and naphthalene in long-term contaminated soil. PloS
- 566 One. 2012; **7**:e40653.
- 567 21. Jiang L, Song M, Luo C, Zhang D, Zhang G. Novel phenanthrene-degrading bacteria
- identified by DNA-stable isotope probing. PloS One. 2015; **10**:e0130846.

- 569 22. Martin F, Torelli S, Le Paslier D, Barbance A, Martin-Laurent F, Bru D et al.
- Betaproteobacteria dominance and diversity shifts in the bacterial community of a PAH-
- contaminated soil exposed to phenanthrene. Environ Pollut. 2012; **162**:345-353.
- 572 23. Regonne RK, Martin F, Mbawala A, Ngassoum MB, Jouanneau Y. Identification of soil
- bacteria able to degrade phenanthrene bound to a hydrophobic sorbent in situ. Environ
- 574 Pollut. 2013; **180**:145-151.
- 575 24. Song M, Jiang L, Zhang D, Luo C, Wang Y, Yu Z, Zhang G. Bacteria capable of
- degrading anthracene phenanthrene and fluoranthene as revealed by DNA based stable-
- isotope probing in a forest soil. J Hazard Mat. 2016; **308**:50-57.
- 578 25. Crampon M, Cébron A, Portet-Koltalo F, Uroz S, Le Derf F, Bodilis J. Low effect of
- phenanthrene bioaccessibility on its biodegradation in diffusely contaminated soil. Environ
- 580 Pollut. 2017; **225**:663-673.
- 581 26. Li J, Zhang D, Song M, Jiang L, Wang Y, Luo C, Zhang G. Novel bacteria capable of
- degrading phenanthrene in activated sludge revealed by stable-isotope probing coupled
- with high-throughput sequencing. Biodeg. 2017; **28**:423-436.
- 584 27. Jones MD, Crandell DW, Singleton DR, Aitken MD. Stable isotope probing of the
- polycyclic aromatic hydrocarbon degrading bacterial guild in a contaminated soil. Environ
- 586 Microbiol. 2011; **13**:2623-2632.
- 28. Zhang S, Wang Q, Xie S. Stable isotope probing identifies anthracene degraders under
- methanogenic conditions. Biodeg. 2012; **23**:221-230.
- 589 29. Singleton DR, Sangaiah R, Gold A, Ball LM, Aitken MD. Identification and
- quantification of uncultivated Proteobacteria associated with pyrene degradation in a
- 591 bioreactor treating PAH contaminated soil. Environ Microbiol. 2006; 8:1736-1745.
- 592 30. Jones MD, Singleton DR, Carstensen DP, Powell SN, Swanson JS, Pfaender FK, Aitken
- MD. Effect of incubation conditions on the enrichment of pyrene-degrading bacteria
- identified by stable-isotope probing in an aged PAH-contaminated soil. Microb Ecol. 2008;
- **56**:341-349.
- 596 31. Song M, Luo C, Jiang L, Zhang D, Wang Y, Zhang G. Identification of benzo [a] pyrene
- 597 (BaP)-metabolizing bacteria in forest soils using DNA-based stable-isotope probing. Appl
- 598 Environ Microbiol. 2015; **81**:7368-7376.
- 599 32. el Zahar Haichar F, Marol C, Berge O, Rangel-Castro JI, Prosser JI, Balesdent J,
- Achouak W. Plant host habitat and root exudates shape soil bacterial community structure.
- 601 ISME J. 2008; **2**:1221.

- 33. Cébron A, Beguiristain T, Faure P, Norini M-P, Masfaraud J-F, Leyval C. Influence of
- vegetation on the in situ bacterial community and polycyclic aromatic hydrocarbon (PAH)
- degraders in aged PAH-contaminated or thermal-desorption-treated soil. Appl Environ
- 605 Microbiol. 2009; **75**: 6322–30.
- 606 34. Biache C, Mansuy-Huault L, Faure P, Munier-Lamy C, Leyval C. Effects of thermal
- desorption on the composition of two coking plant soils: Impact on solvent extractable
- organic compounds and metal bioavailability. Environ Pollut. 2008; **156**: 671-677.
- 35. Biache C, Ouali S, Cébron A, Lorgeoux C, Colombano S, Faure P. Bioremediation of
- 610 PAH-contamined soils: consequences on formation and degradation of polar-polycyclic
- aromatic compounds and microbial community abundance. J Hazard Mat. 2017; **329**:1-10.
- 36. Dunford E, Neufeld JD. DNA stable-isotope probing (DNA-SIP). J Vis Exp. 2010; 2: 1-
- 613 6.
- 614 37. Felske A, Akkermans ADL, De Vos WM. Quantification of 16S rRNAs in complex
- bacterial communities by multiple competitive reverse transcription-PCR in temperature
- gradient gel electrophoresis fingerprints. Appl Environ Microbiol. 1998; **64**: 4581-4587.
- 38. Baker GC, Smith JJ, Cowan DA. Review and re-analysis of domain-specific 16S primers.
- 618 J Microbiol Meth. 2003; **55**:541-555.
- 39. Smit E, Smit E, Leeflang P, Leeflang P, Glandorf B, Glandorf B et al. Analysis of fungal
- diversity in the wheat rhizosphere by sequencing of cloned PCR-Amplied genes encoding
- 621 18S rRNA and temperature gradient gel electrophoresis. Society. 1999; **65**: 2614–2621.
- 40. Vainio EJ, Hantula J. Direct analysis of wood-inhabiting fungi using denaturing gradient
- 623 gel electrophoresis of amplified ribosomal DNA. Mycol Res. 2000; **104**: 927–936.
- 41. Cébron A, Norini M-P, Beguiristain T, Leyval C. Real-Time PCR quantification of PAH-
- ring hydroxylating dioxygenase (PAH-RHDα) genes from Gram positive and Gram
- 626 negative bacteria in soil and sediment samples. J Microbiol Methods. 2008; 73: 148–59
- 42. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-
- 628 index sequencing strategy and curation pipeline for analyzing amplicon sequence data on
- the MiSeq Illumina sequencing platform. Appl Environ Microbiol. 2013; **79**: 5112–20.
- 43. Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD. PANDAseq:
- paired-end assembler for Illumina sequences. BMC Bioinformatics. 2012; **13**: 31.
- 44. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al.
- 633 OIIME allows analysis of high-throughput community sequencing data. Nat Methods.
- 634 2010; **7**: 335–6.

- 45. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics.
- 636 2010; **26**: 2460–1.
- 46. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity
- and speed of chimera detection. Bioinformatics. 2011; **27**: 2194–2200.
- 639 47. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid
- assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol.
- 641 2007; **73**: 5261–7.
- 48. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. An
- improved Greengenes taxonomy with explicit ranks for ecological and evolutionary
- analyses of bacteria and archaea. ISME J. 2012; **6**: 610–8.
- 49. R Core TeamR: A Language and Environment for Statistical Computing. 2013. Vienna,
- 646 Austria.
- 50. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence
- data. Bioinformatics. 2014; **30**: 2114–2120.
- 649 51. Meyer F, Paarmann D, D'Souza M, etal. The metagenomics RAST server—a public
- resource for the automatic phylo- genetic and functional analysis of metagenomes. BMC
- 651 Bioinformatics. 2008; **9**: 386.
- 52. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes:
- A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J
- 654 Comput Biol. 2012; **19**: 455–477.
- 655 53. Patil KR, Roune L, McHardy AC. The phyloPythiaS web server for taxonomic
- assignment of metagenome sequences. PLoS One. 2012; 7: e38581...
- 54. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014. 30:
- 658 2068-2069.
- 55. Duarte M Jauregui R Vilchez-Vargas R Junca H Pieper DH. AromaDeg a novel database
- for phylogenomics of aerobic bacterial degradation of aromatics. Database. 2014; 1–12.
- 661 56. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence
- alignment interactive sequence choice and visualization. Brief Bioinform. 2017; 1–7.
- 57. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2-A
- multiple sequence alignment editor and analysis workbench. Bioinformatics 2009; 25:
- 665 1189–1191.
- 58. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary
- genetics analysis version 6.0. Mol Biol Evol. 2013; **30**: 2725–2729.

- 59. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. DbCAN: A web resource for automated
- carbohydrate-active enzyme annotation. Nucleic Acids Res. 2012; **40**: 445–451.
- 670 60. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for
- functional characterization of genome and metagenome sequences. J Mol Biol. 2016; **428**:
- 672 726–731.
- 673 61. Tabata M, Ohtsubo Y, Ohhata S, Tsuda M, Nagata Y. Complete genome sequence of the
- 674 γ-hexachlorocyclohexane-degrading bacterium Sphingomonas spstrain MM-1. Genome
- Announcements. 2013; 1:e00247-13.
- 676 62. Zipper C, Nickel K, Angst W, Kohler HP. Complete microbial degradation of both
- enantiomers of the chiral herbicide mecoprop [(RS)-2-(4-chloro-2-methylphenoxy)
- propionic acid] in an enantioselective manner by Sphingomonas herbicidovorans sp. nov.
- Appl Environ Microbiol. 1996; **62**: 4318-4322.
- 680 63. Tabata M, Ohhata S, Nikawadori Y, Sato , TKishida K, Ohtsubo Y et al. Complete
- 681 genome sequence of a γ-hexachlorocyclohexane-degrading bacterium Sphingobium
- spstrain MI1205. Genome Announcements. 2016; 4:e00246-16.
- 683 64. Rezek J, Mackova M, Zadrazil F, Macek T. The effect of ryegrass (Lolium perenne) on
- decrease of PAH content in long term contaminated soil. Chemosphere. 2008; **70:**1603-
- 685 1608.
- 686 65. Chaillan F, Chaineau CH, Point V, Saliot A, Oudot J. Factors inhibiting bioremediation
- of soil contaminated with weathered oils and drill cuttings. Environ Pollut. 2006; **144:**255-
- 688 265.
- 689 66. Carmichael LM, Pfaender FK. The effect of inorganic and organic supplements on the
- 690 microbial degradation of phenanthrene and pyrene in soils. Biodeg. 1997; 8:1-13.
- 691 67. Chaineau CH, Rougeux G, Yepremian C, Oudot J. Effects of nutrient concentration on
- the biodegradation of crude oil and associated microbial populations in the soil. Soil Biol
- 693 Biochem. 2005; **37:**1490-1497.
- 694 68. Leys NM, Ryngaert A, Bastiaens L, Verstraete W, Top EM, Springael D. Occurrence and
- 695 phylogenetic diversity of *Sphingomonas* strains in soils contaminated with polycyclic
- aromatic hydrocarbons. Appl Environ Microbiol. 2004; **70:**1944-1955.
- 697 69. Leys NM, Bastiaens L, Verstraete W, Springael D. Influence of the
- 698 carbon/nitrogen/phosphorus ratio on polycyclic aromatic hydrocarbon degradation by
- Mycobacterium and Sphingomonas in soil. Appl Microbiol Biotechnol. 2005; **66:**726-736.

- 700 70. Alonso-Gutiérrez J, Figueras A, Albaigés J, Jiménez N, Vinas M, Solanas AM, Novoa B.
- Bacterial communities from shoreline environments (Costa da Morte Northwestern Spain)
- affected by the Prestige oil spill. Appl Environ Microbiol. 2009; 75:3407-3418.
- 703 71. Leigh MB, Pellizari VH, Uhlík O, Sutka R, Rodrigues J, Ostrom NE et al. Biphenyl-
- utilizing bacteria and their functional genes in a pine root zone contaminated with
- polychlorinated biphenyls (PCBs). ISME J. 2007; 1:134.
- 706 72. Thion C, Cébron A, Beguiristain T, Leyval C. PAH biotransformation and sorption by
- Fusarium solani and Arthrobacter oxydans isolated from a polluted soil in axenic cultures
- and mixed co-cultures. Int Biodeter Biodeg. 2012; **68:**28-35.
- 709 73. Tardif S, Yergeau É, Tremblay J, Legendre P, Whyte LG, Greer CW. The willow
- microbiome is influenced by soil petroleum-hydrocarbon concentration with plant
- 711 compartment-specific effects. Front Microbiol. 2016; 7:1363.
- 712 74. Daane LL, Harjono I, Barns SM, Launen LA, Palleron NJ, Häggblom MM. PAH-
- degradation by Paenibacillus sppand description of Paenibacillus naphthalenovorans
- spnova naphthalene-degrading bacterium from the rhizosphere of salt marsh plants. Int J
- 715 Syst Evol Microbiol. 2002; **52:**131-139.
- 716 75. Waigi MG, Kang F, Goikavi C, Ling W, Gao Y. Phenanthrene biodegradation by
- sphingomonads and its application in the contaminated soils and sediments: a review. Int
- 718 Biodeter Biodeg. 2015; **104:**333-349.
- 719 76. Daane LL, Harjono I, Zylstra GJ, Häggblom MM. Isolation and characterization of
- polycyclic aromatic hydrocarbon-degrading bacteria associated with the rhizosphere of salt
- marsh plants. Appl Environ Microbiol. 2001; **67:**2683-2691.
- 722 77. Johnsen AR, Winding A, Karlson U, Roslev P. Linking of microorganisms to
- phenanthrene metabolism in soil by analysis of 13C-labeled cell lipids. Appl Environ
- 724 Microbiol. 2002; **68:**6106-6113.
- 725 78. Meyer S, Moser R, Neef A, Stahl U, Kämpfer P. Differential detection of key enzymes of
- polyaromatic-hydrocarbon-degrading bacteria using PCR and gene probes. Microbiol.
- 727 1999; **145:**1731-1741.
- 728 79. Ding GC, Heuer H, Smalla K. Dynamics of bacterial communities in two unpolluted soils
- after spiking with phenanthrene: soil type specific and common responders. Front
- 730 Microbiol. 2012; **3:**290.
- 731 80. Kanaly RA, Harayama S. Biodegradation of high-molecular-weight polycyclic aromatic
- hydrocarbons by bacteria. J Bacteriol. 2000; **182**:2059-2067.

- 733 81. Kanaly RA, Harayama S. Advances in the field of high molecular weight polycyclic
- aromatic hydrocarbon biodegradation by bacteria. Microb Biotechnol. 2010; **3:**136-164.
- 735 82. Stolz A. Molecular characteristics of xenobiotic-degrading sphingomonads. Appl
- 736 Microbiol Biotechnol. 2009; **81:**793-811.
- 737 83. Maeda AH, Kunihiro M, Ozeki Y, Nogi Y, Kanaly RA. Sphingobium barthaii spnova
- high molecular weight polycyclic aromatic hydrocarbon-degrading bacterium isolated
- from cattle pasture soil. Int J Syst Evolut Microbiol. 2015; **65:**2919-2924.
- 740 84. Vinas M, Sabaté J, Espuny MJ, Solanas AM. Bacterial community dynamics and
- 741 polycyclic aromatic hydrocarbon degradation during bioremediation of heavily creosote-
- contaminated soil. Appl Environ Microbiol. 2005; **71:**7008-7018.
- 743 85. Weissenfels WD, Beyer M, Klein J. Degradation of phenanthrene fluorene and
- fluoranthene by pure bacterial cultures. Appl Microbiol Biotechnol. 1990; **32:**479-484.
- 745 86. Lal B, Khanna S. Degradation of crude oil by Acinetobacter calcoaceticus and
- 746 Alcaligenes odorans. J Appl Bacteriol. 1996; 81:355-362.
- 747 87. Samanta SK, Singh OV, Jain RK. Polycyclic aromatic hydrocarbons: environmental
- pollution and bioremediation. TRENDS Biotechnol. 2002; **20:**243-248.
- 749 88. Kästner M, Breuer-Jammali M, Mahro B. Enumeration and characterization of the soil
- 750 microflora from hydrocarbon-contaminated soil sites able to mineralize polycyclic
- aromatic hydrocarbons (PAH). Appl Microbiol Biotechnol. 1994; **41:**267-273.
- 752 89. Margesin R, Moertelmaier C, Mair J. Low-temperature biodegradation of petroleum
- hydrocarbons (n-alkanes phenol anthracene pyrene) by four actinobacterial strains. Int
- 754 Biodeter Biodeg. 2013; **84:**185-191.
- 755 90. Uroz S, Ioannidis P, Lengelle J, Cébron A, Morin E, Buée M, Martin F. Functional
- assays and metagenomic analyses reveals differences between the microbial communities
- inhabiting the soil horizons of a Norway spruce plantation. PLoS One. 2013; **8:**e55929.
- 758 91. Mendes LW, Kuramae EE, Navarrete AA, Van Veen JA, Tsai SM. Taxonomical and
- functional microbial community selection in soybean rhizosphere. ISME J. 2014; **8:**1577.
- 760 92. Verastegui Y, Cheng J, Engel K, Kolczynski D, Mortimer S, Lavigne J et al.
- 761 Multisubstrate isotope labeling and metagenomic analysis of active soil bacterial
- 762 communities. MBio. 2014; **5:**e01157-14.
- 763 93. Yergeau E, Sanschagrin S, Maynard C, St-Arnaud M, Greer CW. Microbial expression
- profiles in the rhizosphere of willows depend on soil contamination. ISME J. 2014; 8:344.

- 765 94. Lee SY, Sekhon SS, Ban YH, Ahn JY, Ko JH, Lee L, Kim SY, Kim YC, Kim YH.
- Proteomic analysis of polycyclic aromatic hydrocarbons (PAHs) degradation and
- detoxification in Sphingobium chungbukense. J Microbiol Biotechnol. 2016; 26:1943–
- 768 1950
- 769 95. Kumar R, Verma H, Haider S, Bajaj A, Sood U, Ponnusamy K et al. Comparative
- genomic analysis reveals habitat-specific genes and regulatory hubs within the genus
- 771 Novosphingobium. MSystems. 2017; **2:**e00020-17.
- 772 96. Tao XQ, Lu GN, Dang Z, Yang C, Yi XY. A phenanthrene-degrading strain
- Sphingomonas spGY2B isolated from contaminated soils. Process Biochem. 2007; 42:401-
- 774 408.
- 775
- 776

#### 777 **Tables** 778 Table 1. Proportion of <sup>13</sup>C remaining (after 10 days incubation, D10) compared to the 779 780 amount added at D0 in total soil, soil residue, dichloromethane extract measured using AE-IRMS and of <sup>13</sup>C-phenanthrene in dichloromethane extract measured using GC-MS. 781 782 783 Table 2: Normalized prevalence of genes encoding enzymes for aerobic bacterial degradation of aromatics in <sup>13</sup>C-labeled metagenomes, relative to recA. Values are mean 784 ± s.e.m (n=3), based on features identified using AromaDeg on assembled contigs. The 785 786 difference between bare and planted condition was tested using Welch's test. 787 788 Table 3. Detection of genes encoding putative carbohydrate active enzymes (CE: carbohydrate esterase; GH: glycoside hydrolase; PL: polysaccharide lyase) in <sup>13</sup>C-enriched 789 790 metagenomes from Sphingomonas and Sphingobium in bare and planted SIP incubations, 791 respectively. Counts are given for the three independent replicates. 792 793 794 795 **Figure Legends** 796 **Figure 1. A.** Taxonomic profiling of heavy DNA from <sup>12</sup>C-controls and <sup>13</sup>C-SIP incubations. 797 798 in bare or planted microcosms. Values are means of three independent microcosms. B. Abundance of taxonomic groups significantly enriched in heavy DNA fractions from <sup>13</sup>C-SIP 799 conditions compared to the <sup>12</sup>C-controls in bare (white) and planted conditions (grey). Values 800 801 are depicted on a log-scale and are mean $\pm$ s.e.m.(n=3). Values in brackets beside each bar 802 represent the number of OTUs. 803 804 Figure 2. Functional analysis of shotgun metagenomic paired-end reads from bare and planted <sup>13</sup>C-SIP experiments (mean ± SD, n=3) based on subsystem categories in MG-805 806 **RAST.** Relative abundance of the 28 functional categories (A), the 12 carbohydrate 807 metabolism categories (B) and the 4 categories concerning the metabolism of aromatic 808 compounds (C), with some sub-categories shown when significant differences were obtained 809 between bare and planted conditions. Relative abundances were expressed based on the total

number of hits provided by MG-RAST analysis, i.e. 5 394 642, 5 631 202 and 4 264 598 for triplicates of bare soil metagenomes, and 4 060 525, 5 619 868 and 4 759 581 for triplicates of planted soil metagenomes. Asterisks indicate statistically significant differences (q-values < 0.05 after Benjamini-Hochberg correction) between bare and planted conditions.

**Figure 3.** Maximum-likelihood (ML) phylogenetic reconstructions of dioxygenases (alphasubunit of Rieske non-heme iron oxygenases) of the biphenyl/naphthalene family from **A**) *Proteobacteria* (AromaDeg Clusters XXIV and XXVI) and **B**) *Actinobacteria* (Clusters I, II and V), identified in <sup>13</sup>C-enriched assembled metagenomes from bare (orange; NPA, NPB, NPC) and planted (green; RGA, RGB, RGC) SIP microcosms. Sequences from reference strains are included with their accession number and experimentally validated substrate (Pht: phthalate; Non: unknown; Paa: polycyclic aromatic hydrocarbons (*Actinobacteria*); Pap: polycyclic aromatic hydrocarbons (*Proteobacteria*); Nah: naphthalene; Bph: biphenyl; Phn: phenanthrene; DbtA: dibenzothiophene; Dnt: dinitrotoluene). The LG+G+I (G=1.62, I=0.02) and LG+G (G=0.57) substitution models were respectively used for A) and B) after evaluating the best model in MEGA6. ML bootstrap support (100 re-samplings) are given. Bars represent fraction of sequence divergence.

 **Figure 4.** Maximum-likelihood (ML) phylogenetic reconstruction of extradiol dioxygenases of the vicinal oxygen chelate family (EXDO I) from AromaDeg Cluster XII, identified in <sup>13</sup>C-enriched assembled metagenomes from bare (orange; NPA, NPB, NPC) and planted (green; RGA, RGB, RGC) SIP microcosms. Sequences from reference strains are included with their accession number and experimentally validated substrate (Dhb: 2,3-dihydroxybiphenyl; Dhn: Dihydroxynaphthalene; Dhp: 2,3-Dihydroxybiphenyl and dihydroxylated polycyclic aromatic hydrocarbons (probably dihydroxyphenanthrene); Dhe: 2,3-Dihydroxy-1-ethylbenzene; Thn: 1,2-Dihydroxy-5,6,7, 8-tetrahydronaphthalene; DbtC, 2,3-Dihydroxybiphenyl, probably dihydroxydibenzothiophene and dihydroxylated polycyclic aromatic hydrocarbons; Non: unknown). The LG+G+I substitution model was used (G=1.09, I=0.09) after evaluating the best model in MEGA6. ML bootstrap support (100 re-samplings) are given. The bar represents fraction of sequence divergence.

**Figure 5. Reconstruction of phenanthrene metabolic pathways.** Red arrows represent genes described in Figures 3 - 4 and enzyme families described in AromaDeg. Black arrows are genes found in both metagenomes using GhostKOALA with the reaction identifiers with

enzyme names, EC number, gene names and identification numbers of each reaction. The number of genes identified in the two conditions is added in orange (bare soil, NP) and green (planted soil, RG) above bar graphs with the taxonomic affiliation of identified genes. Note that 32.1% ( $114\,211$  /  $355\,456$ ) and 29.8% ( $225\,994$  /  $758\,006$ ) of the entries could be assigned to known functions for bare and planted soils, respectively.

Table 1. Proportion of  $^{13}$ C remaining (after 10 days incubation, D10) compared to the amount added at D0 in total soil, soil residue, dichloromethane extract measured using AE-IRMS and of  $^{13}$ C-phenanthrene in dichloromethane extract measured using GC-MS.

	proportion of <sup>13</sup> C remaining compared to amount added at D0 (%) <sup>a</sup>				
		GC-MS			
	total soil	soil residue	DCM extract	DCM extract	
$\mathbf{D0}$	$100.0 \pm 6.4$	$3.5 \pm 0.7$	$100.0 \pm 7.6$	$100.0 \pm 9.2$	
Bare D10	$97.8 \pm 6.7$	$13.0 \pm 0.7$	$79.9 \pm 10.7$	$74.3 \pm 1.9$	
Planted D10	$99.1 \pm 2.3$	$5.8 \pm 0.2$	$90.5 \pm 1.2$	$87.9 \pm 0.9$	

DCM: dichloromethane, <sup>a</sup>values are mean ± s.e.m. (n=3)

Table 2: Normalized prevalence of genes encoding enzymes for aerobic bacterial degradation of aromatics in  $^{13}$ C-labeled metagenomes, relative to recA. Values are mean  $\pm$  s.e.m (n=3), based on features identified using AromaDeg on assembled contigs. The difference between bare and planted condition was tested using Welch's test.

Engrano aloga	Prevalence relative to recA			
Enzyme class	Bare	Planted	pval	
Benzoate oxygenase	$6.6\pm0.9$	$2.0\pm0.2$	0.007	
Biphenyl/naphthalene oxygenase	$5.2\pm0.9$	$1.8\pm0.1$	0.018	
Extradiol dioxygenase, vicinal oxygen chelate	$6.1 \pm 1.2$	$2.6\pm0.1$	0.039	
Gentisate oxygenase	$2.3\pm0.7$	$1.2\pm0.2$	0.171	
Homoprotocatechuate oxygenase	$1.0\pm0.0$	$0.5\pm0.0$	0.001	
Protocatechuate oxygenase	$2.0\pm0.9$	$1.0\pm0.1$	0.304	
Phthalate oxygenase	$4.0 \pm 1.5$	$1.2\pm0.0$	0.138	
Salicylate oxygenase	$7.7 \pm 1.1$	$2.5\pm0.2$	0.011	

**Table 3. Detection of genes encoding putative carbohydrate active enzymes** (CE: carbohydrate esterase; GH: glycoside hydrolase; PL: polysaccharide lyase) in <sup>13</sup>C-enriched metagenomes from *Sphingomonas* and *Sphingobium* in bare and planted SIP incubations, respectively. Counts are given for the three independent replicates.

CAZY family	Counts based on dbcan annotation in <sup>13</sup> C-enriched metagenomes					1	- Example of known activities in family
	Sphingomonas Sphingobium in bare soil in planted soil					, , , , , , , , , , , , , , , , , , ,	
CE1	6	6	5	4	4	4	xylanase
CE14	0	0	0	1	1	1	diacetylchitobiose deacetylase
CE3	2	2	0	2	2	2	acetyl xylan esterase
CE4	4	4	3	1	1	1	chitooligosaccharide deacetylase, peptidoglycan N-acetylglucosamine deacetylase, acetyl xylan esterase
CE6	0	0	0	1	1	1	acetyl xylan esterase
CE7	0	0	0	1	1	1	acetyl xylan esterase
GH1	3	4	3	0	0	0	β-glucosidases, β-galactosidases
GH10	0	0	0	1	1	1	endo-xylanase
GH102	2	1	1	1	1	1	peptidoglycan lytic transglycosylase
GH103	2	2	1	1	1	1	peptidoglycan lytic transglycosylase
GH108	1	1	1	1	1	1	N-acetylmuramidase
GH109	2	2	2	2	2	1	α-N-acetylgalactosaminidase
GH115	0	0	0	2	2	2	xylan α-1,2-glucuronidase
GH13	1	1	1	3	3	3	α-glucosidase
GH130	2	2	2	0	0	0	β-1,4-mannosylglucose phosphorylase
GH136	0	0	0	1	1	1	lacto-N-biosidase
GH15	1	1	1	1	1	1	exolytic glucoamylase
GH16	0	0	0	2	2	2	wide variety of activities on β-1,4 or β-1,3 glycosidic bonds
GH23	5	7	6	5	4	4	peptidoglycan lyases
GH24	2	0	1	0	0	0	lysozyme
GH25	0	0	0	1	1	1	lysozyme
GH28	1	1	1	0	0	0	polygalacturonases
GH3	1	1	1	4	4	4	β-D-glucosidases, α-L-arabinofuranosidases, β-D-xylopyranosidases, <i>N</i> -acetyl-β-D-glucosaminidases
GH43_1 2	0	0	0	1	1	1	endo-β-1,4-xylanase / α-L-arabinofuranosidase
GH5	4	4	0	0	0	0	endoqlucanase, endomannanase
	0	0	0	1	1	1	
GH65	0	0	0	1	1	1	maltose phosphorylase, trehalose phosphorylase, kojibiose phosphorylase, and trehalose 6-phosphate phosphorylase
GH67 GH73	0	0	1	0	0	0	xylan α-1,2-glucuronidase
	2	2	1	0	0	0	lysozyme
GH74	1	1	0	0	0	0	xyloglucanase
GH76	0	0	0	2	2	2	α-1,6-mannanase
PL1_2	1	1	1	1	1	1	pectin/pectate lyase
PL12	0	0	0	1	1	1	heparin-sulfate lyase
PL15_2	2	2	2	5	5	5	alginate lyase
PL22	2	2	2	1	1	1	oligogalacturonate lyase
PL6 total							alginate lyase, chondroitinase B
families total	21	20	19	27	27	27	
CAZym es	47	47	36	48	47	46	











