Bacterial diversity on an abandoned, industrial wasteland contaminated by polychlorinated biphenyls, dioxins, furans and trace metals

Françoise Girardot, Séverine Allegra, Stéphane Pfendler, Cyrille Conord, Carine Rey, Benjamin Gillet, Sandrine Hughes, Anne Emilie Bouchardon, Anna Hua, Frédéric Paran, et al.

Bacterial diversity on an abandoned, industrial wasteland contaminated by polychlorinated biphenyls, dioxins, furans and trace metals.

Françoise Girardot*[a,b], Séverine Allegra[a,b], Stéphane Pfendler[a], Cyrille Conord[a], Carine Rey[c,d], Benjamin Gillet[e], Sandrine Hughes[e], Anne Emilie Bouchardon[f], Anna Hua[a], Frédéric Paran[f], Jean Luc Bouchardon[f] and Olivier Faure[f]

[a]Université de Lyon, Université Jean Monnet Saint-Etienne, CNRS, EVS-ISTHME UMR 5600, F-42023 Saint- Étienne, France

[b]Université de Lyon, Université Jean Monnet Saint-Etienne, Institut Universitaire de Technologie, F-42023 Saint- Étienne, France

[c]Université de Lyon, Université Claude Bernard Lyon 1, Ecole Normale Supérieure de Lyon (ENSL), Laboratoire de Biologie et Modélisation de la Cellule, CNRS UMR 5239, F-69634 Lyon, France

[d]Master de Biologie, ENSL, Université Claude Bernard Lyon I, Université de Lyon, F-69342 Lyon, France

[e]Université de Lyon, Université Claude Bernard Lyon 1, ENSL, CNRS, Institut de Génomique Fonctionnelle de Lyon, UMR 5242, F-69007 Lyon, France

[f]Ecole Nationale Supérieure des Mines de Saint-Étienne (ENSM-SE), Centre SPIN-EVS, UMR5600, F-42023 Saint- Étienne tienne, France.

* Corresponding author: Francoise Girardot

Telephone number: + 33 4 77 46 33 49

Email address: francoise.girardot@univ-st-etienne.fr

Abstract

Most former industrial sites are contaminated by mixtures of trace elements–and organic pollutants. Levels of pollutants do not provide information regarding their biological impact, bioavailability and possible interactions between substances. There is genuine interest in combining chemical analyses with biological investigations. We studied a brownfield where several industrial activities were carried out starting in the 1970s, (incineration of pyralene transformers, recovery of copper by burning cables in the open air). Four representative plots showing different levels of polychlorobiphenyls were selected. Organic and trace metal levels were measured together with soil pedological characteristics. The bacterial community structure and functional diversity were assessed by 16S metagenomics with deep sequencing and community-level physiological profiling. Additionally, a vegetation survey was performed. Polychlorobiphenyls (8 $mg.kg^{-1}$ to 1500 $mg.kg^{-1}$) were from 2.4 x 10^3-fold to 6 x 10^5-fold higher than the European background level of 2.5 $\mu g.kg^{-1}$. Polychlorinated dibenzo-p-dioxins and dibenzofurans ranged from 0.5 to 8.0 $\mu g.kg^{-1}$. The soil was also contaminated with trace metals, i.e., Cu>187, Zn>217 and Pb>372 $mg.kg^{-1}$. Location within the study area, trace metal content and soil humidity were stronger determinants than organic pollutants of bacterial community structures and activities. Thus, the highest biological activity and the greatest bacteriological richness were observed in the plot that was less contaminated with trace metals, despite the high level of organic pollutants in the plot. Moreover, trace element pollution was associated with a relatively low presence of Actinobacteria and Rhizobia. The plot with the highest metal contamination was rich in metal-resistant bacteria such as Sphingomonadales, Geodermatophilaceae and KD4-96 (Chloroflexi phylum). Acidobacteria and Sphingomonadales, capable of resisting trace metals and degrading persistent organic pollutants, were dominant in the plots that had accumulated

metal and organic contamination, but bacterial activity was lower in these plots than in the other plots.

Keywords: Brownfield, Polychlorobiphenyls, Polychlorinated dibenzo-p-dioxins and dibenzofurans, Metals, Community-level physiological profiling, 16S metagenomics

1. Introduction

Deindustrialization has left a legacy of many polluted wastelands. It is estimated that in French urban areas, brownfields represent an estimated real estate potential of 140,000 hectares (Plottu and Tendero, 2017). The French Database on Polluted Sites and Soils (BASOL) lists 7331 polluted sites requiring government action. Among them, 58% were polluted by at least 3 compounds and 5% by mixtures of 10 to 16 chemicals (BASOL, 2020). Among pollutants, trace metals are a worldwide issue with more than 5 million polluted sites (He et al., 2015). Given that metals are persistent, they pose a threat to ecosystem functions and human health (Li et al., 2019; Liu et al., 2018). Persistent organic pollutants (POPs), such as polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs), and polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/Fs), are resistant to biodegradation and are thus of major environmental concern in most industrialized countries (Ma et al., 2005, IARC, 2016). Indeed, these ubiquitous toxic molecules may have numerous adverse effects on living organisms, particularly on human health, as they can affect the immune, nervous, endocrine and reproductive systems (Olanca et al., 2014). Furthermore, PAHs and PCDD/Fs are recognized as carcinogenic, while PCBs are probable human carcinogens (Breivik et al., 2004).

PAHs are a family of 136 fused benzene ring molecules naturally present in coal and petroleum products. They are also produced during incomplete combustion of organic compounds (Gupta et al., 2015). PCBs gather 209 organic chlorinated molecules differing by the number (up to 10) and the position of the chlorine substitutes on phenyl rings. These synthetic compounds were used from the 19th century to the 1980s, mainly as heat transfer fluids and insulators in electrical transformers and condensers. Numerous accidental releases or illegal disposal practices have resulted in PCBs being widespread environmental contaminants (Tehrani and Van Aken, 2014). PCDD/Fs, a family of 210 congeners, are by-

products of several industrial processes (Dyke et al., 1997; Hutzinger and Blumich, 1985) or of combustion processes such as those of fossil fuel, PCBs, forest fires, and waste incineration (de Souza Pereira, 2004). In addition, the use of wood preservatives, herbicides and pesticides synthesized from chlorophenols has resulted in the release of large amounts of PCDD/Fs into the environment (Chen et al., 2016).

Precise and rapid analytical methods exist for identifying and quantifying PAHs, PCBs and PCDD/Fs from environmental samples. However, because of technical and financial-constraints, only a reduced set of congeners are routinely monitored. For PCBs, a set of 7 congeners called indicator-PCBs (I-PCB #28, 52, 101, 118, 138, 153 and 180) were selected because they were present at high levels in commercial mixtures. Similarly, 17 PCDD/Fs and 16 PAHs were identified as priorities by the US Environmental Protection Agency due to their toxicity and large presence in the environment (US EPA, 2007). Nonetheless, the only quantification of POPs and trace metals does not provide any information regarding the possible interactions between the different molecules, their bioavailability, and ultimately their biological impact (Fontanetti et al., 2011), which are of prime interest when assessing contaminated sites. Consequently, for a more informative assessment of multi-contaminated soils, there is a genuine interest in combining chemical analysis with biological investigations.

The study site was a brownfield where several industrial activities were successively carried out starting in 1970 (BASIAS, 2020): scrap-yard and car body burning, copper recovery by open-air cable burning, incineration of pyralene transformers, storage of domestic fuel, surface treatment, and wood storage. In 2008, 25000 m^3 of reclaimed wood stored on the site caught fire and burned for 3 months. An investigation carried out in 2009 revealed total I-PCB levels ranging from 10 to 2.10^3 $mg.kg^{-1}$ (unpublished data). In this context, we studied the soil biological functions and the structure of the bacterial populations occurring after 4

years on a small area of the brownfield that was climatically and pedologically homogeneous. Microbial functions were investigated by community-level physiological profiling (CLPP), which allows the discrimination of communities according to the utilization pattern of 31 carbon sources (Aguirre de Carcer et al., 2007; Classen et al., 2003; Insam, 1997; Preston-Mafham et al., 2002). To investigate bacterial biodiversity, 16S metagenomics coupled with new generation sequencing using Illumina was used. We completed our investigation by characterizing plant communities and soil pedology. Furthermore, we determined which pollutant or abiotic parameter was the main driver of bacterial structure and function. Rather than solely reporting levels of contaminants, our study aims to provide a much more precise picture of the impact of abiotic factors on the bacterial fraction taken as an indicator of soil functioning. To our knowledge, this study is the first field investigation using these methods for assessing soils with multiple contaminants.

2. **Materials and methods**

2.1 Site description and sampling plots

The study site is located at Saint-Cyprien (Lambert-93 coordinates: X=795445, Y=6494391), approximately 30 km north of Saint-Etienne (Loire, France). It is listed in the Basol database (BASOL, 2020) under reference 42.0034. The site is an abandoned wasteland and approximately 8,000 m^2. In 2009, the site was divided into 80 plots of 100 m^2, and first I-PCB analyses (samples collected from the surface to 1 m in depth) were performed in the soil as part of the legal investigation (unpublished data). Based on these data, 4 plots were selected for this study (Fig. 1): the highest PCB-contaminated area in the north (F2) and the south (D10) and the lowest PCB-contaminated area in the north (C2) and the south (A10) of the parcel. Unlike the southern areas, the northern areas were flooded during the rainy season.

2.2 Soil sampling

Five sampling campaigns were performed in 2012 (April 19th, June 1st and 22nd) and 2013 (April 8th and July 1st), and 5 soil samples per plot (0-10 cm in depth) were collected during each sampling after removal of the litter. In the soils, the leaching of PCPs is very low, and the molecules tend to accumulate at the surface. Therefore, the choice was made to harvest only the top fraction to avoid an underestimation of organic pollution. For microbial analyses, samples were collected aseptically in sterile containers and stored at 4°C before sieving through a 2-mm mesh sterile screen. Samples for 16S metabarcoding were stored at -80°C. The moisture content was determined in parallel with each analysis using the standardized ISO 11465 method (ISO11465, 1994).

2.3 Soil pedological characteristics

Pedological analyses (particle size distribution, organic carbon, total available phosphorus, total nitrogen, total carbonates, pH in water, and cation exchange capacity (CEC)) were performed by the Laboratory for Soil Analyses at the National Institute for Agronomic Research (INRA Arras, France accreditation n° 1-1380, https://www.cofrac.fr).

2.4 Analyses of organic pollutants

PCBs, PCDD/Fs and PHAs were analysed by the INRA Laboratory for Soil Analyses. The toxicity equivalent concentrations (Teq) for the PCDD/Fs and PAHs were calculated as follows (WSDE, 2007):

$$\mathrm{Teq} = \sum \mathrm{C_n} \times \mathrm{TEF_n}$$

where C_n is the concentration of the individual PCDFs, PCDDs or PAHs, and TEF_n is their toxic equivalency factor. For PCBs, Teq was calculated with only 2 dioxin-like PCBs recorded, i.e., PCB-105 and PCB-118 congeners.

156 *2.5 Trace metal determination*

157 The total trace elements (TM) were determined by X-ray fluorescence. Available trace
158 elements were extracted with 0.01 M calcium chloride as previously described (Houba et al.,
159 2000) and measured by inductively coupled plasma mass spectrometry (ICP-AES).

160

161 *2.6 Plant survey*

162 A survey of the main plant species present in each plot was carried out 5 times in 2012
163 and 2013 by 2 experienced field botanists. Plant names were standardized according to the
164 Plant List Database (http.theplantlist.org). For each species, Raunkiaer's life-form type and
165 Ellenberg's autoecological ratings (preferences for light, temperature, pH, N-P-K nutrients,
166 salinity, and soil humidity) were calculated from the Baseflor database (Julve, 1998).
167 Ellenberg's indicator values (EIVs) for each plot were calculated as the median value of all
168 species identified in the plot.

169

170 *2.7 Community-level physiological profiling*

171 EcoPlates (Biolog, France) contain 96 wells with 3 replicates of 31 carbon substrates
172 and a control well without any carbon source (Insam, 1997). Samples were prepared by
173 mixing 10 g fresh soil with 90 mL of sterile 0.2% w/v sodium hexametaphosphate at room
174 temperature for 30 min. The suspension was then diluted 100-fold in phosphate-buffered
175 saline to minimize background colour. Each well was then inoculated with 150 µL of the
176 suspension. Three plates were used for each plot, and data were collected after 45 h of
177 incubation, corresponding to approximately the mid-exponential growth phase. Formazan
178 production was quantified by measuring the optical density at 590 nm (OD^{590}) with a
179 microtiter plate reader spectrophotometer (Tecan, Switzerland). Initial OD^{590} values were
180 measured immediately after inoculation and were subtracted from subsequent readings. Plates

were then kept at 25°C in the dark, and the raw OD^{590} was read regularly until it reached a plateau. The mean raw OD^{590} of each similar well was calculated, and the raw OD^{590} values were corrected for the background colour in the control well. Wells with OD^{590} values>0.25 were considered positive (Garland, 1996), while others were set to zero. Data were normalized by dividing the OD^{590} of each well by the dry mass of the soil samples (Classen et al., 2003). For each plot, the overall microbial activity was estimated by calculating the average well colour development (AWCD). Functional diversity was estimated by assessing both the substrate richness (S) and Shannon's diversity index (H) as previously described (Garland, 1996; Weber et al., 2007).

2.8 Soil DNA extraction, 16S amplification and sequencing

Total DNA was extracted from 0.25 g of soil for each of the 12 samples (4 plots and 3 sampling dates in 2012 and 2013) using the Power Soil™ DNA extraction kit (Mo BIO Laboratories, Carlsbad, CA, USA) and according to the manufacturer's instructions.

The V3 region of the bacterial 16S rRNA gene was amplified following the protocols detailed in Téfit et al. (2018), except that 25 PCR cycles were performed. After construction of new generation sequencing (NGS) libraries and equimolar mixing, the amplicons were sequenced in a PGM sequencer (Ion torrent) using Ion 318 chips (Téfit et al 2018). Several procedures to avoid and monitor exogenous bacterial DNA contamination with multi-step controls were strictly followed: mainly, experiments were conducted in pre- and post-PCR dedicated rooms, with UV decontamination and mock extraction, and the negative and aerosol PCR controls were performed as detailed elsewhere (Téfit et al., 2018).

2.9 Metagenomics analysis

First, the quality of the reads was checked, and sequences shorter than 300 bp, with quality scores less than 20 or containing one or more unknown bases were deleted using Trimmomatic (Bolger et al., 2014). Then, the reads were re-assigned to each sample according to their unique barcode with no mismatch using FASTX-Toolkit software. The barcodes were removed before subsequent analysis. Finally, taxonomic assignment and operational taxonomic unit (OTU) definition were performed using the Mothur pipeline (Schloss et al., 2009) following the steps detailed below.

First, the sequences were aligned to the SILVA database (Quast et al., 2013) (restricted to the 16S V3 domain) using the Needleman-Wunsch distance. Reads with weak alignment scores were discarded (SearchScore < 60 and SimBtwnQuery&Template < 80). Then, sequences were pre-clustered at 99% to reduce sequencing noise and computational effort during the taxonomic assignment of each cluster representative in the SILVA database. Next, OTUs were defined using abundance-based greedy clustering (AGC) at 97% sequence similarity. All scripts used for the analyses are available upon request. Finally, 122401 OTUs were obtained for the 12 samples. The data have been submitted to the NCBI bioproject database under the number PRJNA647634 and sequence data for all samples are available with the accession numbers going from SAMN15595561 to SAMN15595572.

2.10 Statistical analysis

Statistical analyses were performed using R software v. 1.0.136 (R core team, 2016). Differences were considered significant for p values <0.05. The distribution-free Kruskal-Wallis test, followed by the Steel-Dwass-Critchlow-Flignerwas (SDCF) multiple comparisons tests, was used to compare pedologic descriptors among the samples. Principal component analysis (PCA) was performed to compare the CLPP patterns of substrate utilization and 16S metagenomics data for each plot. Additionally, microbial community composition was

230 calculated using the SHiny Application for Metagenomics Analysis (Shaman) based on the
231 DESeq2 R package (Quereda et al., 2016).
232 For TM analysis, the results were pre-treated to obtain compositional data (relative
233 proportion of the compound in the sample) and were log-ratio centred and scaled to improve
234 the outputs from the multivariate analyses (Hervé et al., 2018). A redundancy analysis (RDA)
235 was used to test the simple and interactive effects of one or more explanatory variables on the
236 differences in the elementary profiles using a permutation test.
237 3. **Results**

238 *3.1 Pedological characteristics of the four sampling plots*
239 Table 1 shows the pedological characteristics of the 4 studied plots. Soils were
240 characterized by a high sand content (>70%) and low clay (6-11%) and silt (9-17%) contents.
241 There was a significant difference ($p<0.05$) in the sand content between plots C2 and D10 and
242 in the clay content between plot A10 and plots C2 and F2. According to the United States
243 Department of Agriculture texture classification, plot C2 was sandy loam, whereas the other
244 plots were loamy sand. The contents of organic carbon, total nitrogen, available phosphorus,
245 total carbonates and CEC were not significantly different between the plots. The C/N ratio
246 was significantly lower in the F2 plot than in the other plots (14 for F2 and > 20 for the other
247 plots). The soil pH in the water was significantly higher in the A10 plot (8.1) than in the F2
248 (6.9) and D10 (7.1) plots.
249

250 *3.2 Levels of POPs in the 4 sampled plots*
251 The levels of POPs in the 4 sampled plots are given in Table 2 (see supplementary
252 material Table S1 and Fig. S1 for more details). In addition to the 7 I-PCBs generally
253 monitored for soil pollution characterization, we analysed 13 supplemental congeners to
254 better cover the range of potential chlorination levels and to obtain insight about highly

chlorinated congeners. On average, the 7 I-PCBs represented 70-75% of the total concentration of the 20 selected PCBs. High heterogeneity in the concentrations was observed, ranging from approximately 8 $mg.kg^{-1}$ DW in the C2 plot to close to 1.5 10^3 $mg.kg^{-1}$ DW in the D10 plot (Table 2). The distribution pattern of the studied congeners was almost the same for the C2, A10 and D10 plots and was characterized by a high proportion (more than 75%) of highly chlorinated molecules with 6 or more chlorine atoms. In contrast, in the F2 plot, the most prominent congeners were tetra-, penta- and hexa-chlorinated PCBs, together representing close to 95% of the total PCBs.

Soil contamination with PCDD/Fs was also highly heterogeneous (supplementary material, Fig. S1c), ranging from approximately 0.5 µg $Teq.kg^{-1}$ DW in the A10 plot to 7.9 µg $Teq.kg^{-1}$ DW in the D10 plot (Table 2). In the C2 plot, more than 80% of the PCDD/Fs were hepta- and octo-chlorinated congeners (supplementary material, Fig. S1d). In contrast, in the D10 plot, more than 80% of the PCDD/Fs were congeners with 4 to 6 chlorine substitutes. In the A10 and F2 plots, approximately one half of the congeners were molecules with 4 to 6 chlorines, and the other half were heavier molecules with 7 or 8 chlorine substitutes.

PAH contamination followed the same trend as PCB contamination, with a clear concentration gradient (supplementary material, Fig. S1e), ranging from 1 µg $Teq.kg^{-1}$ DW in the C2 plot to 260 µg $Teq.kg^{-1}$ in the D10 plot (Table 2). The distribution pattern of the PAH congeners was almost the same in plots A10, F2 and D10, with approximately 75% of the molecules consisting of 4 to 6 rings. By contrast, in plot C2, almost 90% of the congeners were lighter molecules with 2 or 3 rings, the most prominent of them being naphthalene (approximately 60% of the total PAHs).

3.3 Levels of trace metals in the 4 sampled plots

279 The levels of trace metals in the 4 sampled plots are shown in Table 2. In all plots, the
280 levels of Cu, Zn and Pb were higher than 187, 217 and 372 $mg.kg^{-1}$, respectively. Regardless
281 of the TM considered, the highest levels were measured in plot C2, while the lowest levels
282 were found in plot F2. Concerning $CaCl_2$-extractible trace metals, the highest levels were
283 found in the C2 plot for Cu but in the plot D10 for Pb and Zn. TM patterns were analysed by
284 compositional PCA. The two first components explained 71% of the variance (Fig. 2). The C2
285 plot was separated from the other plots on the component 1 axis; by contrast, the second
286 principal component clearly separated the F2 plot from the A10 and D10 plots, which were
287 grouped. The C2 plot differed from the other plots by its levels of Cu, Pb and Zn (Fig. 2;
288 Table 2), and the F2 plot was characterized by a globally lower content of TM.
289
290 *3.4 Plant community structure*
291 The species composition of the plant communities found in each plot is shown in the
292 supplemental material (Table S2). Within the 4 investigated plots, 71 different species from
293 25 different families were identified. The number of species per plot varied from 33 (F2 plot)
294 to 41 (A10 plot). Thirteen species-were found in each plot, while 33 were identified in only
295 one plot. The visual estimate of plant cover was approximately 50% in all plots except at F2,
296 where it was close to 70%. The Asteraceae family was predominant in all plots except in the
297 F2 plot, where Fabaceae were the most represented. Most species were Hemicryptophytes
298 (*i.e.*, perennial grasses and herbs), accounting for 45% (F2 plot) to 62% (D10 plot).
299 Therophytes (*i.e.*, annual grasses and herbs) accounted for 27% (F2 plot) to 37% (A10 plot) of
300 the plant communities. Geophytes (*i.e.*, perennial plants with bulbs or rhizomes) accounted
301 for between 5% (D10 plot) and 15% (F2 plot) of the plants, and Phanerophytes (*i.e.*, shrubs
302 and bushes) accounted for between 0% (A10 and D10 plots) and 12% (F2 plot).

303 The EIVs (Fig. 3) showed that plant communities in each plot had the same ecological
304 requirements except for edaphic humidity, which was slightly higher in plots C2 and F2 than
305 in the other plots.
306
307 *3.5 Community-level physiological profiling*
308 The assessment of the bacterial communities by CLPP showed that in all plots, 7
309 substrates out of 31 were never used (D-xylose, phenylethylamine, L-threonine, D-
310 glucosaminic acid, glycyl-L-glutamic acid, α-ketobutyric acid and α-D-lactose), while 9 of
311 them were catabolized in all the plates (D-galacturonic acid, L-asparagine, tween 40, tween
312 80, D-mannitol, 4-hydroxybenzoic acid, L-serine, N-acetyl-D-glucosamine, and putrescine).
313 The other substrates were differentially metabolized according to the samples. Substrate
314 richness (S), Shannon's diversity index (H) and AWCD values (Table 3) were the highest in
315 the F2 plot, followed by the C2 plot. The lowest values were found in the A10 and D10 plots
316 and were not significantly different.
317 Substrate utilization patterns were further compared using PCA (Fig. 4a). The two first
318 components explained 83% of the variance. The first axis (62%) was mostly driven by 4-
319 hydroxybenzoic acid, D-galactonic acid-γ-lactone, pyruvic acid methyl ester, D-galacturonic
320 acid, D-mannitol, itaconic acid and putrescine, while the second axis (21%) was mostly
321 driven by 2-hydroxybenzoic acid, D-malic acid, D,L-α-glycerol phosphate, α–cyclodextrine
322 and L-serine (supplementary material,Fig. S2a). The scatter plot of principal component
323 scores showed a clear separation of the studied plots. The C2 plot was mainly separated from
324 the other plots on the first axis; by contrast, the second principal component clearly separated
325 plot F2 from the other studied plots.
326
327 *3.6 Taxonomic composition of the soil samples and comparison of bacterial communities*

328 A total of 122,401 OTUs were obtained from the data analysed. The OTUs were
329 grouped into 32 phyla, 923 families and 983 genera. The bacterial community (Fig. 5) was
330 mainly represented by Proteobacteria (30 to 50%), Acidobacteria (10 to 40%) and
331 Actinobacteria (12 to 35%). The profiles obtained for the different sampling campaigns
332 showed similar patterns. In each plot, the 3 main orders were Sphingomonadales, Rhizobiales
333 and Acidobacteriales. However, their respective contributions to the bacterial community
334 varied between plots: Acidobacteriales was dominant in the A10 and D10 plots with 27% and
335 48% OTUs, respectively; Sphingomonadales was dominant in the C2 plot (16%), and
336 Rhizobiales was dominant in the F2 plot (26%). The fourth bacterial group was
337 Actinobacteria (6 to 7%), which was present in all plots except C2, where it was Chloroflexi
338 (8.5%) was dominant. The family richness was similar between the F2 and C2 plots (477±20
339 and 470±34, respectively) but significantly lower for the A10 and D10 plots (395±32 and
340 403±3, respectively).
341 The PCA carried out at the family level is shown in Fig. 4b and supplementary material,
342 fig. S2b. The first axis (31.5%) clearly separated plots A10 and D10 from plots C2 and F2. In
343 addition, the second axis-(19.4%) separated the F2 plot from the 3 other plots.
344 4. **Discussion**
345 *4.1 Study site has unusual high levels of PCBs and PCDD/Fs*
346 In European countries, the background concentration for PCBs in soils is
347 approximately 2.5 $\mu g.kg^{-1}$ (INERIS, 2012; Li et al., 2010). In comparison with these reference
348 values, the values in the 4 studied plots had marked PCB contamination (Table 2), reaching
349 between 2.4 x 10^3-fold (int the C2 plot) and 6 x 10^5-fold (in the D10 plot) the background
350 level. It is noteworthy that no site listed to date at a global scale had such levels of PCBs
351 (IARC, 2016). The studied site also showed very high PCDD/F contamination, with
352 concentration values ranging from 250-fold (the A10 plot) to 3,900-fold (the D10 plot) the

353 background level of 0.002 µg Teq.kg^{-1}, as determined for European soils (Bodenan and
354 Michel, 2013). In contrast, PAH contents were in the range of the background levels for
355 European soils, that is, <2 10^3 µg TE.kg^{-1} (Nam et al., 2008). Moreover, the levels of PAHs
356 did not exceed the predicted non-effect concentration (PNEC) of 320 µg.kg^{-1} established for
357 benzo(a)pyrene (INERIS, 2006) or the ecological soil screening level (Eco-SSL) of 18 mg.kg^{-}
358 ^{1} established by the US EPA (US EPA, 2007) for high molecular weight PAHs. Therefore,
359 although no PNEC or Eco-SSL has been published to date for PCBs and PCDD/Fs, we c
360 conclude that the POPs of concern at the study site are PCBs and PCDD/Fs but not PAHs.
361 Moreover, the 4 sampled plots differed in their POP profiles. This was particularly
362 noticeable for the PCBs in the F2 plot, which showed a higher proportion in minimally
363 chlorinated congeners compared with that in the other plots. This was also the case for
364 PCDD/Fs, which were enriched in highly chlorinated congeners in the C2 plot, while they
365 were mostly minimally chlorinated in the D10 plot. It has been demonstrated (Correa et al.,
366 2010) that PCB toxicity towards bacteria varies among congeners and is higher for highly
367 chlorinated congeners. At the other end of the spectrum, the bioavailability of PCBs and
368 PCDD/F in soils, *i.e*., their capacity to interact with biological targets, generally decreases
369 with the level of chlorination (Shiu and Mackay, 1986). In addition, the toxicity of POPs
370 depends on the content and form of organic matter in soil (Kim and Lee, 2002) and could
371 increase due to synergistic interactions with trace metals.
372 Trace metal contamination was identified for Pb, Cu and Zn compared to their
373 background levels in agricultural French soils (Baize, 2000). Due to the local geology, some
374 Zn and Pb could be linked to the presence of zincblende (ZnS) and galena (PbS) in detrital
375 materials. However, the high quantities of Zn and Pb in the study area are more likely due to
376 former car-burning activities, as paint residues, tire debris and motor oil are sources of Zn and

377 Pb (Garba and Abubakar, 2018). The high copper concentrations are probably related to the
378 burning of electrical cables.

379 In our study, 0.1 to 0.3% Cu, less than 0.02% Pb and 0.05 to 1.6% Zn were extracted
380 by $CaCl_2$. These low percentages are similar to those obtained in a study with loamy sand-
381 and sandy loam-contaminated and uncontaminated Japanese soils (Kashem et al., 2007).

382 Metal availability depends on metal concentrations and soil properties such as pH,
383 organic matter, CEC or clay content (Almas and Singh, 2017, Giller et al., 2009). Although
384 plot C2 had the highest total TE levels, the $CaCl_2$-extractable fraction of Pb and Zn was
385 higher in plot D10. This is likely because of heterogeneities in the pedological characteristics
386 since metal availability depends on metal concentrations as well as on soil properties (Almas
387 and Singh, 2017, Giller et al., 2009). In addition to PCB and PCDD/F toxicity, Cu, Zn and Pb
388 toxicity towards soil biota could be expected.

389 In summary, the 4 studied plots showed a clear concentration gradient for PCBs
390 increasing in the order of C2<A10<F2<D10, PCDD/Fs increased in the order of
391 C2/A10<F2<D10, and total trace metals followed the order of F2<A10/D10<C2. However,
392 the PCAs had 2 different major patterns: (*i*) the A10 and D10 plots, located in the south part
393 of the parcel, showed similar metal pollution levels and could be compared based on their
394 difference in POP levels, and (*ii*) the F2 and C2 plots, located on the north of the site, in a
395 periodically flooded area, showed either high metal pollution (C2) or very high POP levels
396 (F2). Comparing the F2 and C2 plots could help determine the main driver among the metal
397 and organic pollutants shaping-bacterial communities.

398

399 *4.2 Plants*

400 The vegetation at the site was dominated by perennial and annual grasses and herbs.
401 Except in the F2 plot, where the plant cover was significantly higher than in the other plots,

the vegetation was sparse and exhibited areas with bare soil, as is commonly observed in metal-polluted soil (Heckenroth et al., 2016; Nagendran et al., 2006). Vegetation also included plants able to accumulate Zn (*Silene vulgaris* and *Hypericum perforatum*) or Pb (*Festuca ovina*), 50% of the plants were typical of abandoned lands (e.g., *Bromus tectorum* L. and *Senecio inaequidens* DC) and 5 pioneer plants (e.g., *Silene vulgaris* and *Alisma plantago-aquatica* L.) were found. We also found 1/4 of annual commensal plants associated with crops (e.g., *Senecio vulgaris* and *Sonchus asper*)

According to Ellenberg index values (EIVs), the difference in the moisture content between the plots was the main driver of vegetation. Similarly, studies comparing EIVs with field measurements have shown discrepancies (Chytry et al., 2009*)*, and EIVs do not take pollution into account.

4.3 Functional diversity of microbial communities depends on plot localization

CLPP is regarded as a culture-based assay even if both culturable and non-culturable bacteria can contribute to tetrazolium salt reduction. Some authors have claimed that CLPP highlights a very small fraction (i.e., <1%) of a bacterial community (Torsvik and Ovreas, 2002) and that fast-growing species skew the test (Preston-Mafham et al., 2002). Moreover, some bacteria are not able to reduce the tetrazolium dye, and therefore, the functional ability of the entire community is not measured (Ros et al., 2008). On the other hand, some researchers have reported that culture methods are more appropriate for studying the effect of contaminants because cultivable bacteria "may constitute the ecologically relevant genetic pool in soils" (Puglisi et al., 2012). In this study, CLPP showed that the bacterial activities in the A10 and D10 plots were very similar and clearly separated from those in the C2 and F2 plots. Therefore, CLPP reflected differences related to the presence of TM.

Substrate richness (S) and Shannon's diversity index (H) AWCD were highest in the F2 plot, where POP levels were very high, and the lowest in the A10 and D10 plots, which had both high metal and organic pollution. The C2 plot, which was the most contaminated with TM, showed intermediate S, H and AWCD. Taken together, these results suggest that neither TM alone nor POPs alone were major limiting factors for bacterial activities. In contrast, the combination of TM plus POPs had a significant impact on bacterial communities at this site. Studies on the effects of metal or organic pollutants on bacterial communities are contradictory (Puglisi et al., 2012, Valentin et al.2013), but it is assumed that many-bacterial functions are redundant and can quickly replace those lost due to a given environmental stressor (Allison and Martiny, 2008, (Maron et al., 2018). However, in the case of multiple stressors, such as metals plus organic pollutants, the response is more complex, depending on soil properties and the type of stressors (Azarbad et al., 2015). In our study, the combination of metal stress and PCBs-PCDD/Fs was detrimental to bacterial activities.

4.4 Pattern of bacterial taxa may be affected by pollution

Proteobacteria are usually the dominant phylum in soils, followed by Acidobacteria, Actinobacteria and Verrucomicrobia (Bruns, 2017). Firmicutes, Planctomycetes, Chloroflexi, Gemmatimonadetes, and Cyanobacteria are generally found at levels of 1 to 5% (Bruns, 2017). In an Italian wasteland, Ventorino et al. (2018) showed that uncontaminated areas consisted of 39% Actinobacteria. However, in areas contaminated by Zn, Cu-and PHAs, the abundance of Actinobacteria, Firmicutes, Gemmatimonadetes and Chloroflexi significantly dropped, while the relative abundance of Proteobacteria, Acidobacteria, Verrucomicrobia and Bacteroidetes increased. In another study comparing river sediments (Yin et al., 2015), in comparison to other phyla, Firmicutes, Chloroflexi and Acidobacteria were more represented in highly TM contaminated samples. In our study, Proteobacteria, Acidobacteria and

Actinobacteria were the most abundant phyla in each plot. However, Verrucomicrobia was hardly detected (0.7% of OTUs), and Chloroflexi was the fourth most dominant phylum. At the study site, Actinomycetes did not exceed 7% of the 10 major taxa. TM pollution could be responsible, as it has been shown that Actinobacteria are sensitive to metals (Lugauskas et al., 2005; Yin et al., 2015).

Additionally, Rhizobiales (a subgroup of α-Proteobacteria) were approximately 2-fold higher in the F2 plot (which also had the greatest richness of OTUs) than in the other plots. Several studies (Chaudri et al., 2008; Giller et al., 1989; Hirsch et al., 1993) have shown that most Rhizobiales are Zn sensitive (killed or lose their ability to fix nitrogen). Because F2 had a relatively low Zn content, we may assume that Zn pollution could be responsible for the relatively low abundance of Rhizobiales in metal-contaminated plots.

4.5 Bacterial communities are shaped by metal and POP pollutants

Acidobacteria (*e.g.*, RB41 family, subgroup 6, *Candidatus Solibacter* and *Candidatus Koribacter*) were dominant in the D10 plot and second most dominant in the A10 plot, although the soil pH was neutral to alkaline. It has been demonstrated that exposure to highly chlorinated PCBs (such as those found in the D10 and A10 plots at a high level) results in a higher abundance of members of Acidobacteria to which the most PCB degraders belong (Aguirre de Carcer et al., 2007). Moreover, Acidobacteria subgroup 6 is known for its high metal resistance (Yin et al., 2015), and a positive correlation between Cu, Pb and Zn and the abundance of Acidobacteria, such as *Candidatus Koribacter*, has been recorded (Wang et al., 2018). Therefore, the abundance of Acidobacteria in the A10 and D10 plots could be linked to their capability to cope with metals and PCBs simultaneously.

Geodermatophilaceae and KD4-96, which belong to Chloroflexi, were well represented in the C2 plot. Both groups are linked in the literature with highly heavy metal- or

TE-polluted soils (Normand et al., 2014; Yin et al., 2015). Sphingomonadales were very abundant in all plots. Members of this taxa are able to degrade minimally chlorinated PCBs and dioxins using aerobic pathways (Abraham et al., 2002; Field and Sierra-Alvarez, 2008), as is also the case for *Hyphomicrobium* present in the F2 plot.

5. Conclusions

The study site was one of the most PCB- and PCDD/F-polluted sites ever described worldwide (IARC, 2016), and it has accumulated significant metallic pollution. Our study showed clear biological and chemical differences between plots. The assessed characteristics of the plots, including total elements, microbial activities, bacterial composition and, to a lesser extent, vegetation, provided evidence that the C2 and F2 plots were clearly different and distinct from each other from the A10 and D10 plots, which appeared similar. The combined presence of both metals, PCBs and PCCD/Fs (e.g., in the D10 and A10 plots), was probably the cause of the poorer functional diversity in these plots than in the other plots. TM impaired the bacterial community structure through the inhibition of Rhizobia and Actinobacteria. However, we determined that the bacterial groups present were able to cope with multi-contamination of the soil (Acidobacteria, Sphingomonadales, and Chloroflexi).

These observations show that this brownfield site should be popular choice for further studies.

6-Acknowledgements

This project was financially supported by Saint Etienne Metropole. We thank the University of Saint Étienne, IGFL and Mines Saint Etienne for their support. We thank Thibault Munoz for his technical assistance and the students of the Biology Master (ENS

Lyon/UCBL) from the UE “Practicals in NGS” 2018 and its organizer Marie Sémon for the metagenomic analysis.

Table
Click here to download Table: Tables_ Anapop_revised.docx

Tables

Table 1: Median (median absolute deviation between brackets) of selected pedological characteristics of the 4 studied plots (n=5). In a same column, different letters indicate significant differences between plots at $p<0.05$. C org: organic carbon; N tot: total nitrogen; $CaCO_3$: total carbonates; CEC: cationic exchange capacity; P_2O_5: available phosphorus.

Plots	**Clay** (%)	**Silt** (%)	**Sand** (%)	**C org** ($g.kg^{-1}$)	**N Tot** ($g.kg^{-1}$)	**C/N** ratio	**pH_{water}**	**$CaCO_3$** ($g.kg^{-1}$)	**CEC** ($cmol+.kg^{-1}$)	**P_2O_5** ($mg.kg^{-1}$)
A10	6(0)b	15(2)a	79(1)a,b	13(2)	0.6(0.1)	21(1)a	8.1(0.3)a	5.3(3.9)	5(1)	21(1)
D10	9(1)a,b	9(1)a	81(3)a	24(4)	0.9(01)	26(3)a	7.1(0.3)b	0.8(0.3)	5(1)	22(4)
C2	11(2)a	17(7)a	70(4)b	76(57)	3.4(0.3)	22(6)a	7.5(0.5)a,b	44(43)	10(4)	27(3)
F2	11(5)a	9(6)a	82(1)a,b	24(22)	1.7(0.3)	14(1)b	6.9(0.4)b	3.6(4.9)	4(2)	24(4)

Table 2 : Levels of organic (POPs) and total concentration of selected TM (n=4, standard deviation between brackets). Teq : toxic equivalents. According to usual TM levels in agricultural soils (Baize, 2000), * indicates moderate anomalies, ** major anomalies and *** above major anomalies. Lowest and highest values are shown in the table in light grey and dark grey, respectively.

		Soil plots			
		A10	D10	C2	F2
POPs	∑20 PCBs (mg.kg^{-1})	71	1480	8	738
	∑7 I-PCBs (mg.kg^{-1})	49	1088	6	557
	∑2 PCBs-DL (105,118) (mg.kg^{-1})	0.8	23	0.5	174
	∑5 Dioxins + 10 Furans (µg. kg^{-1})	7	60	29	37
	∑ 16 PAHs (mg.kg^{-1})	0.9	1.7	0.7	1.1
	Teq PCB-DL 105 +118 (ng.kg^{-1})	25	678	14	5226
	Teq PCDD/Fs (µg.kg^{-1})	0.5	7.9	0.7	2.4
	Teq PAHs (µg.kg-^{1})	116	260	1	167
TM (mg.kg^{-1})	Cu	516(140)***	814(164)***	1676(394)***	187(86)***
	Pb	1445(413)**	1599(534)**	2343(608)**	372(86)**
	Zn	500(144)**	563(109)**	2461(984)**	217(35)*

Table 3: Bacterial activities as determined by CLPP. Substrate richness (S), Shanon's diversity index (H), and Average well color development (AWCD). Mean values, n=3, standard deviation between brackets. Different letters in a same column indicate significant differences at $p<0.05$.

Plots	S	H	AWCD
A10	13(1)b	2.48(0.0)8b	0.25(0.06)b
D10	11(1)b	2.30(0.11)b	0.19(0.01)b
C2	16(1)a	2.65(0.07)a	0.45(0.03)a
F2	19(1)c	2.86(0.07)c	0.58(0.03)c

Figure
Click here to download Figure: Figures_revision.docx

Figures

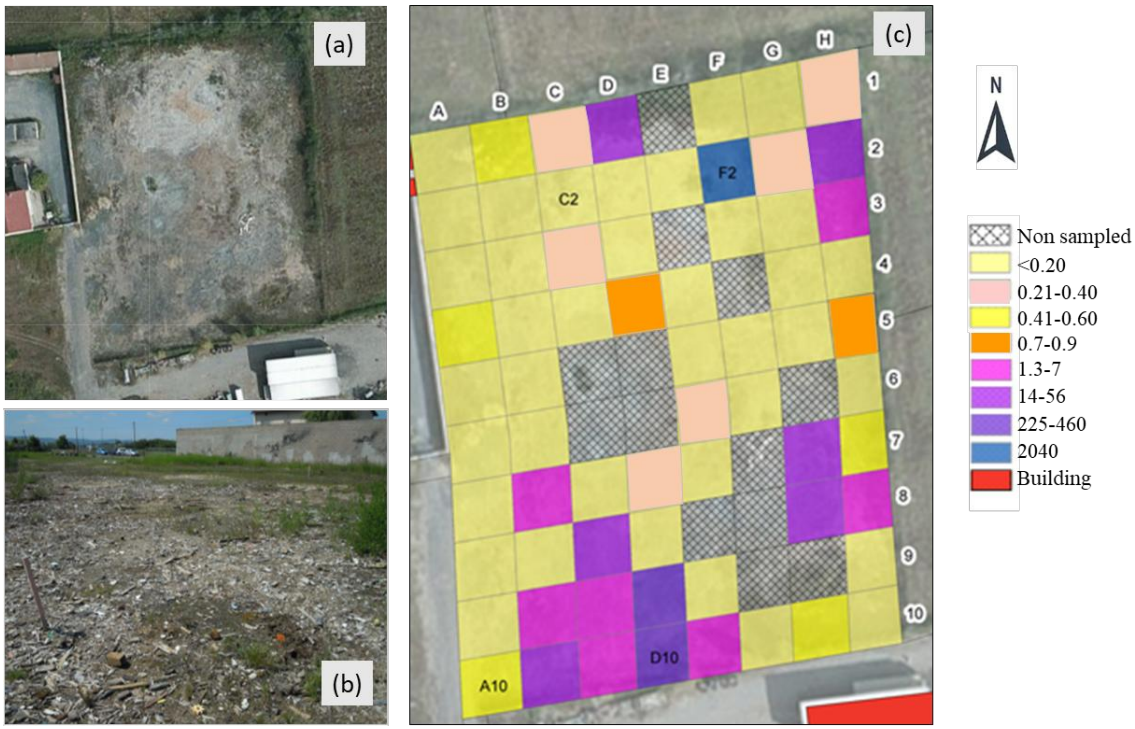

Fig. 1: (a) Aerial view of the study site. (b) General view of the site. (c) Ranges of I-PCBs contamination (mg.kg^{-1} DW), as measured in 2009 in a first pollution assessment (hatched areas were not sampled). Plots selected in the present study are A10, C2, D10 and F2.

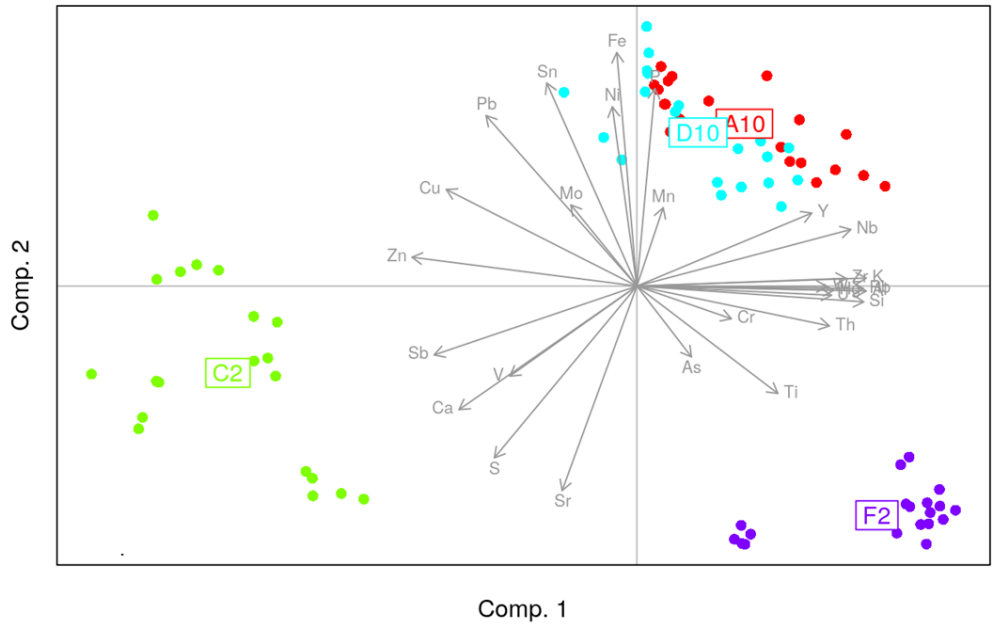

Fig. 2: Compositional principal component analysis of total TM values from the four studied plots.

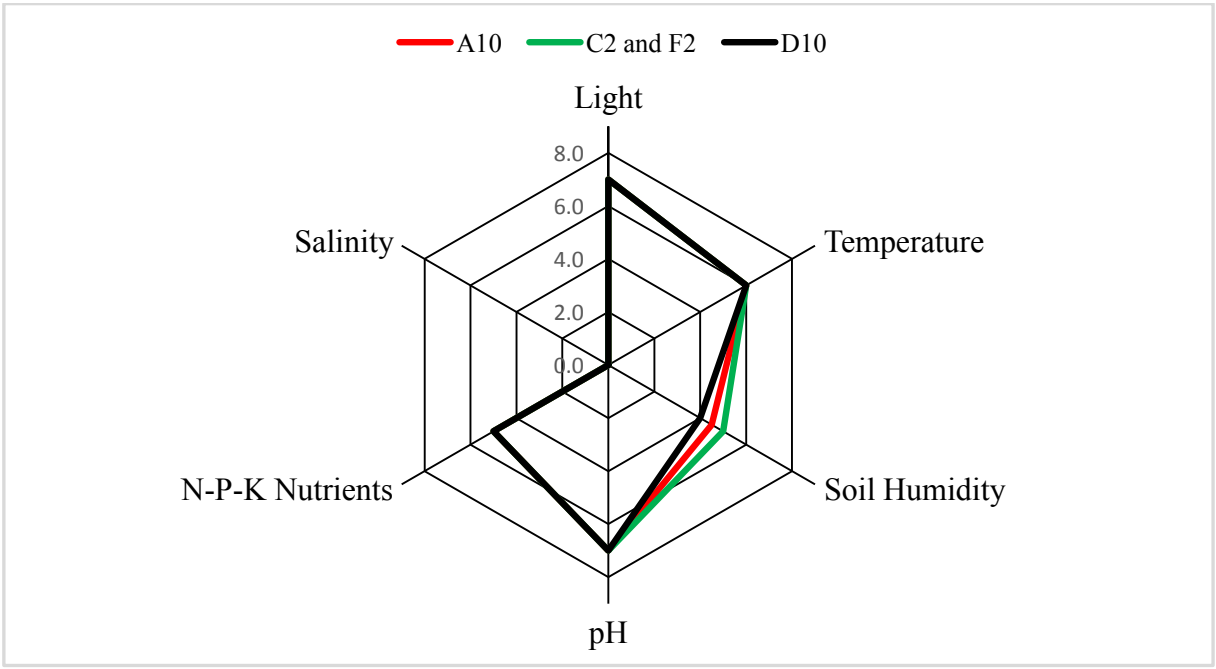

Fig. 3: Median of Ellenberg's Indicator values of plant communities at each plot. For C2 and F2, EIVs were identical.

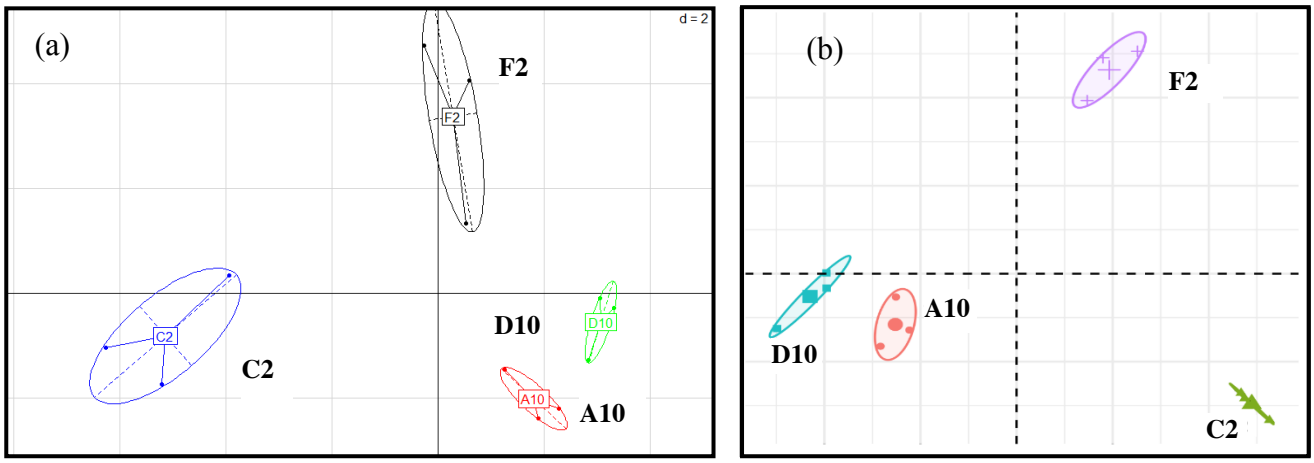

Fig. 4: Principal component analysis of Biolog ecoplates values (a) and bacterial diversity observed with 16S sequencing at the family level for the 12 taxa that contribute the most to the axis diversity (b) from the four studied plots.

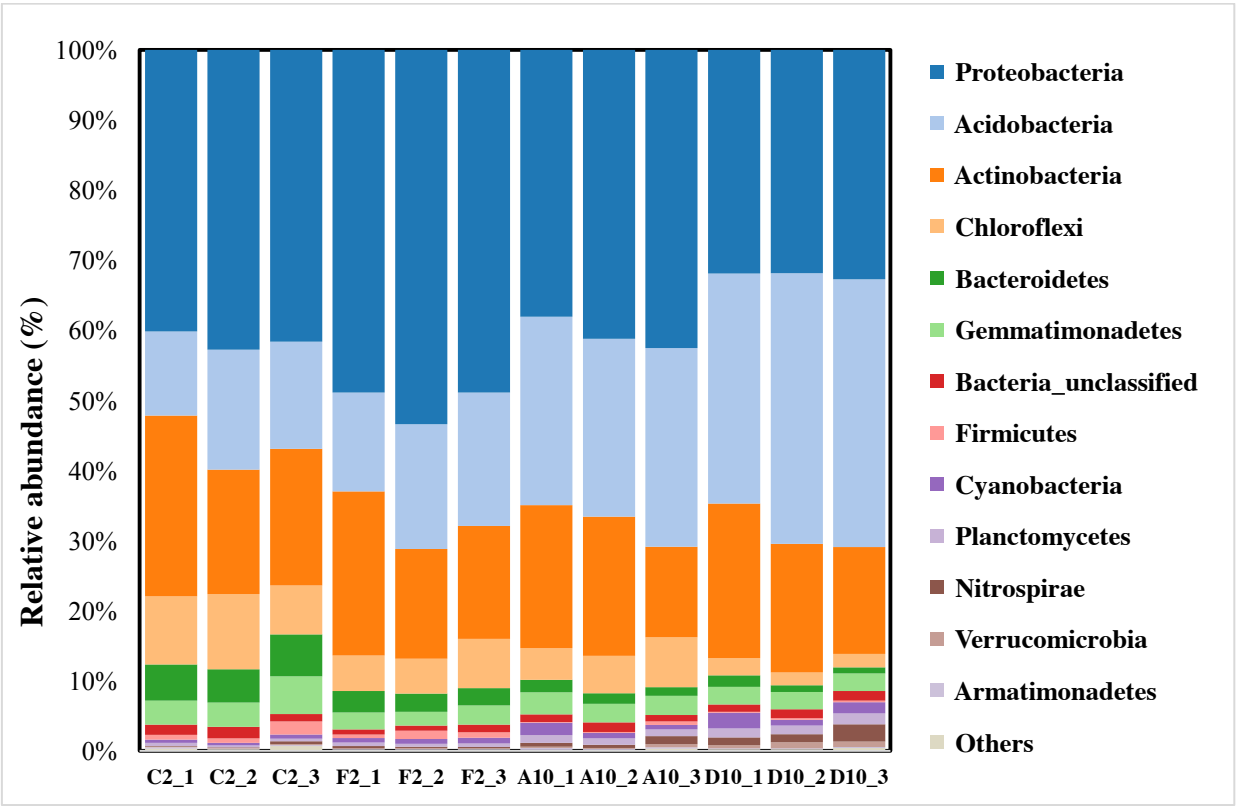

Fig. 5: Microbial community composition based on 16S rRNA gene sequencing for the most represented phyla and for the 4 plots studied at 3 different sampling times.