



HAL
open science

Small-Scale Variability in Bacterial Community Structure in Different Soil Types

Mylène Hugoni, Naoise Nunan, Jean Thioulouse, Audrey Dubost, Danis Abrouk, Jean M.F. Martins, Deborah Goffner, Claire Prigent-Combaret, Geneviève Grundmann

► **To cite this version:**

Mylène Hugoni, Naoise Nunan, Jean Thioulouse, Audrey Dubost, Danis Abrouk, et al.. Small-Scale Variability in Bacterial Community Structure in Different Soil Types. *Microbial ecology*, 2021, 82 (2), pp.470-483. 10.1007/s00248-020-01660-0 . hal-03763569

HAL Id: hal-03763569

<https://cnrs.hal.science/hal-03763569v1>

Submitted on 1 Sep 2022

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

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



HAL
open science

Small-Scale Variability in Bacterial Community Structure in Different Soil Types

Mylène Hugoni, Naoise Nunan, Jean Thioulouse, Audrey Dubost, Danis Abrouk, Jean M F Martins, Deborah Goffner, Claire Prigent-Combaret, Geneviève Grundmann

► **To cite this version:**

Mylène Hugoni, Naoise Nunan, Jean Thioulouse, Audrey Dubost, Danis Abrouk, et al.. Small-Scale Variability in Bacterial Community Structure in Different Soil Types. Microbial ecology, Springer Verlag, 2021, 10.1007/s00248-020-01660-0 . hal-03224926

HAL Id: hal-03224926

<https://hal-cnrs.archives-ouvertes.fr/hal-03224926>

Submitted on 17 May 2021

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.



Distributed under a Creative Commons Attribution| 4.0 International License

1 **Small scale variability in bacterial communities structure in different soil**
2 **types**

3
4 Mylène Hugoni¹, Naoise Nunan^{2,3}, Jean Thioulouse⁴, Audrey Dubost¹, Danis Abrouk¹, Jean
5 M.F. Martins⁵, Deborah Goffner⁶, Claire Prigent-Combaret¹, Geneviève Grundmann¹

6
7 ¹ Univ Lyon, Université Claude Bernard Lyon 1, CNRS, INRAE, VetAgro Sup, UMR Ecologie
8 Microbienne, F-69622 Villeurbanne, France

9 ² Institute of Ecology and Environmental Sciences – Paris, CNRS – Sorbonne Université, 4
10 place Jussieu, 75005 Paris, France

11 ³ Department of Soil and Environment, Swedish University of Agricultural Sciences, 75007
12 Uppsala, Sweden

13 ⁴ Biométrie et Biologie Evolutive, Université Lyon 1; CNRS UMR 5558; 69622 Villeurbanne
14 Cedex, France

15 ⁵ Université Grenoble Alpes; CNRS; IRD; IGE UMR 5001; 38000 Grenoble, France

16 ⁶ Unité Mixte Internationale CNRS 3189 « Environment, Health and Societies », Faculté de
17 Médecine, 51 Bd Pierre Dramard, 13344 Marseille, France

18
19 **Correspondance:** Mylène Hugoni, Université Lyon 1; CNRS, UMR5557; Ecologie
20 Microbienne; INRA, UMR1418; 69220 Villeurbanne Cedex, France. Email:
21 mylene.hugoni@univ-lyon1.fr ; ORCID number 0000-0002-2430-1057

22
23 **Keywords:** Bacterial diversity / Community structure / Soil / Spatial scale / Ecological
24 strategies

26 **Abstract**

27 Microbial spatial distribution has mostly been studied at field to global scales (*i.e.* ecosystem
28 scales). However, the spatial organization at small scales (*i.e.* centimeter to millimeter scales),
29 which can help improve our understanding of the impacts of spatial communities structure on
30 microbial functioning, has received comparatively little attention. Previous work has shown
31 that small scale spatial structure exists in soil microbial communities, but these studies have not
32 compared soils from geographically distant locations. Nor have they utilized community
33 ecology approaches, such as the core and satellite hypothesis and/or abundance-occupancy
34 relationships, often used in macro-ecology, to improve the description of the spatial
35 organization of communities. In the present work, we focused on bacterial diversity
36 (*i.e.* 16SrRNA gene sequencing) occurring in micro-samples from a variety of locations with
37 different pedo-climatic histories (*i.e.* from semi-arid, alpine and temperate climates) and
38 physicochemical properties. The forms of ecological spatial relationships in bacterial
39 communities (*i.e.* occupancy-frequency and abundance-occupancy) and taxa distributions (*i.e.*
40 habitat generalists and specialists) were investigated. The results showed that bacterial
41 composition differed in the four soils at the small scale. Moreover, one soil presented a satellite
42 mode distribution whereas the three others presented bimodal distributions. Interestingly,
43 numerous core taxa were present in the four soils among which 8 OTUs were common to the
44 four sites. These results confirm that analyses of the small-scale spatial distribution are
45 necessary to understand consequent functional processes taking place in soils, affecting thus
46 ecosystem functioning.

47

48

49

50 **1. Introduction**

51 Microbial diversity is central to soil ecosystem functioning and the ecosystem services soils
52 deliver, as microbial communities are involved in many biogeochemical processes and their
53 various interactions may affect these activities [1–3]. Despite this fundamental role, we still do
54 not have a full understanding of the processes underpinning the emergence and maintenance of
55 soil microbial diversity. Hubbell (2001) defined biodiversity as being synonymous with species
56 richness and relative species abundance in space and time, where relative species abundance
57 refers to the commonness or rarity of a species in relation to others, in a given community. The
58 question of what makes a species common or rare has long been of empirical and theoretical
59 interest in community ecology [4]. Despite a growing number of surveys interested in rare and
60 core microbial communities, only a few recent studies have targeted bacterial spatial
61 distributions, in a range of different ecosystems, such as water, guts, soil or even the surface of
62 stone [5–9]. These studies have highlighted the fact the “rare biosphere” is constituted of a
63 myriad of species. The rare communities are of interest as they may be of functional
64 significance: it has been shown that rare taxa are a resource pool for responding to
65 environmental changes [10], maintaining species diversity [11], and promoting functional
66 redundancy [12]. Other studies conducted on aquatic ecosystems at a global scale have shown
67 the presence of a rich and diverse rare biosphere among Archaea and Bacteria [13, 14]. Hugoni
68 *et al.* (2013) reported that the rare archaeal biosphere in the ocean should not solely be
69 characterized as a seed bank of dormant cells; rather, it is a complex association of indigenous
70 and itinerant cell types of different origins and with different fates that might contribute to
71 microbial interaction networks and metabolic processes in the environment.

72 Although these have yielded a host of information at larger scales, microbial
73 distribution patterns and their consequences for ecosystem functioning are far less documented
74 at small scales [15–17], especially in soils. However, it is clear that cell-to-cell, cell-substrate

75 and cell-substratum interactions that all take place at small scales, are likely to have significant
76 effects on overall processes [18, 19]. Previous studies have attempted to study bacterial
77 communities and soil properties at sub-millimeter scales, in different contexts as different soil
78 aggregate size classes, aggregate surfaces vs interiors, various soil fractions, or different soil
79 pore distributions [20–25]. These studies, using several methods such as fingerprinting [21],
80 clone libraries [24], pyrosequencing [20, 22, 25] or particulate organic matter characterization
81 (POM) [23] suggested that microaggregates represent micro-environments that select specific
82 bacterial lineages [26]. In order to evaluate the regulatory role of such interactions, the problem
83 of spatial patterns and scale needs to be addressed and models derived from the community
84 ecology of macro-organisms might give the tools to understand small-scale microbial
85 biodiversity.

86 The analysis of how biodiversity is distributed across space has employed various
87 spatial approaches in community ecology, based on habitat generalists and specialists [27, 28],
88 and on the core-satellite hypothesis [29, 30]. Classifying species into habitat generalists and
89 specialists is the first step toward examining the underlying biological and ecological factors
90 leading to the differential distribution of species among habitats [31]. Habitat generalists are
91 defined as broadly distributed microbial taxa whereas specialists are rare taxa that can be locally
92 abundant [28, 32]. The “core-satellite hypothesis” (or “Hanski’s hypothesis”) proposes that
93 species fall into one of two categories: “core” species, which are widespread and locally
94 abundant, and “satellite” species consisting mostly of rare species that are present in a limited
95 number of sites and at low abundances [29, 30]. Core and satellite taxa are classically identified
96 to predict communities’ occupancy-frequency distributions (with occupancy defined as the
97 percentage of samples in which a taxon is present and frequency defined as the percentage of
98 taxa found for each occupancy). This characterization of communities’ ecology in terms of
99 spatial distribution often showed a bimodal distribution (*i.e.* most species are present in either

100 most patches or only a small number of patches). In contrast, generalists and specialists refer to
101 the ability of a taxon to tolerate a range of habitats by considering the spatial distribution of
102 taxa. Previous work focusing on soil bacteria at the small scale showed that a occupancy-
103 frequency relationship was detectable at this small scale, following Hanski's core-satellite
104 hypothesis [8]. This work was done using microarray and 454 pyrosequencing data, using
105 samples of a few milligrams taken along a 22 cm long transect at one cm intervals. Here, we
106 extend this study to include a range of soils and use Illumina sequencing in order to derive an
107 more in depth analysis of the distribution of individual taxa. We chose to include a number of
108 soils with different abiotic properties because it is known that these properties influence the
109 composition and distribution of microbial communities.

110 Other works have attempted to identify which environmental variables shape microbial
111 diversity. Indeed, previous work reported a relationships between microbial distributions and
112 abiotic variables [33–35] and aimed to determine the main spatial community driver at a larger
113 scale.

114 The objective of this study was to investigate in different soils, at the small scale (from
115 1cm to 1m intervals and about 100 to 500mg weight) whether (1) soil bacterial communities
116 exhibit spatial structures and, if so, whether (2) these structures varied among soils from
117 different environments. We used Hanski's hypothesis (the core and satellite hypothesis) to
118 structure the analysis. The four soils that were used in the study were from four pedo-climatic
119 situations and from three different regions: two Alpine polygonal soils (one from the centers
120 and the other from the sides of polygons), a soil from the Rhône area and a Sahel soil. Bacterial
121 diversity analyses, evaluated using a 16S rRNA gene metabarcoding approach, were carried
122 out. Community ecology approaches were used to assess the spatial structures using occupancy-
123 frequency distribution patterns (*i.e.* core or satellite mode distributions) and abundance-
124 occupancy relationships (abundance being calculated as the average percent presence of a taxon

125 in samples). An in-depth analysis of bacterial OTU spatial distribution patterns was conducted
126 by studying core taxa (defined as OTUs present in 100% of the micro-samples for a considered
127 soil) and satellite taxa (defined as present in only 1 micro-sample for a considered soil, these
128 satellite taxa were therefore spatially rare taxa). Moreover, in order to gain a better
129 understanding of the ecological strategies employed by the taxa with regard to their habitat, we
130 classified them according to their distribution across micro-samples (generalist taxa defined as
131 broadly distributed bacterial taxa and specialists taxa were satellite taxa with more than 0.1%
132 of sequences per micro-sample).

133

134 **2. Materials and Methods**

135 ***2.1. Soil sampling***

136 Soils were sampled in regions presenting four pedo-climatic combinations in three geographic
137 locations: La Vanoise (Van), a loamy polygonal soil under alpine climate, sampled in the center
138 of polygons (Van_PC) and from the sides of polygons (Van_PS) (polygons being typical
139 structures retrieved in soil that experience freezing and thawing alternately); La Dombes (LD),
140 a silty loam soil under temperate climate and Grande Muraille Verte (GMV), a sandy soil in a
141 semi-arid climate.

142 Thirty nine micro-samples were collected at La Vanoise (col de la Laysse, La Vanoise,
143 France, 2200m altitude, N 45° 27' 19.56'', L 6° 51' 54.12''), characterized by an average
144 temperature in the summer of 5.3°C and an annual rainfall of 637mm. Seventeen micro-samples
145 were collected in Van_PC and 22 micro-samples were collected in Van_PS in September 2013
146 and September 2014 to take into account the soil convection movements typical of this type of
147 soil with polygons (Supplementary Table 1 [36]). Twenty-two LD soil micro-samples were
148 collected in an undisturbed agricultural soil in September 2013 (Saint André de Corcy, Ain,
149 France, 45° 51' 04.7'' N 4° 57' 30.5'') characterized by an average annual min-max

150 temperatures of 8.1°C-16.9°C and annual rain of 832mm. Finally, twenty soil micro-samples
151 were collected in September 2014 in a Savanna soil (Widou, Grande Muraille Verte, Senegal,
152 15°58.549'N 15°17.221'W). Maximum and minimum temperatures are 36.9°C and 20.5°C,
153 respectively, and the annual rainfall is 300mm.

154 The sampling procedure for the Van_PC, Van_PS and LD soils proceeded as follows:
155 soil cores, 1cm diameter and 2cm depth, were taken with a corer along a longitudinal transect,
156 with 1cm to 10m lags, 0.2 to 10m long (Supplementary Table 2). The top 1mm of the cores
157 (sample depth named “a”) and a 1 mm thick slice underneath (sample depth named “b”) were
158 cut with a scalpel blade. The soil water content allowed the sample to remain intact. Samples
159 were freeze-dried and stored at -80°C for DNA extraction. Two-centimeter depth core was an
160 intermediary size to sample in the field and to bring it back to lab and subsample the final
161 sample (*i.e.* the micro-sample) with spatial coordinate. Slice *a* was analyzed systematically and
162 slice *b* was taken in case there was a need to verify data.

163 In La Vanoise, micro-samples were collected along two 10m transects with lags from
164 1cm to 1m, to ensure that the sampling covered several polygons, from the centers (Van_PC,
165 yellow coloured) and sides of polygons (Van_PS, brown colored), respectively (Supplementary
166 Table 2). Micro-sample weights were 119 mg on average (min 70mg - max 288mg) for Van_PS
167 and 129 mg on average (from 120 to 150 mg) for Van_PC.

168 In LD, the micro-samples were collected 1cm lags along a 20m transect, from an area
169 at the edge of a maize field that had not been ploughed for at least 10 years. Micro-sample
170 weights were 100mg on average (77 to 144 mg).

171 For the GMV soil, the sampling procedure 1m lags along to a 20m transect, needed to
172 be adapted as it is a sandy soil and we could not proceed with cores. As the soil was very sandy,
173 10g of soil were sampled every meter at 20 spots in the field and then subsampled in the
174 laboratory (250 to 500 mg), lyophilized and stored at -80°C for DNA extraction.

175 Soil physicochemical characteristics (texture, pH, organic carbon, organic matter, C/N,
176 assimilable P, K, Ca, Mg) were measured by CESAR (Centre Scientifique Agricole Regional,
177 Ceyzeriat, France, <http://www.labo-cesar.com>) on pooled micro-samples from Van_PC,
178 Van_PS, GMV and LD to get average characteristics of the site (Supplementary Table 1) using
179 a standard methods that followed the AFNOR French standard (AFNOR, 2004).

180

181 **2.2. DNA extraction**

182 DNA extraction from micro-samples was performed following Michelland *et al.* (2016).
183 Briefly, soil micro-samples were suspended (1 g mL^{-1}) in PBS (pH=8) and centrifuged at 5,700g
184 for 1 min. The pellet was suspended in 1 mL washing solution (50 mmol L^{-1} Tris-HCl pH = 7.7,
185 25 mmol L^{-1} EDTA, 0.1 % SDS, 0.1 % PVP, H_2O) and centrifuged at 5700g for 1min. The pellet
186 was then suspended in $35 \mu\text{L}$ lysis solution (50 mmol L^{-1} Tris-HCl, pH = 8, 25 mmol L^{-1} EDTA,
187 3% SDS, 1.2% PVP) and microwaved at 600-700W for 45s (Orsini and Romano-Spica, 2001).
188 Four hundred microliters of extraction solution (10 mmol L^{-1} Tris HCl, pH =8, 1 mmol L^{-1}
189 EDTA, 0.3 mol L^{-1} sodium acetate, 1.2% PVP), pre-heated (75°C), were then added before
190 extraction with one volume phenol-choloroform-isoamylic alcohol (25:24:1) and centrifugated
191 for 10min at 13,000rpm. One volume of chloroform was then added, the solution centrifuged
192 for 5min at 13,000rpm, and the aqueous phase was retained. Ten percent of the total volume of
193 sodium acetate 3M (pH 5) and 2 volumes of cold absolute ethanol were added and, after cooling
194 on ice for 20min, the micro-samples were centrifuged for 30min at 13,000rpm and 4°C . The
195 supernatant was then discarded and the DNA pellets were washed with EtOH 70% and
196 resuspended in $20 \mu\text{L}$ water.

197

198 **2.3. Illumina sequencing and sequence processing**

199 Amplification of the bacterial V4 region of the 16S rRNA genes was performed using the
200 universal primers 515F/806R . High-throughput sequencing was carried out after a multiplexing
201 step using a MiSeq 2x300bp PE technology (MR DNA lab, www.mrdnalab.com, Shallowater,
202 TX, USA).

203 According to the MR DNA analysis pipeline (MR DNA, Shallowater, TX, USA), paired-
204 end sequences were merged and denoising procedures, that consisted in discarding reads
205 containing ambiguous bases (N) or reads that were outside the range of expected length (*i.e.*
206 <150bp), were carried out using Uchime [37]. The remaining sequences were clustered at a
207 97% similarity threshold [38] with Usearch [39]. Chimeras were detected using Uchime [37]
208 and removed from the dataset. Final OTUs were taxonomically affiliated using BLASTn
209 against a cured database derived from Greengenes, RDPII and NCBI (www.ncbi.nlm.nih.gov,
210 <http://rdp.cme.msu.edu>) [40].

211 Singletons were removed from the dataset. In order to compare micro-samples and to
212 retain the largest possible number of micro-samples, a random resampling normalization step
213 at a depth of 5474 sequences per sample was carried out. It should be noted that this procedure
214 resulted in the exclusion of 1 sample from the Van_PC and 3 samples from the GMV soils as
215 they did not contain 5474 sequences (Table 1).

216 The sequencing data of the LD soil have been published in Michelland *et al.* (2016) and
217 the dataset is available under the accession number PRJEB14534 in the SRA database. Data for
218 other soils have been deposited under the following accession number ERP016178 at EBI.

219

220 ***2.4. Statistical analyses***

221 The soil parameters (Supplementary Table 1, parameters in column) were analyzed by a
222 standardized PCA. Standardization was needed because of the differences in parameter units.
223 The data table was transformed to row percentages (row sums = 1) before doing the PCA, to

224 reduce the impact of the high sequence count variability among OTUs in the soil micro-
225 samples.

226 The table of diversity data (OTUs in columns) was analyzed by a PCoA on the Jaccard
227 distance matrix.

228 Computations were performed with the ade4 package [41] for the R Statistical Software
229 (R Core Team, 2019).

230

231 **3. Results**

232 ***3.1. Physico-chemical characteristics and bacterial diversity***

233 Rarefaction curves indicated that the sequencing effort was sufficient to describe bacterial
234 diversity in Van_PC and GMV soils while in Van_PS and LD, some samples do not appear to
235 have reached an asymptote (Supplementary Figure 1). Moreover, it showed that bacterial
236 richness in the Van_PC micro-samples was more variable than in the other soils
237 (Supplementary Figure 1). There were 418 OTUs common to the four soils, equivalent to
238 16.54% of the total richness (Figure 1A).

239 The PCA ordination, based on the physical and chemical properties of the soils (texture,
240 pH, organic carbon, organic matter, C/N, assimilable P, K, Ca, Mg, Supplementary Table 1), is
241 shown in Figure 2A. The percentage variability explained by axes 1 and 2 was 59% and 27%,
242 respectively. There was a clear separation among the soils, the LD soil being related to coarse
243 silt (Csi), K and assimilable P, the GMV soil related to coarse sand, the Van_PC soil to high
244 C/N ratios, and the Van_PS soil to clay and total nitrogen. Both La Vanoise soils (i.e. Van_PC
245 and Van_PS soils) were associated with high fine silt, OM, Mg, Ca contents. The pH of the La
246 Vanoise soils was close to or above 8 (Supplementary Table 1).

247 The PCoA ordination graph of bacterial diversity data is shown in Figure 2B. The first
248 two axes account for 11% and 9% of the total variability. This analysis highlighted that GMV
249 and LD soils exhibited lower intra-diversity compared to Van_PC and Van_PS soils. On the

250 first axis, GMV and LD soils were distinguished from the Van_PC and Van_PS soils while on
251 the second axis, GMV and LD soils were distinguished.

252

253 **3.2. Abundance-occupancy and occupancy-frequency relationships**

254 The abundance-occupancy diagrams for each soil showed a positive relationship between the
255 OTU average relative abundances in micro-samples and the proportion of micro-samples in
256 which they were found (Figure 3). The patterns in the occupancy-frequency diagrams were
257 different however (Figure 4). The distribution in Van_PC was unimodal with a satellite mode
258 (a large proportion of taxa occupied only one micro-sample), as highlighted by the ratio of
259 OTUs number_{core} over OTUs number_{satellite} equal to 0.027. Van_PS, GMV and LD presented
260 bimodal distributions with number_{core} OTUs: number_{satellite} OTUs ratios of 0.301, 0.323 and 0.571,
261 respectively. Interestingly, compared to Van_PC, Van_PS and GMV, a smaller proportion of
262 the taxa that were limited to a small number of micro-samples associated with a larger
263 proportion of taxa were present in all micro-samples is characteristic of the LD soil (*i.e.* ratio
264 of OTUs number_{core} over OTUs number_{satellite} of 0.571). Van_PS, GMV and LD core taxa
265 number are always important, 105, 86 and 155 respectively but they were less numerous than
266 the satellite taxa. The average abundances (*i.e.* % of sequences) of satellite OTUs in micro-
267 samples were higher in Van_PC soil than in Van_PS, GMV and LD soils (Table 1), whereas
268 the average abundances of core OTUs in Van_PC soil (*i.e.* 21.72%) were much lower compared
269 to other soils (80.96%, 78.80% and 82.78% for Van_PS, GMV and LD soils, respectively,
270 Table 1). Core OTUs in Van_PC soil exhibited the lowest frequency compared to the other soils
271 (7.65, 9.11 and 12.86% for Van_PS, GMV and LD soils, respectively, Table 1), while satellite
272 OTUs frequency in Van_PC soil was the highest (*i.e.* 36.80%) compared to the other soils
273 (25.34, 28.18 and 22.49% for Van_PS, GMV and LD soils, respectively, Table 1).

274 The PCoA ordination based on OTU patterns, which explained 11 and 9% of total variability
275 on the first and second axes, respectively, separated soils with positive coordinates for LD and
276 Van_PS soils and negative coordinates for GMV and Van_PC. Moreover, on the second axis
277 (Figure 2B) it separated LD and GMV. Moreover, the PCoA ordination based on OTU patterns
278 separated the four soils as the PCA ordination of the physico-chemical properties did (Figure
279 2). Thus, the spatial distribution type, which separates Van_PC (satellite mode) from Van_PS,
280 GMV and LD (bimodal), presents a typology that is different from the typology of soil physico-
281 chemical characteristics and to the typology based on diversity. The first axis of soil PCA
282 (which explains 59% total variability, Figure 2A) separates GMV and LD soils from Van_PS
283 and Van_PC. This corresponds to the opposition between semi-arid sandy soil of GMV and
284 organic matter rich soils from La Vanoise while LD is “positioned” by coarse silt. The second
285 axis (27% explained variability) separates LD and Van_PS from Van_PC and GMV. The
286 community spatial distribution type may thus be driven by a combination of factors different
287 from the ones driving diversity, thus bringing more information about the soil function.

288

289 ***3.3. Composition of core and satellite bacterial communities***

290 Taxonomic investigations, at the phylum_class level (representing > 2% of sequence
291 abundance) on satellite and core OTUs, revealed that some classes, as *Actinobacteria*,
292 *Alphaproteobacteria* and *Betaproteobacteria*, were ubiquitous (Figure 5). *Actinobacteria* were
293 more abundant among GMV and LD core OTUs than among satellite OTUs, and
294 *Deltaproteobacteria* were less abundant among core taxa in GMV and LD than among satellite
295 taxa. Globally, Bacilli and Clostridia represented, on average, 10.2 and 0.7% of all sequences,
296 respectively, in the satellite taxa of the four soils, but only 4.1 and 2.8 %, on average, in the
297 core taxa (to be noted that Bacilli were enriched in GMV core OTUs representing 13.9% of
298 sequences).

299 Core and satellite taxa in bacterial families that accounted for > 2% of the sequences
300 were also analyzed (Table 2). Among those 52 families, 28 were specific to core taxa while 14
301 families were specific of satellite taxa (Table 2). The core specific families represented 65.9,
302 43.3, 44.3 and 33.4% of core taxa sequences in Van_PC, Van_PS, GMV and LD, respectively,
303 while the satellite-specific families represented 40.2, 14.2, 19.3 and 7.5% of satellite taxa
304 sequences in Van_PC, Van_PS, GMV and LD, respectively.

305 Among the core taxa of the four soils (357 OTUs in total), only 2.2% (*i.e.* 8 OTUs) were
306 common to all of them (Figure 1B): *Acidimicrobiales* spp. (affiliated with Actinobacteria),
307 *Azospira* spp. (affiliated with Betaproteobacteria, Rhodocyclales order), *Flexibacter* spp.
308 (affiliated with Bacteroidetes, Cytophagales order), *Gemmatimonas* spp. (affiliated with
309 Gemmatimonadetes), *Nitrosococcus* spp. (affiliated with Gammaproteobacteria, Chromatiales
310 order), *Rubrobacter* spp (affiliated with Actinobacteria, Rubrobacterales order), *Sphaerobacter*
311 *thermophilus* (affiliated with Chloroflexi, Sphaerobacterales order), *Sphingomonas* spp.
312 (affiliated with Alphaproteobacteria, Sphingomonadales order). Moreover, Van_PS, GMV and
313 LD soils had 9.2%, 9.5% and 23.2% specific core taxa, respectively, whereas Van_PC
314 presented no specific core taxa (Table 1, Figure 1B). Relatively to common taxa retrieved in
315 the four soils (*i.e.* 418 OTUs, Figure 1A), core OTUs in Van_PC which is satellite mode
316 represented 2.4% of OTUs and in bimodal soils, Van_PS, GMV and LD soil it represented
317 24.1, 15.6% and 28.7% of OTUs respectively (Table 1).

318 In the same way, among the 1288 OTUs constituting the satellite taxa retrieved in the
319 four soils (Figure 1C), only 2 OTUs were common to the four soils (*i.e.* 0.16% of OTUs):
320 *Ancalomicrobium adetum* (affiliated with Alphaproteobacteria, Rhizobiales order) and
321 *Nonomuraea terrinata* (affiliated with Actinobacteria, Actinomycetales order). Moreover,
322 Van_PC, Van_PS, GMV and LD soils had 20.3%, 17.6, 13.9 and 14.4% specific satellite taxa,
323 respectively (Table 1, Figure 1C). Relatively to common taxa retrieved in the four soils (*i.e.*

324 418 OTUs, Figure 1A), satellite OTUs in Van_PC, Van_PS, GMV and LD represented 11.2,
325 3.8, 18.9 and 8.4% of OTUs, respectively (Table 1).

326

327 ***3.4. Bacterial ecological strategy at the small scale: habitat generalists and specialists***

328 Concerning generalists, among core taxa retrieved in the four studied soils (*i.e.* 229 OTUs),
329 34.7% were common to at least two soils, 20.2% common to at least 3 soils and only 3.5%
330 common to the four soils (Figure 1B).

331 Fifty-three specialist OTUs were found in Van_PC, representing 4.8% of the total richness in
332 Van_PC (Table 1). To be noted that in Van_PC, 45.3% of specialist taxa belonged to
333 Proteobacteria and 15.1% belonged to Actinobacteria. In GMV, 9 OTUs were considered
334 specialist, whilst there were 2 in Van_PS and only 1 in LD (Table 1). Interestingly, specialists
335 OTUs in LD and GMV are specific to these soils while one specialist OTU was common to
336 Van-PC and Van_PS soils.

337

338 **4. Discussion**

339 We use a community ecology approach, drawn from the study of macroorganisms, was
340 used to investigate bacterial community spatial distributions at the micro-scale across four
341 different soils. The micro-sampling (few hundred milligrams: from 100 to 500mg) was carried
342 out on undisturbed soil so as to take the spatial distribution of taxa in the micro-samples into
343 account (within and between micro-samples). The analysis of bacterial diversity data combined
344 with soil micro sampling suggested that bacteria can be classified in each soil as core or satellite
345 taxa, depending on their distribution at the small scale.

346

347 ***4.1. Spatial distribution patterns***

348 The separation of the soils on the two PCA and PCoA ordinations, suggests that satellite mode
349 distribution, retrieved in Van_PC, is related to coarse sand, high C/N and pH. Unfortunately,
350 we could not directly correlate bacterial taxa with environmental variables as micro-samples
351 were too small to perform physico-chemical analyses. To date, quantifying environmental
352 variables in micro-samples remains a big technical issue which needed to be resolved in further
353 decades to gain insight into relationships between microbial communities and fine soil physico-
354 chemical properties. Moreover, other complementary variables could be drivers of community
355 spatial structure such as bulk density, soil structural heterogeneity, pore size distributions and
356 connectivity and organic features (i.e. root architecture or soil invertebrate activity).

357 While the existence of a rare microbial biosphere, across different ecosystems, is now
358 accepted [43, 44], there is no widely accepted definition of rare taxa in the literature. Indeed,
359 arbitrary relative abundance cut-offs, depending on the study, range from 0.1% to 0.01% [45].
360 In the present work, conducted at the small scale, we used a spatial definition of rare taxa based
361 on micro-sample occupancy, where taxa found in only one micro-sample were considered
362 spatially rare, or satellite taxa [32]). This corresponds to an average relative abundance of
363 0.027% of the total sequences, when including all micro-samples in this study. Compared to
364 the definition of Lynch and Neufeld (2015), our percentage of rare taxa identified is quite
365 conservative. Because rare taxa are substantial contributors to bacterial community ecology
366 [43, 45], their spatial distribution may be an important factor impacting ecological processes
367 and thus soil functioning. Studies focusing on spatial ecology of soil bacteria [46] concluded
368 that the relative importance of the underlying processes contributing to the establishment of
369 bacterial distributions can change with soil characteristics. More generally, Lindh *et al.* (2017)
370 suggested that analyses of occupancy-frequency patterns can be a highly valuable approach
371 allowing the definition of microbial biomes across environmental gradients [7].

372 Although a positive small scale abundance-occupancy relationship has already been
373 identified in the LD soil [8], this study confirmed this relationship and extended it to three other
374 soils, suggesting a certain prevalence of the relationship in soil microbial communities. This is
375 hardly surprising as it is one of the most fundamental patterns in ecology [2]. Moreover, the
376 bacterial community spatial distributions also conformed to Hanski's core and satellite
377 hypothesis [29]. However, the present work was conducted in four soils for which samples
378 differed a bit in the size and/or mass collected as there was a loamy, a silty loam and a sandy
379 soil that do not allow the strictly same sampling procedure. Soil heterogeneity and thus, the
380 requirement for different experimental procedures may have an impact on the results.

381 Gleason (1929) showed that quadrats of the most serviceable size should be chosen to
382 measure frequency distributions of plants, and that the information is obscured or lost if the
383 quadrats are either too large or too small [47]. The fact that different ecological strategies were
384 identifiable here suggests that the sample size used was appropriate for the study of bacterial
385 spatial distributions in soils [16]. Compared to other studies, the present work showed the
386 relevance of the small scale approach in spite of the very low organism's size to sample size
387 ratio [48]. Previous studies investigating the type of occupancy-frequency patterns have
388 suggested that unimodal distributions, that are characterized by an excess of rare, endemic
389 species, occur more frequently than bimodal distributions [49, 50]. Even when bimodal
390 distributions are detected, the magnitude of the core mode, representing cosmopolitan species,
391 is generally smaller than that of the rare species [50]. The three bimodal distributions that we
392 observed, indicate that it may be a common community spatial distribution of soil bacteria at
393 the small scale. This suggested that OTUs are either adapted to the majority of micro-habitats
394 in this soil, or that these OTUs have a high dispersal capabilities favouring the colonization of
395 empty ecological niches. Indeed, core species are believed to owe their regional commonness
396 to the rescue effect (*i.e.* reduction in local extinction probability [51]) while satellite species are

397 maintained by the habitat heterogeneity [52]. Cadotte and Lovett-Douste (2007) suggested that,
398 via the rescue effect, widespread and abundant species should have reduced risk of local
399 extinction. In their work conducted on trees, different community spatial structures were
400 observed [53]. Indeed, trees in degraded environments showed a satellite mode distribution,
401 contrary to the other studied environments. In the Van_PC soil, the satellite mode distribution
402 might represent a functional mode (*i.e.* corresponding to a stressful state) as in Cadotte and
403 Lovett-Douste (2007).

404

405 ***4.2. Bacterial ecological strategies***

406 For some authors, identifying core taxa is the first step in defining a “healthy” community and
407 predicting community responses to perturbation [54]. Interestingly, it has been suggested that
408 due to the great microbial diversity among ecosystems, the possibility of any species being
409 present at high abundance in a large range of samples is low and it is possible that the focus
410 should instead be on higher taxonomic levels or on functional genes [55]. The results presented
411 here are in accordance with this theory as only a few core OTUs are common to the four soils.
412 However, when considering the commonness in only two or three soils, their richness increases,
413 suggesting a patchy distribution rather than a ubiquitous one.

414

415

416

417 **5. Conclusions**

418 There is an increasing recognition that interactions within bacterial communities can have an
419 impact on community functioning [56, 57]. In this contribution we have shown that the types
420 of small-scale distributions of bacterial communities vary across soils with different physico-
421 chemical features. This suggests that the type and range of interactions that might be expected
422 to occur within microbial communities are likely to vary depending on the pedo-climatic

423 context. We have shown that one soil presented a core mode, another presented a satellite mode
424 and the three others presented bimodal distributions. We also identify among bacterial richness
425 retrieved in those four soils micro-samples that up to 13% can be classified as core taxa, that
426 could be responsible for the stability of soil bacterial communities. Thus the considering of
427 microbial spatial distribution, that could play a major ecological role, should further improve
428 predictions and modelling on diversity modification risks, on soil bio-conservation and
429 epidemiology.

430 **Conflicts of interest**

431 The authors declare that they have no known competing financial interests or personal
432 relationships that could have appeared to influence the work reported in this paper.

433

434 **Funding**

435 This work was cofunded by the EC2CO MicrobiEn AO2012- 779949 'Les communautés
436 bactériennes dans les sols extrêmes: les paramètres de leur structure et leur composition' and
437 the Labex DRIIHM, French program "Investissements d'Avenir" (ANR-11-LABX-0010)
438 which is managed by the ANR. We thank La Vanoise National Parc for soil sampling
439 authorization.

440

441 **References**

- 442 1. Delgado-Baquerizo M, Maestre FT, Reich PB, et al (2016) Microbial diversity drives
443 multifunctionality in terrestrial ecosystems. *Nat Commun* 7:10541.
444 <https://doi.org/10.1038/ncomms10541>
- 445 2. Gaston KJ, Blackburn TM, Greenwood JJD, et al (2000) Abundance–occupancy
446 relationships. *J Appl Ecol* 37:39–59. <https://doi.org/10.1046/j.1365-2664.2000.00485.x>
- 447 3. Graham EB, Knelman JE, Schindlbacher A, et al (2016) Microbes as Engines of
448 Ecosystem Function: When Does Community Structure Enhance Predictions of Ecosystem
449 Processes? *Front Microbiol* 7:. <https://doi.org/10.3389/fmicb.2016.00214>
- 450 4. Hubbell SP (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*
451 (MPB-32). Princet Univ Press

- 452 5. Cariveau DP, Elijah Powell J, Koch H, et al (2014) Variation in gut microbial
453 communities and its association with pathogen infection in wild bumble bees (*Bombus*). *ISME*
454 *J* 8:2369–2379. <https://doi.org/10.1038/ismej.2014.68>
- 455 6. Hugoni M, Escalas A, Bernard C, et al (2018) Spatiotemporal variations in microbial
456 diversity across the three domains of life in a tropical thalassohaline lake (Dziani Dzaha,
457 Mayotte Island). *Mol Ecol* 27:4775–4786
- 458 7. Lindh MV, Sjöstedt J, Ekstam B, et al (2017) Metapopulation theory identifies
459 biogeographical patterns among core and satellite marine bacteria scaling from tens to
460 thousands of kilometers. *Environ Microbiol* 19:1222–1236. [https://doi.org/10.1111/1462-](https://doi.org/10.1111/1462-2920.13650)
461 [2920.13650](https://doi.org/10.1111/1462-2920.13650)
- 462 8. Michelland R, Thioulouse J, Kyselková M, Grundmann GL (2016) Bacterial
463 Community Structure at the Microscale in Two Different Soils. *Microb Ecol* 72:717–724.
464 <https://doi.org/10.1007/s00248-016-0810-0>
- 465 9. Cutler NA, Chaput DL, Oliver AE, Viles HA (2015) The spatial organization and
466 microbial community structure of an epilithic biofilm. *FEMS Microbiol Ecol* 91:.
467 <https://doi.org/10.1093/femsec/fiu027>
- 468 10. Jones SE, Lennon JT (2010) Dormancy contributes to the maintenance of microbial
469 diversity. *Proc Natl Acad Sci U S A* 107:5881–5886. <https://doi.org/10.1073/pnas.0912765107>
- 470 11. Dohrmann AB, Küting M, Jünemann S, et al (2013) Importance of rare taxa for bacterial
471 diversity in the rhizosphere of Bt- and conventional maize varieties. *ISME J* 7:37–49.
472 <https://doi.org/10.1038/ismej.2012.77>
- 473 12. Louca S, Polz MF, Mazel F, et al (2018) Function and functional redundancy in
474 microbial systems. *Nat Ecol Evol* 2:936–943. <https://doi.org/10.1038/s41559-018-0519-1>
- 475 13. Galand PE, Casamayor EO, Kirchman DL, Lovejoy C (2009) Ecology of the rare
476 microbial biosphere of the Arctic Ocean. *Proc Natl Acad Sci U S A* 106:22427–22432.
477 <https://doi.org/10.1073/pnas.0908284106>
- 478 14. Hugoni M, Taib N, Debros D, et al (2013) Structure of the rare archaeal biosphere and
479 seasonal dynamics of active ecotypes in surface coastal waters. *Proc Natl Acad Sci U S A*
480 110:6004–6009. <https://doi.org/10.1073/pnas.1216863110>
- 481 15. Franklin RB, Mills AL (2003) Multi-scale variation in spatial heterogeneity for
482 microbial community structure in an eastern Virginia agricultural field. *FEMS Microbiol Ecol*
483 44:335–346. [https://doi.org/10.1016/S0168-6496\(03\)00074-6](https://doi.org/10.1016/S0168-6496(03)00074-6)
- 484 16. Grundmann GL (2004) Spatial scales of soil bacterial diversity--the size of a clone.
485 *FEMS Microbiol Ecol* 48:119–127. <https://doi.org/10.1016/j.femsec.2004.01.010>
- 486 17. Morris SJ, Boerner REJ (1999) Spatial distribution of fungal and bacterial biomass in
487 southern Ohio hardwood forest soils: scale dependency and landscape patterns. *Soil Biol*
488 *Biochem* 31:887–902. [https://doi.org/10.1016/S0038-0717\(99\)00002-4](https://doi.org/10.1016/S0038-0717(99)00002-4)
- 489 18. Wilpiseski RL, Aufrecht JA, Retterer ST, et al (2019) Soil Aggregate Microbial
490 Communities: Towards Understanding Microbiome Interactions at Biologically Relevant
491 Scales. *Appl Environ Microbiol* 85:.
492 <https://doi.org/10.1128/AEM.00324-19>
- 493 19. Nunan N, Schmidt H, Raynaud X (2020) The ecology of heterogeneity: soil bacterial
494 communities and C dynamics. *Philos Trans R Soc Lond B Biol Sci* 375:20190249.
495 <https://doi.org/10.1098/rstb.2019.0249>
- 496 20. Bailey VL, McCue LA, Fansler SJ, et al (2013) Micrometer-scale physical structure and
497 microbial composition of soil macroaggregates. *Soil Biol Biochem* 65:60–68.
498 <https://doi.org/10.1016/j.soilbio.2013.02.005>
- 499 21. Blaud A, Chevallier T, Virto I, et al (2014) Bacterial community structure in soil
500 microaggregates and on particulate organic matter fractions located outside or inside soil
501 macroaggregates. *Pedobiologia* 57:191–194. <https://doi.org/10.1016/j.pedobi.2014.03.005>
- 502 22. Davinic M, Fultz LM, Acosta-Martinez V, et al (2012) Pyrosequencing and mid-

502 infrared spectroscopy reveal distinct aggregate stratification of soil bacterial communities and
503 organic matter composition. *Soil Biol Biochem* 46:63–72.
504 <https://doi.org/10.1016/j.soilbio.2011.11.012>

505 23. Kravchenko AN, Negassa W, Guber AK, Schmidt S (2014) New Approach to Measure
506 Soil Particulate Organic Matter in Intact Samples Using X-ray Computed Microtomography.
507 *Soil Sci Soc Am J* 78:1177–1185. <https://doi.org/10.2136/sssaj2014.01.0039>

508 24. Mummey D, Holben W, Six J, Stahl P (2006) Spatial Stratification of Soil Bacterial
509 Populations in Aggregates of Diverse Soils. *Microb Ecol* 51:404–411.
510 <https://doi.org/10.1007/s00248-006-9020-5>

511 25. Negassa WC, Guber AK, Kravchenko AN, et al (2015) Properties of Soil Pore Space
512 Regulate Pathways of Plant Residue Decomposition and Community Structure of Associated
513 Bacteria. *PLOS ONE* 10:e0123999. <https://doi.org/10.1371/journal.pone.0123999>

514 26. Rillig MC, Muller LA, Lehmann A (2017) Soil aggregates as massively concurrent
515 evolutionary incubators. *ISME J* 11:1943–1948. <https://doi.org/10.1038/ismej.2017.56>

516 27. Barberán A, Bates ST, Casamayor EO, Fierer N (2012) Using network analysis to
517 explore co-occurrence patterns in soil microbial communities. *ISME J* 6:343–351.
518 <https://doi.org/10.1038/ismej.2011.119>

519 28. Pandit SN, Kolasa J, Cottenie K (2009) Contrasts between habitat generalists and
520 specialists: an empirical extension to the basic metacommunity framework. *Ecology* 90:2253–
521 2262. <https://doi.org/10.1890/08-0851.1>

522 29. Hanski I (1982) Dynamics of Regional Distribution: The Core and Satellite Species
523 Hypothesis. *Oikos* 38:210–221. <https://doi.org/10.2307/3544021>

524 30. Ulrich W, Zalewski M (2006) Abundance and co-occurrence patterns of core and
525 satellite species of ground beetles on small lake islands. *Oikos* 114:338–348.
526 <https://doi.org/10.1111/j.2006.0030-1299.14773.x>

527 31. Chazdon RL, Chao A, Colwell RK, et al (2011) A novel statistical method for
528 classifying habitat generalists and specialists. *Ecology* 92:1332–1343.
529 <https://doi.org/10.1890/10-1345.1>

530 32. Monard C, Gantner S, Bertilsson S, et al (2016) Habitat generalists and specialists in
531 microbial communities across a terrestrial-freshwater gradient. *Sci Rep* 6:37719.
532 <https://doi.org/10.1038/srep37719>

533 33. Bahram M, Hildebrand F, Forslund SK, et al (2018) Structure and function of the global
534 topsoil microbiome. *Nature* 560:233–237. <https://doi.org/10.1038/s41586-018-0386-6>

535 34. Wang C, Michalet R, Liu Z, et al (2020) Disentangling Large- and Small-Scale Abiotic
536 and Biotic Factors Shaping Soil Microbial Communities in an Alpine Cushion Plant System.
537 *Front Microbiol* 11:. <https://doi.org/10.3389/fmicb.2020.00925>

538 35. Martínez-Olivas MA, Jiménez-Bueno NG, Hernández-García JA, et al (2019) Bacterial
539 and archaeal spatial distribution and its environmental drivers in an extremely haloalkaline soil
540 at the landscape scale. *PeerJ* 7:e6127. <https://doi.org/10.7717/peerj.6127>

541 36. Taş N, Prestat E, Wang S, et al (2018) Landscape topography structures the soil
542 microbiome in arctic polygonal tundra. *Nat Commun* 9:777. <https://doi.org/10.1038/s41467-018-03089-z>

543 37. Edgar RC, Haas BJ, Clemente JC, et al (2011) UCHIME improves sensitivity and speed
544 of chimera detection. *Bioinformatics* 27:2194–2200.
545 <https://doi.org/10.1093/bioinformatics/btr381>

546 38. Kim M, Morrison M, Yu Z (2011) Evaluation of different partial 16S rRNA gene
547 sequence regions for phylogenetic analysis of microbiomes. *J Microbiol Methods* 84:81–87.
548 <https://doi.org/10.1016/j.mimet.2010.10.020>

549 39. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST.
550 *Bioinformatics* btq461. <https://doi.org/10.1093/bioinformatics/btq461>

551

- 552 40. DeSantis TZ, Hugenholtz P, Larsen N, et al (2006) Greengenes, a Chimera-Checked
553 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl Environ Microbiol*
554 72:5069–5072. <https://doi.org/10.1128/AEM.03006-05>
- 555 41. Thioulouse J, Dray S, Dufour A-B, et al (2018) *Multivariate Analysis of Ecological*
556 *Data with ade4*. Springer-Verlag, New York
- 557 42. R Core Team (2016) *R: A Language and Environment for Statistical Computing*. R
558 Foundation for Statistical Computing, Vienna, Austria
- 559 43. Pedrós-Alió C (2012) The rare bacterial biosphere. *Annu Rev Mar Sci* 4:449–466.
560 <https://doi.org/10.1146/annurev-marine-120710-100948>
- 561 44. Sogin ML, Morrison HG, Huber JA, et al (2006) Microbial diversity in the deep sea and
562 the underexplored “rare biosphere.” *Proc Natl Acad Sci* 103:12115–12120.
563 <https://doi.org/10.1073/pnas.0605127103>
- 564 45. Lynch MDJ, Neufeld JD (2015) Ecology and exploration of the rare biosphere. *Nat Rev*
565 *Microbiol* 13:217–229. <https://doi.org/10.1038/nrmicro3400>
- 566 46. Raynaud X, Nunan N (2014) Spatial Ecology of Bacteria at the Microscale in Soil.
567 *PLOS ONE* 9:e87217. <https://doi.org/10.1371/journal.pone.0087217>
- 568 47. Gleason HA (1929) The Significance of Raunkiaer’s Law of Frequency. *Ecology*
569 10:406–408. <https://doi.org/10.2307/1931149>
- 570 48. Fierer N, Lennon JT (2011) The generation and maintenance of diversity in microbial
571 communities. *Am J Bot* 98:439–448. <https://doi.org/10.3732/ajb.1000498>
- 572 49. Levin SA (1974) Dispersion and Population Interactions. *Am Nat* 108:207–228.
573 <https://doi.org/10.1086/282900>
- 574 50. Tokeshi M (1992) Dynamics of distribution in animal communities: Theory and
575 analysis. *Res Popul Ecol* 34:249–273. <https://doi.org/10.1007/BF02514796>
- 576 51. Eriksson A, Elías-Wolff F, Mehlig B, Manica A (2014) The emergence of the rescue
577 effect from explicit within- and between-patch dynamics in a metapopulation. *Proc R Soc B*
578 *Biol Sci* 281:20133127. <https://doi.org/10.1098/rspb.2013.3127>
- 579 52. Czaran T (1998) *Spatiotemporal Models of Population and Community Dynamics*.
580 Springer US
- 581 53. Cadotte MW, Lovett-Doust J (2007) Core and Satellite Species in Degraded Habitats:
582 an Analysis Using Malagasy Tree Communities. *Biodivers Conserv* 16:2515–2529.
583 <https://doi.org/10.1007/s10531-006-9027-8>
- 584 54. Shade A, Handelsman J (2012) Beyond the Venn diagram: the hunt for a core
585 microbiome. *Environ Microbiol* 14:4–12. <https://doi.org/10.1111/j.1462-2920.2011.02585.x>
- 586 55. Hamady M, Knight R (2009) Microbial community profiling for human microbiome
587 projects: Tools, techniques, and challenges. *Genome Res* 19:1141–1152.
588 <https://doi.org/10.1101/gr.085464.108>
- 589 56. Pande S, Kost C (2017) Bacterial Unculturability and the Formation of Intercellular
590 Metabolic Networks. *Trends Microbiol* 25:349–361. <https://doi.org/10.1016/j.tim.2017.02.015>
- 591 57. Pascual-García A, Bonhoeffer S, Bell T (2020) Metabolically cohesive microbial
592 consortia and ecosystem functioning. *Philos Trans R Soc B Biol Sci* 375:20190245.
593 <https://doi.org/10.1098/rstb.2019.0245>

594

595

596

597

598 **Figures and Tables Legends**

599 **Figure 1.** Venn Diagrams presenting the shared and specific OTUs (A) detected in the four
600 soils, (B) among core OTUs and (C) among satellite OTUs.

601 **Figure 2.** PCA map of soils characteristics (A) and PCoA map of bacterial diversity (B) for
602 LD, GMV, Van_PC, Van_PS soils. On Figure 2A, arrows correspond to the four soil samples
603 and labels to the soil physico-chemical parameters. On Figure 2B, the position of ellipse centers
604 are given by the means of the coordinates of micro-samples on the two axes of the factor map
605 for each soil sample origins (LD, GMV, Van_PC and Van_PS). The width and height of ellipses
606 are given by the variance of the coordinates of the micro-samples on the two axes and the ellipse
607 slope is equal to their covariance.

608 **Figure 3.** Relationship between relative abundance (average percentage taxa presence in micro-
609 samples) and occupancy (percentage of micro-samples in which each taxa is present) in
610 Van_PC, Van_PS, LD and GMV soils.

611 **Figure 4.** Occupancy-frequency relationship for Van_PC, Van_PS, LD and GMV soils.
612 Frequency is expressed as the percent taxa found for each possible occupancy.

613 **Figure 5.** Distribution of bacterial classes representing more than 2% of total taxa in core and
614 satellite taxa, in Van_PC, Van_PS, LD and GMV soils.

615

616 **Table 1.** Sequence analyses for Van_PC, Van_PS, LD and GMV soils: sample number, raw
617 sequence number and normalization threshold. Detailed analyses of core and satellite taxa:
618 OTUs number, sequences percentage, frequency, average abundance. Detailed analyses of
619 ecological strategies: specific core OTUs percentage, percentage of core OTUs found in
620 common taxa and number of specialist OTUs.

621 **Table 2.** Distribution of bacterial families representing more than 2% of sequences in at least
622 one of the condition (core or satellite of each site). The ten most abundant families per
623 conditions were highlighted in grey.

624 **Supplementary Figure 1.** Rarefaction curves for Van_PC, Van_PS, LD and GMV soils, using
625 the normalized datasets (*i.e.* 5474 sequences).

626 **Supplementary Table 1.** Physico-chemical characteristics of Van_PC, Van_PS, LD and GMV
627 soils. The unit is g.kg⁻¹ except for pH and C/N ratio.

628 **Supplementary Table 2.** La Vanoise soil sampling: Distances between micro-samples (cores)
629 in Van_PC and in Van_PS starting from an initial point, named 0 m. Micro-samples were taken
630 along two transects 10 cm apart, in September 2013 and September 2014. Results obtained on
631 all micro-samples were gathered as no difference could be shown between the two sampling
632 dates. Micro-samples named “a” corresponded to the top 1 mm of the cores and sample “b”
633 corresponded to the 1 to 2 mm below sample “a”. “b*”: in two cases, a second core was taken
634 beside core “b”, at the same distance from the origin.

635

636

637

638

639

640

641

642

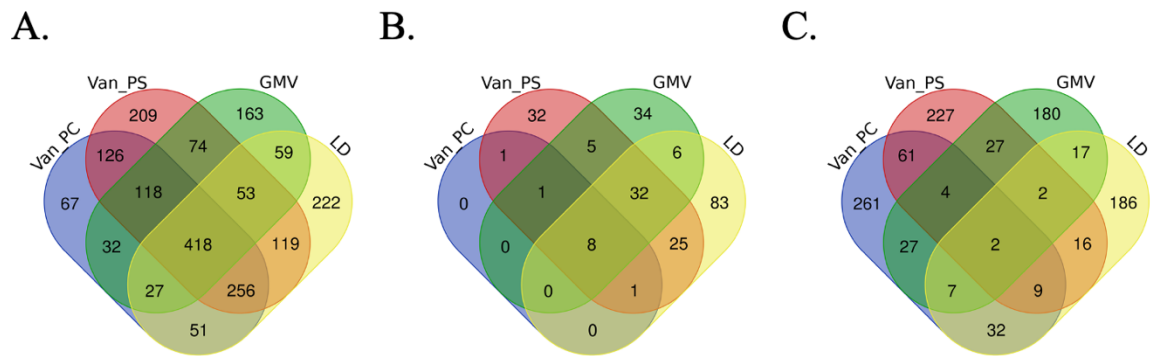
643

644

645

646

647



649

650

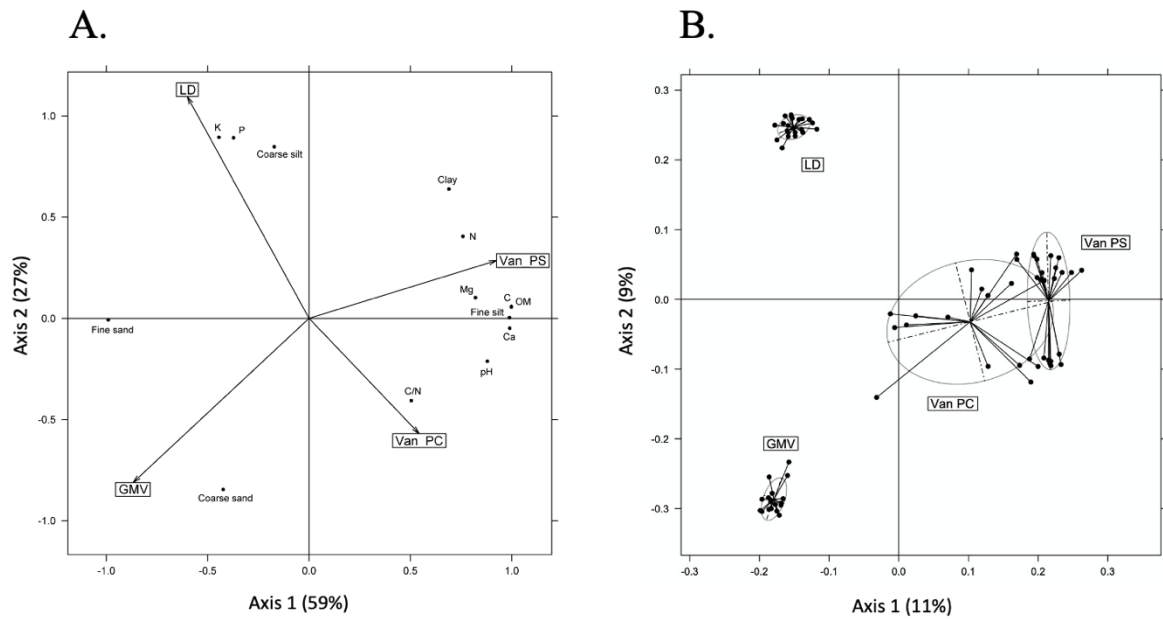
651

652 **Figure 1.** Venn Diagrams presenting the shared and specific OTUs (A) detected in the four
 653 soils, (B) among core OTUs and (C) among satellite OTUs.

654

655

656



657

658

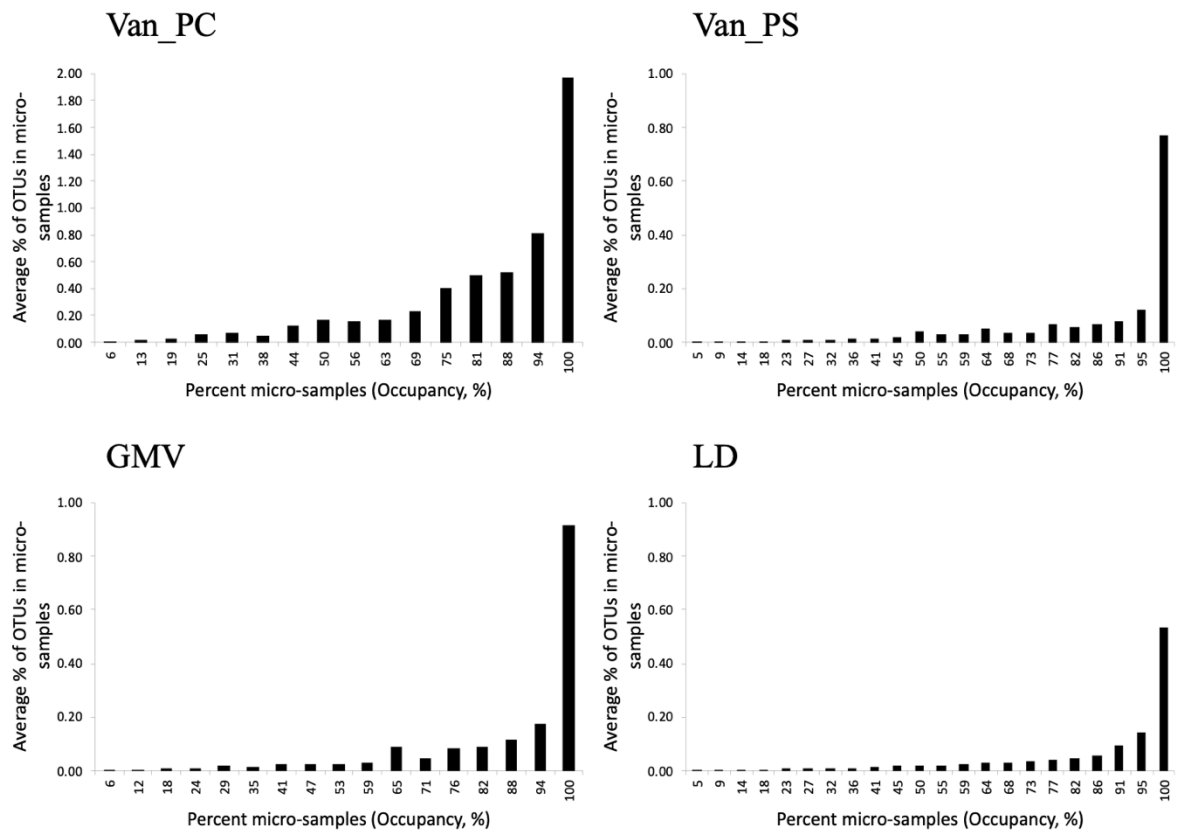
659 **Figure 2.** PCA map of soils characteristics (A) and PCoA map of bacterial diversity (B) for
 660 LD, GMV, Van_PC, Van_PS soils. On Figure 2A, arrows correspond to the four soil samples
 661 and labels to the soil physico-chemical parameters. On Figure 2B, the position of ellipse centers
 662 are given by the means of the coordinates of micro-samples on the two axes of the factor map
 663 for each soil sample origins (LD, GMV, Van_PC and Van_PS). The width and height of ellipses
 664 are given by the variance of the coordinates of the micro-samples on the two axes and the ellipse
 665 slope is equal to their covariance.

666

667

668

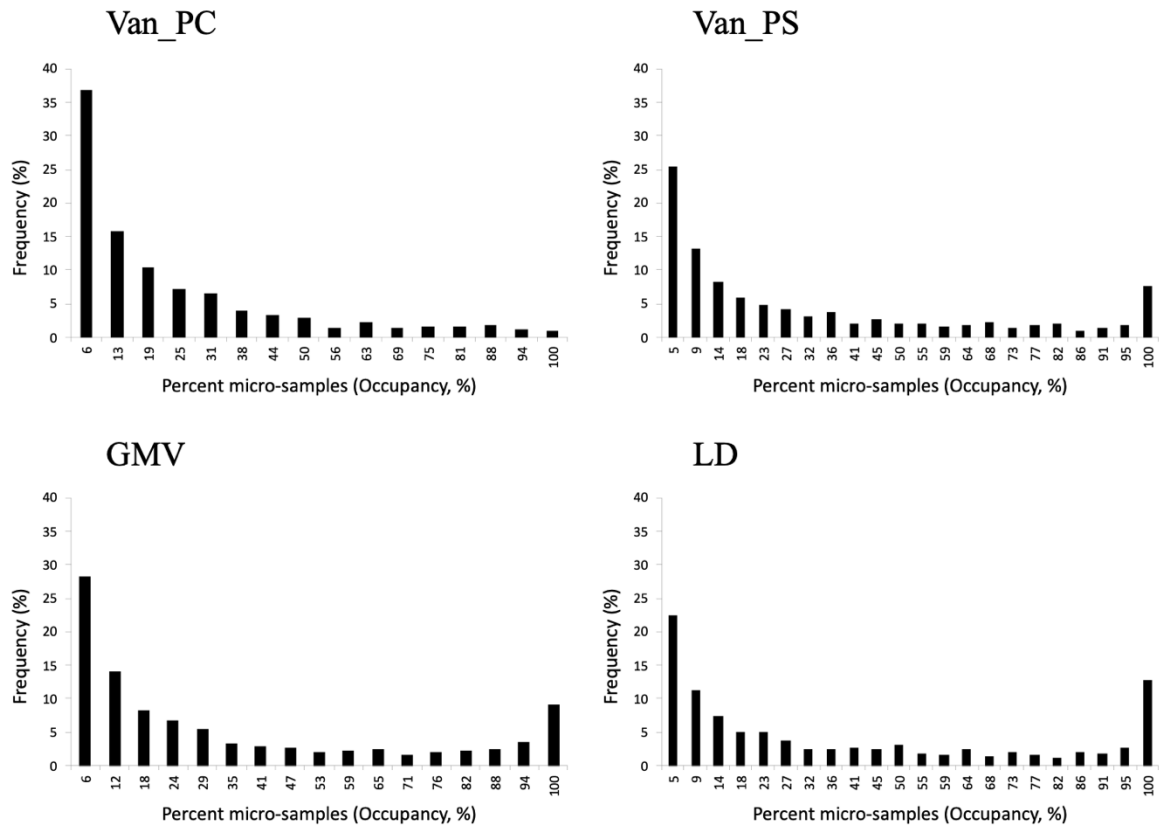
669



670

671 **Figure 3.** Relationship between relative abundance (average percentage taxa presence in micro-
 672 samples) and occupancy (percentage of micro-samples in which each taxa is present) in
 673 Van_PC, Van_PS, LD and GMV soils.

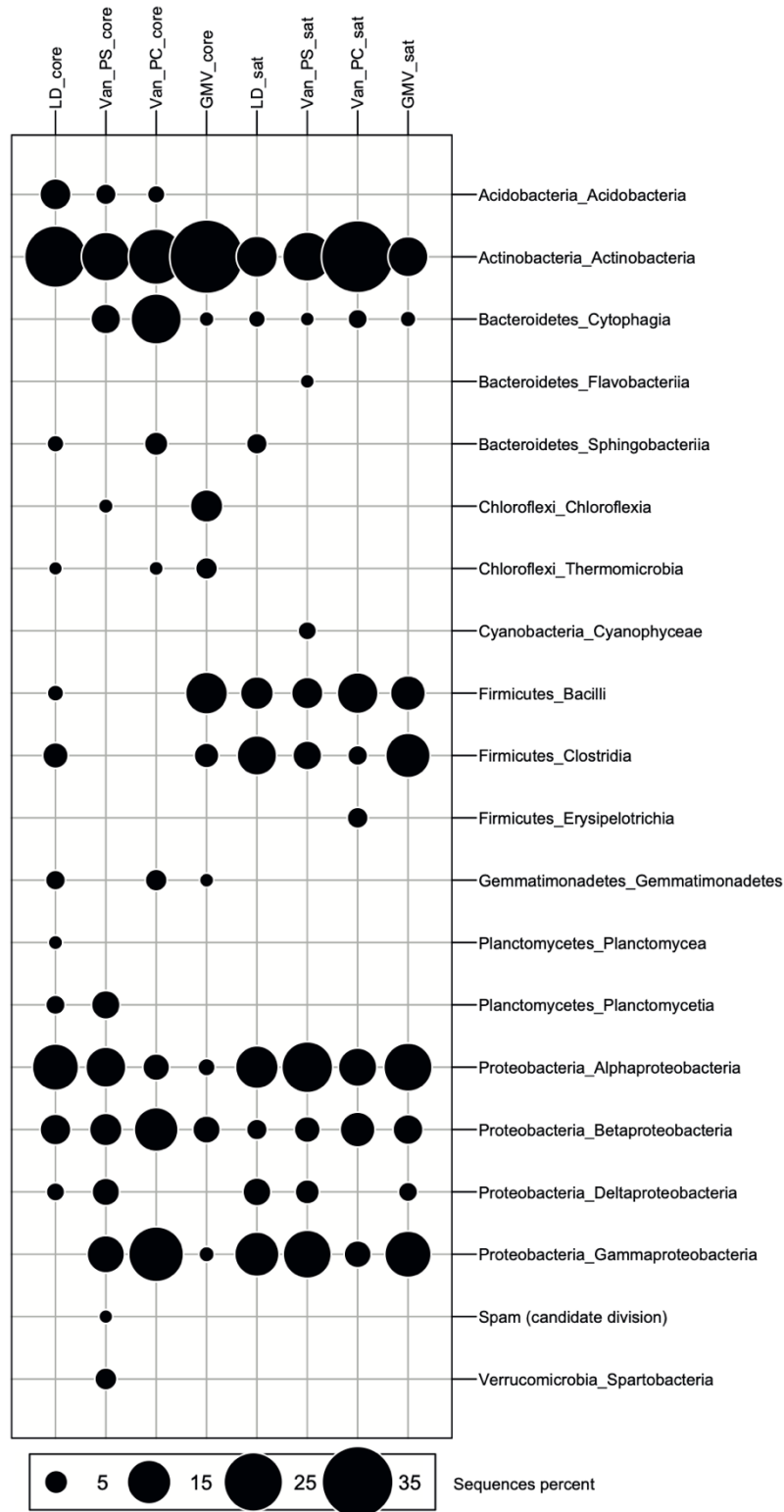
674



675

676 **Figure 4.** Occupancy-frequency relationship for Van_PC, Van_PS, LD and GMV soils.

677 Frequency is expressed as the percent taxa found for each possible occupancy.



678

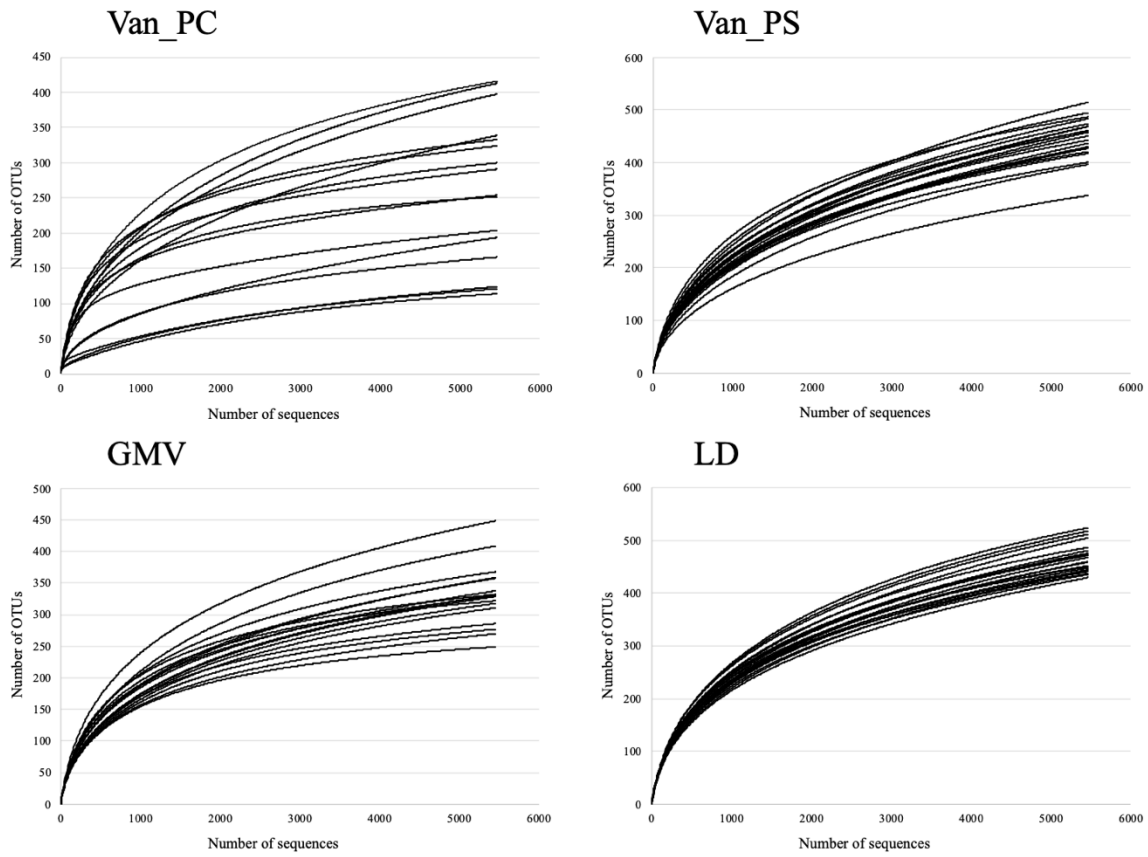
679 **Figure 5.** Distribution of bacterial classes representing more than 2% of total taxa in core and
 680 satellite taxa, in Van_PC, Van_PS, LD and GMV soils.

681

682

683

684



685

686 **Supplementary Figure 1.** Rarefaction curves for Van_PC, Van_PS, LD and GMV soils, using

687 the normalized datasets (*i.e.* 5474 sequences).

688

689

690

691

692

693

694

695 **Supplementary Table 1.** Physico-chemical characteristics of Van_PC, Van_PS, LD and GMV
 696 soils. The unit is g.kg⁻¹ except for pH and C/N ratio.

	LD	Van_PS	Van_PC	GMV
Coarse sand	40	127	179	437
Fine sand	392	115	140	476
Coarse silt	337	105	149	38
Fine silt	124	483	458	13
Clay	107	170	74	35
pH	6.82	7.77	8.41	6.61
Organic C	9.1	40.6	30.3	1.9
Organic matter	15.6	69.9	52.2	3.4
Total N	0.9	2.9	0.7	0.2
C/N	10	13	45	9
Assimilable P	0.261	0.032	0.013	0
K	0.138	0.056	0.015	0.044
Ca	1.94	9.54	9.1	0.33
Mg	0.055	0.62	0.174	0.084

697

698 **Supplementary Table 2.** La Vanoise soil sampling: Distances between micro-samples (cores)
 699 in Van_PC and in Van_PS starting from an initial point, named 0 m. Micro-samples were taken
 700 along two transects 10 cm apart, in September 2013 and September 2014. Results obtained on
 701 all micro-samples were gathered as no difference could be shown between the two sampling
 702 dates. Micro-samples named “a” corresponded to the top 1 mm of the cores and sample “b”
 703 corresponded to the 1 to 2 mm below sample “a”. “b*”: in two cases, a second core was taken
 704 beside core “b”, at the same distance from the origin.

Van_PC (17 samples)	Distance, meter												
	0 m	0.3 m	0.6 m	0.9 m	1.2 m	1.5 m	1.8 m	2.1 m	2.4 m	2.7 m	3 m	3.3 m	
September 2013		a, b	b, b*					b	b	b		b	
September 2014		a, b, b*	b	b	b	b	b				b		
Van_PS (22 samples)	0 m	0.32 m	1 m	2 m	2.5 m	3 m	3.5 m	4 m	4.5 m	5 m	7 m	9 m	10 m
September 2013	b		b				b	b	b	b	b	b	
September 2014	a, b	a, b	b	b	b	b	b	b	b	b	b		b

705