*Manuscript with continous line numbering
Click here to view linked References

1 **Environmental filtering: a case of bacterial community assembly in soil**

2

3

4 Yan Yan[a,b,c], Peter G.L. Klinkhamer[b], Johannes A. van Veen[a], Eiko E. Kuramae[a#]

5

6 [a]Department of Microbial Ecology, Netherlands Institute of Ecology NIOO-KNAW,
7 Wageningen, The Netherlands
8 [b]Department of Plant Ecology and Phytochemistry, Institute of Biology Leiden, Leiden
9 University, The Netherlands
10 [c]Biostatistics Department, Harvard T. H. Chan School of Public Health, Boston, USA

11

12

13 [#]Corresponding author: Eiko E. Kuramae. Department of Microbial Ecology, Netherlands
14 Institute of Ecology NIOO-KNAW. Droevendaalsesteeg 10, 6708 PB, Wageningen, The
15 Netherlands. Tel: +31-317-473502. Fax: +31-317-473675. E-mail address:
16 E.Kuramae@nioo.knaw.nl

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

Abstract

Soil has a strong effect on the assembly of bacterial communities, as revealed by studies in which sterilized soil is inoculated with dilution series of bacterial suspensions from the same soil. However, the impact of soil on the assembly of bacterial communities after inoculation with suspensions from different soils is not clear. Here, we explored environmental filtering of bacterial community assembly. Diluted suspensions from different soils harboring different bacterial diversities were used to cross-inoculate three pre-sterilized soils. The main differences in the abiotic factors of the soils were organic matter, ammonium, nitrate, and phosphorus content, pH and the C:N ratio. We used 16S rRNA gene amplicon sequencing to determine the bacterial community structure of the suspensions and the soils. When the different diluted suspensions were used to inoculate their native soils, the regrown soil bacterial communities clustered together; by contrast, the communities were separated when the same suspensions were used to inoculate soils other than their native soils. The diversity indices of the suspensions decreased significantly upon dilution. The strength of selection of bacterial communities by soil was stronger for the 10^{-1} diluted soil samples than for the 10^{-9} diluted soil samples. Thus, differences in abiotic factors shape and explain the variation in bacterial community assemblage among these soils.

56 ## 1. Introduction

57 Soil-borne bacteria are an essential component of terrestrial ecosystems and key to
58 many vital ecosystem functions (Philippot et al., 2013). The diversity of soil-borne bacteria
59 comprises hundreds of thousands of taxa per gram of soil and represents the majority of
60 biodiversity in terrestrial ecosystems (Torsvik and Ovreas, 2002). Abiotic factors such as
61 soil properties (Kuramae et al., 2012b; Kuramae et al., 2014), moisture (Brockett et al.,
62 2012; Lupatini et al., 2019), salinity (Crump et al., 2004), nutrient (Pan et al., 2014; Leite et
63 al., 2017) and organic matter content (Verbruggen et al., 2010; Suleiman et al., 2016;
64 Lourenço et al., 2018; Suleiman et al., 2018), and biotic interactions with local communities
65 of macro- and micro-organisms, drive the activity of soil-borne bacteria and shape their
66 community structure (Cassman et al., 2016; Lupatini et al., 2017)

67 Assessing and predicting the dynamics of microbial communities in soil and other
68 natural environments under variable conditions requires a proper understanding of the
69 mechanisms of microbial community assembly is a long-standing goal of microbial ecology.
70 Microorganisms are dispersed globally and can propagate in any habitat with suitable
71 environmental conditions (Martiny et al., 2006). Upon their arrival in a new environment, or
72 upon drastic changes in their current environment, microorganisms may either assemble into
73 distinct new community profiles (Panke-Buisse *et al.*, 2015) or functionally adapt to the
74 local habitat without dramatic changes in community composition (Comte and del Giorgio,
75 2010; Navarrete et al., 2015).

76 Many studies have addressed the factors responsible for the structuring of microbial
77 communities in soils (Pavon-Jordan et al., 2013), but such studies typically focused on the
78 importance of a single factor without considering the full complexity of all edaphic
79 properties, which accumulating evidence indicates are critical drivers shaping microbial
80 communities (Kuramae et al., 2012a; Navarrete et al., 2013). The restricted depth of
81 analyses in the few studies that have examined the development of communities after the
82 addition of diverse inocula in soils has limited the level of detail of the resulting taxonomic
83 information (Garland and Lehman, 1999; Franklin et al., 2001; Franklin and Mills, 2006;
84 Wertz et al., 2006). Soil microbiome transplantation has been recognized as a promising
85 method for soil restoration and ecosystem repair. However, it is difficult to predict with
86 confidence how soil microbial communities are assembled in different soil habitats.

87

88 The major aim of this study was to determine the role of the soil environment in
89 filtering the bacterial community after inoculating sterilized soils with different suspensions
90 of varying bacterial diversity. Here, the term 'environmental filtering' refers to the
91 prevention of the establishment of a bacterial community by the abiotic environment in the
92 absence of existing biotic interactions. A previous study demonstrated that the dilution
93 approach, in which suspensions of different dilutions were inoculated into sterilized soils in
94 order to establish initial microbial communities of different diversity, is suitable for
95 manipulating the diversity of bacterial communities, and that soil has strong selective power
96 in shaping the microbial community after inoculation, leading to a rather uniform structure
97 of the regrown microbial community (Yan et al., 2015). As we observed in our previous
98 study, inoculation with 10^{-1} diluted suspensions resulted in final bacterial communities with
99 higher diversity than inoculation with 10^{-9} diluted suspensions (Yan et al., 2015). In this
100 study, we took this approach one step further by inoculating three soils with two dilutions of
101 the suspensions of these three different soils in a cross-inoculation design. We hypothesized
102 that soil exerts a dominant selection of bacterial communities after incubation with
103 suspensions of different bacterial diversity. In particular, we were interested in which soil
104 characteristics determined the structure of the final bacterial community.

105

106 **2. Material and Methods**

107 *2.1. Soil sampling*

108 Three field loamy sand soil soils (Table 1) were selected across The Netherlands: soils
109 from the surroundings of Utrecht (52°03′N, 5°13′E), from the so-called Clue fields
110 (52°03′N, 5°45′E) and from the Meijendel dunes (52°9′N, 4°22′E). Ten liters of each soil
111 were collected by scoops at a depth of around 15 cm in each field. The soil was sieved
112 (5mm) and homogenized, and aliquots of 50 g were stored in plastic bags (Whirl-Pak
113 sampling bag, 100 ml; Sigma-Aldrich). One bag of each soil was kept separately to serve as
114 inoculum. The bags with soil were sterilized by gamma irradiation (> 35 kGy; Isotron, Ede,
115 The Netherlands). As compared to autoclaving and freezing, this way of sterilizing soils
116 minimizes the effects on abiotic soil properties. Sterility was checked by spreading 0.5 g of
117 sterilized soil onto Reasoner's 2A (R2A) and potato-dextrose agar (PDA) media and

118 incubated for one week. Neither bacterial nor fungal growth was observed on agar plates
119 after incubation of the sterilized soil samples (n=6). To control the sterility during the
120 experiment, plates were further incubated for the duration of the experiment at 28 °C, and
121 no colonies were observed on R2A medium during the entire incubation period.
122
123 *2.2. Experimental design*
124 The experiment was designed as a cross inoculation experiment, in which we
125 inoculated suspensions of the three soils in each of the sterilized soils (Fig. S1). The
126 experiment consisted of twelve treatments in a factorial design, with 108 samples, *i.e*. three
127 soils × three inocula × two dilutions × six replicates, and 9 sterile controls (three soils ×
128 three replicates). Three sterilized bags of each soil were inoculated with autoclaved
129 demineralized water to be used as controls. A subsample of the fresh soil was taken to
130 determine soil moisture (24 h, 105 °C). The six replicates of each soil type were
131 homogenized, sieved through 4 mm mesh size sieve, and 400 g soil was analyzed at
132 Chemical Laboratory of the Netherlands Institute of Ecology (NIOO-KNAW) to determine
133 soil pH, total carbon and nitrogen concentrations, nitrate and ammonium, total organic
134 matter and available phosphate.
135 Soil suspensions were made by mixing 20 g fresh soil and 190 ml autoclaved
136 demineralized water with a blender for 2 minutes. This procedure was repeated 3 times and
137 in between the blender was cooled down on ice for 2 minutes. The obtained suspension was
138 called the 10^{-1} dilution. 100 ml of 10^{-1} dilution was transferred to a bottle containing 900 ml
139 of autoclaved demineralized water and subsequently shaken by hand for 1 min. This
140 procedure was repeated until 10^{-9} dilutions were obtained. Subsequently, 2.5 ml of the
141 respective dilutions were mixed gently with 25 g of soil in the bags and additional
142 demineralized water was given to bring the moisture level of the inoculated soil at around
143 20%, which is roughly similar to the average level of the prevailing moisture conditions at
144 the sites from where the soil was taken.
145 The remaining suspensions were centrifuged at 3,000 g for 10 min at 4 °C and the
146 pellets were stored at -20 °C for further analysis. After inoculation, soil bags were incubated
147 at 20 °C using sterilized cotton plug caps to ensure gas exchange with 70% humidity in the

148 climate chamber. The soils were turned over regularly once a week to enable homogeneous
149 microbial growth.
150
151 *2.3. Bacterial abundance after incubation*
152 After 9 weeks of incubation, soil samples were taken under laminar flow conditions,
153 to determine the bacterial abundance in all treatments by quantitative real time PCR (qPCR).
154 For this analysis total DNA was extracted from the incubated soil using the MoBio Power
155 Soil Extraction Kit according to the supplier's manual (Mo Bio Laboratories, USA). Total
156 DNA concentration was quantified and qualified on ND-1000 spectrophotometer (Nanodrop
157 Technology, Wilmington, DE). Amplification of the 16S rRNA gene was performed using
158 the primer set Eub 338 and Eub 518 (Muyzer et al 1993). Each 25 µl reaction solution
159 consisted of 12.5 µl Sybergreen mix (Bioline, GC-Biotech) with 4 mg/ml BSA in a total
160 volume of 25 µl, 5 µM of each primer, 5 µl template DNA (5 ng/µl). For bacteria, the
161 standard curves were generated using 10-fold dilution series from 10^{-8} to 10^{-3} of plasmid
162 DNA contained 16S rRNA gene sequence of one genus of *Firmicutes*. Polymerase Chain
163 Reactions were run on a Rotor-Gene 3000 (Qiagen) and started with 15 min at 95°C,
164 followed by 40 amplification cycles each of 95 °C for 60 sec, 53 °C 50 sec and 72 °C 60
165 sec. A subsample of the soil from each bag was stored at -20 °C for further analysis.
166 Triplicate reactions per DNA sample and the appropriate set of standards were used. For
167 qPCR assays, a linear relationship was presumed between the log of the plasmid DNA copy
168 number and the calculated threshold cycle (Ct value). PCR efficiencies were 99% and
169 correlation coefficients for standard curves were $R^2 = 0.99$. Because there were differences
170 of bacterial abundance between the 10^{-1} diluted samples and the 10^{-9} diluted samples after 9
171 weeks of incubation (Fig. S2, T-test, $p < 0.05$) in all soil samples, the relative abundance of
172 species was used for analysis and comparison among samples.
173
174 *2.4 16S rRNA amplicon sequencing*
175 Total DNA was extracted from the soil suspensions and incubated soil, as described
176 above, to determine the composition of the respective microbial communities by *16S rRNA*
177 Illumina Miseq (Argonne National Lab, USA). The PCR was carried out using barcoded
178 primers 515F and 806R (Caporaso et al., 2012). The PCR program used included incubation

179 at 95 °C for 5 min followed by 25 cycles each of 95 °C for 30 s, 52 °C 1 min and 72 °C 10
180 min. For PCR reactions 5 mM dNTPs (Invitrogen, Carlsbad, CA), 1 unit of *Taq* polymerase
181 (Roche, Indianapolis, IN), and 5 ng/µl of sample DNA as the template in a total volume of
182 25 µl were used. To control for contamination during PCR amplification, one negative
183 control (water instead of DNA) was included for all PCR reactions. PCR products of each
184 subsample from the barcoded primers were generated in six replicates per sample and
185 purified using the Agencourt AMPure XP beads kit (Beckman Coulter, Inc.). Equimolar
186 purified PCR products that were quantified by fragment analyzer (Advanced Analytical
187 Technologies, GmbH, Germany) were mixed and sequenced using Miseq sequencing from
188 Illumina. All MiSeq data were uploaded to the ENA and are publicly accessible under study
189 accession number PRJEB31454.
190
191 *2.5 Sequence analysis*
192 The RDP extention to PANDASeq (Masella et al., 2012) named Assembler (Cole et
193 al., 2014) was used to merge paired-end reads with a minimum overlap of 10 bp and at least
194 a PHRED score of 25. Primer sequences were removed from the per sample FASTQ files
195 using Flexbar version 2.5 (Dodt et al., 2012). Sequences were converted to FASTA format
196 and concatenated into a single file. All reads were clustered into OTUs using the UPARSE
197 strategy by dereplication, sorted by abundance with at least two sequences and clustered
198 using the UCLUST smallmem algorithm (Edgar et al., 2011) with a phylotype defined at the
199 97% sequence similarity level. These steps were performed with VSEARCH version 1.0.10
200 (Rognes et al., 2015), which is an open-source and 64-bit multithreaded compatible
201 alternative to USEARCH. Subsequently, chimeric sequences were detected using the
202 UCHIME algorithm (Edgar et al., 2011) implemented in VSEARCH. All reads before the
203 dereplication step were mapped to OTUs using the usearch_global method implemented in
204 VSEARCH to create an OTU table and converted to BIOM format 1.3.1 (McDonald et al.,
205 2012). Finally, taxonomic information for each OTU was added to the BIOM file by using
206 the RDP Classifier version 2.10 (Cole et al., 2014). All steps were implemented in a
207 Snakemake workflow (Koster and Rahmann, 2012).
208 The OTU table was rarefied by the lowest number of OTUs in a sample using the
209 QIIME script. Alpha diversity was calculated based on the rarefied OTU table. Shannon

210 diversity indices were determined with the "vegan" package (Dixon, 2003) in R (The R
211 Foundation for Statistical Computing). The percentage of coverage was calculated by
212 Good's method using the formula % = [1-(n/N)] $\times$ 100, where n means the number of
213 phylotypes represented by singletons and N is the total number of sequences (Good, 1953).
214 Good's method equation gives an estimate of the coverage of an entire sampled community.
215 We determined how environmental variables such as the C:N, soil pH, nitrate and
216 ammonium, total organic matter and available phosphate correlated with bacterial OTU
217 abundance by performing Canonical Correspondence Analyses (CCA) on both dilution of
218 10^{-1} and 10^{-9} communities.

219 To check the treatment effects on the composition of bacterial communities, between-
220 class analysis (BCA) was carried out using the package "ade4" (Dray & Dufour, 2007) in R
221 . To explore the dissimilarities of treatments in each community, a Principal Component
222 Analysis (PCA) was used to create BCA tables using the function "bca". Between-class
223 analysis (BCA) measured the amount of variance restricted to the grouping factor as a
224 percentage of the total inertia (Dray and Jombart, 2011, Thioulouse *et al.*, 2012). BCA is an
225 alternative method to linear discriminant analysis for which the number of samples is
226 smaller than the number of variables (Thioulouse *et al.*, 2012). Principal Component
227 Analysis (PCA) was applied to each community data set prior to BCA. Monte-Carlo tests of
228 the treatment groups were performed with 999 permutations. The results were visualized by
229 ordination plot. To identify the bacterial taxa significantly responsible for the dissimilarities
230 among treatments, multiple factor analysis (MFA) with ascending hierarchical classification
231 (AHC) was performed using the FactoMineR package and visualized using factoextra
232 package. P-values were adjusted by false discovery rate ($p < 0.05$). The effects of soil and
233 dilution on bacterial community composition were tested by a two-way PERMANOVA test
234 for each inoculum, respectively. All the multivariate analyses were performed using the
235 PAST software (Hammer et al, 2001).

236

237 **3. Results**

238 *3.1. Chemical characteristics of the three field soils*

239 The three field soils, Meijendel, Utrecht, and Clue, were chemically different.
240 Meijendel soil had the highest pH and contents of NO_3^-, NH_4^+ and organic matter compared

241 with the other soils, while Utrecht soil had the highest C:N ratio and lowest pH. Clue field
242 soil had the highest phosphorus content (Table 2). Between-class analysis (BCA) explained
243 89% (p=0.001) of the differences among the soils (Fig. 1a). To extract major environmental
244 factors among the combination of explanatory variables in each soil communities, we
245 performed canonical correspondence analyses (Fig. S3). Bacterial communities that were
246 grown in the Meijendel soil were driven mainly by organic matter and high pH, as well as
247 contents of NO_3^- and NH_4^+, in Utrecht soil the structure of bacterial communities was
248 mainly affected by C:N ratio and phosphorus was the major factor controlling the bacterial
249 community structure in Clue field soil. This is true for communities assembled from both
250 10^{-1} diluted and 10^{-9} diluted suspensions.
251
252 *3.2. Effect of dilution and soil on bacterial community diversity*
253 The alpha diversity indices of each of the three soil suspensions decreased
254 significantly upon dilution (Supplementary Table S1; t-test, $p<0.05$), suggesting strong
255 impact of dilution. These changes remained substantial after incubation in the different soils
256 especially for the communities formed after incubation of the high diversity, i.e. 10^{-1} diluted
257 inocula from the Clue and Meijendel soils. Good's coverage estimator was above 99% for
258 all samples, indicating that the sequencing depth was sufficient to detect most species. In
259 general, the diversity of the soil communities formed after inoculation with 10^{-9} diluted
260 suspensions increased during incubation compared with the original 10^{-9} suspension (t-test,
261 $p<0.05$).
262
263 *3.3. Effects of dilution and soil on bacterial community composition*
264 BCA of the bacterial community compositions grouped by soil types after incubation
265 of in both 10^{-1} diluted suspensions (Fig. 1b) and 10^{-9} diluted suspensions (Fig. 1c)
266 independent of the origin of the suspension, indicating the major role of the soil and dilution
267 treatment to the assemblage of bacterial communities (Monte-Carlo test of groups, all p =
268 0.001). The relative positions of the communities in the soil incubated with 10^{-1} inocula and
269 10^{-9} inocula were equal to the relative positions of the soil factors.
270 We quantified the individual effects of soil and dilution on the regrown bacterial
271 communities after incubation with the different inocula. Based on the associated *F*-values,

272 for all three soils, the dilution effect was slightly stronger than the soil effect, and both
273 individual effects were larger than their interactions. PERMANOVA of these two factors
274 yielded statistically significant results regarding dilutions and soils, as well as their
275 interactions (Table 3).

276 We next compared the differences between the within-soil and between-soil diversity
277 (Fig. 2) by calculating the Bray-Curtis dissimilarity distance between individual samples.
278 Soil bacterial compositions were similar when a soil suspension was used to inoculate the
279 same soil compared with cross-inoculation in another soil. This pattern was especially
280 obvious for the soil with 10^{-1} inocula. For 10^{-9} inocula, there was large variations between
281 replicates when a suspension was used to inoculate the same soil.

282 The most dominant phyla in the 10^{-1} diluted suspension samples were *Proteobacteria*,
283 *Acidobacteria*, *Actinobacteria*, *Planctomycetes*, *Bacteroidetes* and *Verrucomicrobia*
284 (Supplementary Fig. S4). Multiple factor analysis (MFA) at the family taxonomic level
285 revealed the effect of soil on the three different inocula types (Fig. 3 and Supplementary
286 Tables S2). In the 10^{-1} Utrecht soil suspension, dim 1 and dim 2 explained 39.1% and 30%
287 of the separation among the three soil clusters, respectively (Fig 3a), with contributions of
288 *Paenibacillaceae*, *Acidobacteria subdivisions* Gp1, *Rhodospirillales*, *Acetobacteraceae* and
289 *Micrococcaceae* along dim 1 and *Solirubrobacterales*, *Clostridiaceae*,
290 *Gemmatimonadaceae*, *Burkholderiales* and *Bdellovibrionaceae* along dim 2 (Table S2). In
291 the 10^{-9} Utrecht soil suspension, dim 1 and dim 2 explained 29.4% and 25.5% of the
292 differences between the clusters, respectively (Fig. 3d), and the dissimilarity among the
293 three soils was dominated by *Acidobacteria subdivisions* Gp2, *Ktedonobacteraceae*, and
294 *Xanthobacteraceae* along dim 1 and *Clostridia* clades, *Gracilibacteraceae* and
295 *Sphingobacteriaceae* along dim 2 (Table S2).

296 In the 10^{-1} Clue suspension, MFA revealed that the clusters associated with dim 1 and
297 dim 2 explained 42.4% and 31% of the total variance among the three soils, respectively
298 (Fig. 3b). The dissimilarity among the three soils was caused mostly by changes in
299 *Acidobacteria subdivision* Gp1, *Hyphomicrobiaceae*, *Comamonadaceae*, *Micrococcaceae*
300 and *Bdellovibrionaceae* along dim 1 and *Myxococcales*, *Burkholderiales*, *Acidobacteria*
301 subdivision Gp4, *Betaproteobacteria*, *Gemmatimonadaceae* along dim 2 (Table S2). In the
302 10^{-9} Clue suspension, the clusters associated with dim 1 and dim 2 explained 31.7% and

303 27.9% of the differences among the three soils, respectively (Fig. 3e), with contributions of
304 *Gaiellaceae*, *Nitrospiraceae*, *Gemmatimonadaceae*, *Cystobacteraceae* and *Clostridiaceae*
305 along dim 1 and *Deinococcaceae*, *Ktedonobacteria* clades, *Acidobacteria* subdivision Gp2,
306 and *Solirubrobacteraceae* along dim 2 (Table S2).

307 In the 10^{-1} Meijendel soil suspension, dim 1 and dim 2 explained 40.8% and 28.6% of
308 the separation of the clusters among the three soils, respectively (Fig. 3c). These differences
309 were mainly due to *Hyphomicrobiaceae*, *Bacillales* clades, *Micrococcaceae* and
310 *Phyllobacteriaceae* along dim 1 and *Pseudonocardiaceae*, *Sphingobacteriaceae*,
311 *Intrasporangiaceae*, unclassified *Clostridiales* and *Kinrosporiaceae* along dim 2 (Table S2).
312 In the 10^{-9} Meijendel soil suspension, dim 1 and dim 2 explained 33.6% and 27.9% of the
313 separation among the three soils (Fig. 3f) due to contributions of *Acidobacteria* subdivision
314 Gp2, *Rhodospirillales*, *Ktedonobacterales*, and Actinobacteria clades, including
315 Actinomycetales and *Solirubrobacterales*, along dim 1 and *Micrococcaceae*,
316 *Thermoactinomycetaceae*, *Cystobacteraceae*, *Clostridiaceae* and *Paenibacillaceae* along
317 dim 2 (Table S2).
318
319

320 **4. Discussion**

321 In a previous study (Yan et al., 2015), we found strong indications that the overall soil
322 environment drives the structure of the bacterial communities that develop after inoculation
323 with suspensions of different bacterial diversities and compositions. Thus, soil acts as an
324 environmental filter that determines whether arriving species fail or successfully establish in
325 the new abiotic environment. The concept of environmental filtering has mainly been
326 applied in plant ecology (Kraft et al., 2015) and the current study was designed to test the
327 concept of environmental filtering by soil in the assembly of bacterial communities. The
328 results showed that soil abiotic factors led to homogeneous bacterial community structures
329 even after inoculation with suspensions from different soils and thus with different bacterial
330 communities. The compositions of the regrown soil bacterial communities were strongly
331 determined by the soil in which they were inoculated; that is, the communities showed great
332 similarity within each of the three inoculated soils but large differences across the three
333 inoculated soils. This result confirmed our earlier suggestion of the strong impact of soil on

334 the of bacterial community assemblage. The soil effect was clear at both the phylum
335 (Supplementary Fig. S4) and OTU (Figs. 2 and 3) taxonomic levels.

336 In addition, we observed that dilution had a strong effect on bacterial diversity in the
337 different soils (Table 3). Similar to our previous observations (Yan et al., 2015), diluting soil
338 suspensions 10^{-9} led to a significant reduction of the diversity of the bacterial communities.
339 The soil was sterilized before inoculation and thus did not add a substantial inoculum to the
340 community. The observed increase in the diversity of some of the communities that
341 developed after inoculation of the diluted 10^{-9} suspensions compared with the original
342 suspension may reflect the failure of the technological approach to detect all organisms in a
343 sample, although Good's coverage was large enough to assume with confidence that the
344 largest proportion of the present community was included in sequencing. Remarkably, the
345 diversity of the diluted 10^{-9} suspension of Clue soil was much lower than that of the other
346 10^{-9} diluted suspensions (Table S1) and was associated with a dominant proportion of
347 *Proteobacteria* (Supplementary Fig. S4). The dominance of *Proteobacteria* might be
348 explained by the lower organic matter content in this soil compared with the other two soils,
349 which favors fast-growing bacteria such as *Proteobacteria* (Suleiman et al., 2016; Lupatini
350 et al., 2017). This dominance of *Proteobacteria* may also explain why the diversity indices
351 were lowest when 10^{-9} diluted Clue soil suspensions when regrown in Meijendel soil, which
352 had the richest nutrient content, as *Proteobacteria* were selected in the dilution procedure.

353 Generally, the bacterial community composition similarity among the soils was higher
354 after inoculation of the more diverse in the 10^{-1} diluted samples compared with the less
355 diverse in the 10^{-9} diluted samples (Fig. 2). This similarity was particularly evident when the
356 suspensions were inoculated into their native soils. Furthermore, the BCA showed different
357 responses of the bacterial compositions of the 10^{-1} and 10^{-9} inocula to soil type.
358 *Proteobacteria* were dominant in each soil after incubation with 10^{-1} diluted samples, which
359 may explain the relative similarity of the communities after incubation with the 10^{-1} diluted
360 soil suspensions, in contrast to the communities that developed after incubation with the 10^{-9}
361 diluted soil suspensions, in which neither *Proteobacteria* nor any other phylum was
362 consistently dominant (Supplementary Fig. S4).

363 It is not surprising that dilution resulted in a dramatic reduction of species diversity.
364 However, independent of the original composition and diversity of the microbial pools from

365 the different soils, soil reshaped the community assemblage process, indicating that dilution
366 was not solely responsible for the differences in the resulting communities. The diversity
367 indices of the incubated samples were not always highest in their own native soil
368 (Supplementary Table S1), which indicates that only a selected fraction of the original
369 community is able to establish in soil, even in native soil. The diluted inoculum is a mixture
370 of dormant, dead, and active microbes, and after addition of this diluted pool of microbes to
371 a new abiotic environment, in this case sterilized soils, the new microbial community will be
372 assembled from those present in the dilution. Moreover, the composition of the new
373 community is very dependent on the resources available in the new environment, especially
374 available nutrients. Thus, the environmental conditions filter the new population based not
375 on its taxonomic composition but rather on the traits the members of the population carry in
376 their genomes. For instance, Meijendel soil, which is rich in organic matter, nitrate and
377 ammonium at a desirable pH (neutral) for the availability of nutrients to plants, provides
378 favorable conditions for microbes that can decompose organic matter, particularly those
379 whose genomes are rich in genes encoding enzymes (glycosyl hydrolases). This nutrient
380 availability probably explains the enrichment of *Acidobacteria* (Kielak et al., 2016a; Kielak
381 et al., 2016b) and *Actinobacteria* in Meijendel soil independent of the source and dilution of
382 the inoculum (Table S2). Meijendel soil also contains high ammonium and nitrate content,
383 which explains the establishment of nitrifiers such as *Nitrospiraceae,* denitrifiers
384 (*Comamonadaceae*, *Hyphomicrobiaceae* (Osaka et al., 2006)) and nitrogen fixers
385 (Acetobacteraceae; Rhizobiales *Hyphomicrobiaceae*). Thus, in this particular soil, bacterial
386 guilds related to C and N cycles were selected due to the chemical conditions that this soil
387 offered. Other groups that were common in the three soils, like *Ktedonobacterales*, are
388 related to the decomposition of highly recalcitrant organic matter (Hug et al., 2013), as was
389 more evident in Utrecht soil, which had the highest C:N ratio.

390 In addition to having the highest C:N content, Utrecht soil had the lowest pH. These
391 soil conditions select bacteria with the capacity to decompose recalcitrant organic matter,
392 especially bacterial groups that possess different glycosyl hydrolases, such as *Acidobacteria*
393 (Kielak et al., 2016b) and *Burkholderiales*, and bacteria associated with environments with
394 high C:N ratios and organic matter content (*Sphingobacteriaceae)* (He et al., 2010).

395 A previous study reported that the abundance of *Firmicutes* is highly correlated with
396 phosphorus content (Kuramae et al., 2012a), and indeed, we found that the *Clostridiales*
397 clades were among the most important in explaining the variations among the three soils.
398 Thus, our study shows that specific soils present resources that serve as determinants of
399 bacterial community structure.

400 In conclusion, we have shown that soil characteristics have a strong impact on
401 bacterial community assemblage. The overall soil environment rather than individual abiotic
402 factors functions as the driver shaping bacterial community structure independent of the
403 diversity of the original suspension inoculated in the soil. The three soils examined in this
404 study modified the bacterial community structure differently by providing specific habitats
405 suitable for the growth of the inocula, confirming the overriding impact of the
406 physicochemical nature of the soil on the assemblage of bacterial communities in terrestrial
407 ecosystems. Thus, approaches such as "microbiome transplantation" and "donor-soil
408 microbes" for soil restoration or ecosystem repair are not simple and straightforward and
409 will be unsuccessful if the history of the soils, namely chemistry and physical
410 characteristics, are not considered when establishing the desired microbial community.

411

412 **Acknowledgments**

413 The authors acknowledge Agata Pijl for laboratory assistance and Mattias de Hollander for
414 bioinformatics assistance. Publication number XXXX of the NIOO-KNAW, Netherlands
415 Institute of Ecology.

416

417 **References**

418 Brockett, B.F.T., Prescott, C.E., Grayston, S.J. 2012. Soil moisture is the major factor
419 influencing microbial community structure and enzyme activities across seven
420 biogeoclimatic zones in western Canada. Soil Biology and Biochemistry 44, 9-20.
421 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N.,
422 Owens, S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G.,
423 Knight, R. 2012. Ultra-high-throughput microbial community analysis on the Illumina
424 HiSeq and MiSeq platforms. The ISME Journal 6, 1621-1624.
425 Cassman, N.A., Leite, M.F.A., Pan, Y., de Hollander, M., van Veen, J.A., Kuramae, E.E.
426 2016. Plant and soil fungal but not soil bacterial communities are linked in long-term
427 fertilized grassland. Scientific Reports 6, 23680.

428 Cole, J.R., Wang, Q., Fish, J.A., Chai, B.L., McGarrell, D.M., Sun, Y.N., Brown, C.T.,
429 Porras-Alfaro, A., Kuske, C.R., Tiedje, J.M. 2014. Ribosomal Database Project: data and
430 tools for high throughput rRNA analysis. Nucleic Acids Research 42, D633-D642.
431 Comte, J., del Giorgio, P.A. 2010. Linking the patterns of change in composition and
432 function in bacterioplankton successions along environmental gradients. Ecology 91, 1466-
433 1476.
434 Crump, B.C., Hopkinson, C.S., Sogin, M.L., Hobbie, J.E. 2004. Microbial biogeography
435 along an estuarine salinity gradient: Combined influences of bacterial growth and residence
436 time. Applied and Environmental Microbiology 70, 1494-1505.
437 Dixon, P. 2003. VEGAN, a package of R functions for community ecology. Journal of
438 Vegetation Science 14, 927-930.
439 Dray, S., Dufour, A-B. 2007. The ade4 package: implementing the duality diagram for
440 ecologists. Journal of Statatistics Software 22,1-20.
441 Dray, S., Jombart, T. 2011. Revisiting Guerry's data: Introducing spatial constraints in
442 multivariate analysis. The Annals of Applied Statistics 5, 2278-99.
443 Dodt, M., Roehr, J.T., Ahmed, R., Dieterich, C. 2012. FLEXBAR—Flexible barcode and
444 adapter processing for next-generation sequencing platforms. Biology 1, 895-905.
445 Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R. 2011. UCHIME improves
446 sensitivity and speed of chimera detection. Bioinformatics 27, 2194-2200.
447 Franklin, R.B., Garland, J.L., Bolster, C.H., Mills, A.L. 2001. Impact of dilution on
448 microbial community structure and functional potential: Comparison of numerical
449 simulations and batch culture experiments. Applied and Environmental Microbiology 67,
450 702-712.
451 Franklin, R.B., Mills, A.L. 2006. Structural and functional responses of a sewage microbial
452 community to dilution-induced reductions in diversity. Microbial Ecology 52, 280-288.
453 Garland, J.L., Lehman, R.M. 1999. Dilution/extinction of community phenotypic characters
454 to estimate relative structural diversity in mixed communities. FEMS Microbiology Ecology
455 30, 333-343.
456 Good, I.J. 1953. The population frequencies of species and the estimation of population
457 parameters. Biometrika 40, 237-264.
458 Hammer, O., Harper, D.A.T., Ryan, P.D. 2001. PAST: paleontological statistics software
459 package for education and data analysis. Palaeontologia Electronica 4, 9 p.
460 He, X., Xiao, T., Kuang, H., Lan, X., Tudahong, M., Osman, G., Fang, C., Rahman, E.
461 2010. Sphingobacterium shayense sp. nov., isolated from forest soil. International Journal of
462 Systematic and Evolutionary Microbiology 60, 2377-2381.
463 Hug, L.A., Castelle, C.J., Wrighton, K.C., Thomas, B.C., Sharon, I., Frischkorn, K.R.,
464 Williams, K.H., Tringe, S.G., Banfield, J.F. 2013. Community genomic analyses constrain
465 the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in
466 sediment carbon cycling. Microbiome 1, 22.
467 Kielak, A., Scheublin, T., Mendes, L., Van Veen, J.A., Kuramae, E.E. 2016a. Bacterial
468 community succession in pine-wood decomposition. Frontiers in Microbiology 7, 231.
469 Kielak, A.M., Barreto, C.C., Kowalchuk, G.A., van Veen, J.A., Kuramae, E.E. 2016b. The
470 ecology of Acidobacteria: moving beyond genes and genomes. Frontiers in Microbiology 7,
471 744.
472 Koster, J., Rahmann, S. 2012. Snakemake-a scalable bioinformatics workflow engine.
473 Bioinformatics 28, 2520-2522.

474 Kraft, N.J.B., Adler, P.B., Godoy, O., James, E.C., Fuller, S., Levine, J.M. 2015.
475 Community assembly, coexistence and the environmental filtering metaphor. Functional
476 Ecology 29, 592-599.
477 Kuramae, E.E., Yergeau, E., Wong, L.C., Pijl, A.S., van Veen, J.A., Kowalchuk, G.A. 2012.
478 Soil characteristics more strongly influence soil bacterial communities than land-use type.
479 FEMS Microbiology Ecology 79, 12-24.
480 Kuramae, E.E., Zhou, J.Z., Kowalchuk, G.A., van Veen, J.A. 2014. Soil-borne microbial
481 functional structure across different land uses. The Scientific World Journal 216071, 8 p.
482 Leite, M.F., Pan, Y., Bloem, J., Berge, H.T., Kuramae, E.E. 2017. Organic nitrogen
483 rearranges both structure and activity of the soil-borne microbial seedbank. Scientific
484 Reports 7, 42634.
485 Lourenço, K.S., Suleiman, A.K.A., Pijl, A., van Veen, J.A., Cantarella, H., Kuramae, E.E.
486 2018. Resilience of the resident soil microbiome to organic and inorganic amendment
487 disturbances and to temporary bacterial invasion. Microbiome 6, 142.
488 Lupatini, M., Korthals, G.W., de Hollander, M., Janssens, T.K.S., Kuramae, E.E. 2017. Soil
489 microbiome is more heterogeneous in organic than in conventional farming system.
490 Frontiers in Microbiology 7, 2064.
491 Lupatini, M., Suleiman, A.K.A., Jacques, R.J.S., Lemos, L.N., Pylro, V.S., Van Veen, J.A.,
492 Kuramae, E.E., Roesch, L.F.W. 2019. Moisture is more important than temperature for
493 assembly of both potentially active and whole prokaryotic communities in subtropical
494 grassland. Microbial Ecology 77, 460–470.
495 Martiny, J.B., Bohannan, B.J., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L.,
496 Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., Morin, P.J., Naeem, S.,
497 Ovreas, L., Reysenbach, A.L., Smith, V.H., Staley, J.T. 2006. Microbial biogeography:
498 putting microorganisms on the map. Nature Reviews in Microbiology 4, 102-112.
499 Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G., Neufeld, J.D. 2012.
500 PANDAseq: PAired-eND Assembler for Illumina sequences. BMC Bioinformatics 13.
501 McDonald, D., Clemente, J.C., Kuczynski, J., Rideout, J.R., Stombaugh, J., Wendel, D.,
502 Wilke, A., Huse, S., Hufnagle, J., Meyer, F., Knight, R., Caporaso, J.G. 2012. The
503 biological observation matrix (BIOM) format or: how I learned to stop worrying and love
504 the ome-ome. Gigascience 1, 7.
505 Navarrete, A.A., Kuramae, E.E., de Hollander, M., Pijl, A.S., van Veen, J.A., Tsai, S.M.
506 2013. Acidobacterial community responses to agricultural management of soybean in
507 Amazon forest soils. FEMS Microbiology Ecology 83, 607-621.
508 Navarrete, A.A., Tsai, S.M., Mendes, L.W., Faust, K., de Hollander, M., Cassman, N.A.,
509 Raes, J., van Veen, J.A., Kuramae, E.E. 2015. Soil microbiome responses to the short-term
510 effects of Amazonian deforestation. Molecular Ecology 24, 2433-2448.
511 Osaka, T., Yoshie, S., Tsuneda, S., Hirata, A., Iwami, N., Inamori, Y. 2006. Identification of
512 acetate- or methanol-assimilating bacteria under nitrate-reducing conditions by stable-
513 isotope probing. Microbial Ecology 52, 253-266.
514 Pan, Y., Cassman, N., de Hollander, M., Mendes, L.W., Korevaar, H., Geerts, R., van Veen,
515 J.A., Kuramae, E.E. 2014. Impact of long-term N, P, K, and NPK fertilization on the
516 composition and potential functions of the bacterial community in grassland soil. FEMS
517 Microbiology Ecology 90, 195-205.

518 Pavon-Jordan, D., Karell, P., Ahola, K., Kolunen, H., Pietiainen, H., Karstinen, T.,
519 Brommer, J.E. 2013. Environmental correlates of annual survival differ between two
520 ecologically similar and congeneric owls. Ibis 155, 823-834.
521 Philippot, L., Raaijmakers, J.M., Lemanceau, P., van der Putten, W.H. 2013. Going back to
522 the roots: the microbial ecology of the rhizosphere. Nature Reviews Microbiology 11, 789-
523 799.
524 Suleiman, A.K.A., Gonzatto, R., Aita, C., Lupatini, M., Jacques, R.J.S., Kuramae, E.E.,
525 Antoniolli, Z.I., Roesch, L.F.W. 2016. Temporal variability of soil microbial communities
526 after application of dicyandiamide-treated swine slurry and mineral fertilizers. Soil Biology
527 and Biochemistry 97, 71-82.
528 Suleiman, A.K.A., Lourenco, K.S., Pitombo, L.M., Mendes, L.W., Roesch, L.F.W., Pijl, A.,
529 Carmo, J.B., Cantarella, H., Kuramae, E.E. 2018. Recycling organic residues in agriculture
530 impacts soil-borne microbial community structure, function and N_2O emissions. Science of
531 the Total Environment 631-632, 1089-1099.
532 Thioulouse. J., Prin, Y., Duponnois, R. 2012. Multivariate analyses in soil microbial
533 ecology: a new paradigm. Environmental and Ecological Statistics 19, 499-520.
534 Torsvik, V., Ovreas, L. 2002. Microbial diversity and function in soil: from genes to
535 ecosystems. Current Opinion in Microbiology 5, 240-245.
536 Verbruggen, E., Roling, W.F., Gamper, H.A., Kowalchuk, G.A., Verhoef, H.A., van der
537 Heijden, M.G. 2010. Positive effects of organic farming on below-ground mutualists: large-
538 scale comparison of mycorrhizal fungal communities in agricultural soils. New Phytologist
539 186, 968-979.
540 Wertz, S., Degrange, V., Prosser, J.I., Poly, F., Commeaux, C., Freitag, T., Guillaumaud, N.,
541 Roux, X.L. 2006. Maintenance of soil functioning following erosion of microbial diversity.
542 Environmental Microbiology 8, 2162-2169.
543 Yan, Y., Kuramae, E.E., Klinkhamer, P.G., van Veen, J.A. 2015. Revisiting the dilution
544 procedure used to manipulate microbial biodiversity in terrestrial systems. Applied and
545 Environmental Microbiology 81, 4246-4252.

546

547 **Figure legends**

548 **Fig. 1**. Between-class analysis (BCA) explaining the variability of (a) the soil factors and
549 the bacterial community compositions at the family level in soil incubated with (b) 10^{-1} and
550 (c) 10^{-9} (Fig. 1c) inocula. csoil = Clue field soil, usoil = Utrecht soil, msoil = Meijendel
551 soil).

552 **Fig. 2**. Boxplot of Bray-Curtis distances for within- and between-soil bacterial diversity in
553 different soils incubated with (b) 10^{-1} and (c) 10^{-9} (Fig. 1c) inocula.

554 **Fig. 3.** Multiple factor analysis (MFA) at the family level of the soil effect in three field
555 soils (U=Utrecht, C=Clue, M=Meijendel) incubated with 10^{-1} and 10^{-9} inocula. (a) and (d)
556 10^{-1} and 10^{-9} suspensions of Utrecht soil (U) inoculated in Clue field soil (Uc), Utrecht soil

557 (Uu) and Meijendel soil (Um), respectively; (b) and (e) 10^{-1} and 10^{-9} suspensions of Clue
558 field soil (C) inoculated in Clue field soil (Cc), Utrecht soil (Cu) and Meijendel soil (Cm),
559 respectively; (c) and (f) 10^{-1} and 10^{-9} suspensions of Meijendel soil (M) inoculated in Clue
560 field soil (Mc), Utrecht soil (Mu) and Meijendel soil (Mm), respectively. Numbers are the
561 replicates (n=5).

562 **Tables**

563

564 **Table 1.** Textural composition of the loamy sand soils used in this study.

565

Physical soil properties	Utrecht soil	Clue field soil	Meijendel soil
Sand (%)	62	70	78
Silt (%)	21	18	10
Clay (%)	14	8	8
Coarse fragment (%)	3	4	4

566

567

568

569 **Table 2.** Soil chemical properties.

570

Chemical Properties	Utrecht soil	Clue field soil	Meijendel soil
OM (%)	4.67±0.18 a	3.97±0.29 a	9.11±0.36 b
NO_3^- (mg/kg)	0.02±0.02 a	6.50±0.51 b	30.43±0.85 c
NH_4^+ (mg/kg)	0.92±0.20 a	1.21±0.18 ab	2.23±0.25 b
Phosphorus (mg/kg)	2.28±0.35 a	80.84±3.56 b	15.16±0.41 c
C:N ratio	20.30±1.22 a	14.81±0.69 ab	12.16±0.26 b
pH (H_2O)	4.61±0.023 a	5.77±0.015 b	7.47±0.005 c

571 Values are means ± SE, n = 6. Within columns, means followed by the same letter are not
572 significantly different ($p < 0.05$) based on Tukey's HSD test. OM means organic matter. The
573 data were transformed to fit a normal distribution when necessary.

574

575

576

577

578

579 **Table 3.** Results of two-way PERMANOVA using Bray-Curtis similarity showing the
580 effects of soil, dilution and their interaction on bacterial community composition.

Inoculum	Factor	Sum of sqrs	df	Mean Square	*F*	p
U suspension	Dilution	1.59	1	1.59	14.79	0.0001
	Soil	2.28	2	1.14	10.61	0.0001
	Interaction	1.28	2	0.64	5.98	0.0001
C suspension	Dilution	2.22	1	2.21	33.03	0.0001
	Soil	2.11	2	1.06	15.76	0.0001
	Interaction	1.49	2	0.74	11.10	0.0001
M suspension	Dilution	2.15	1	2.15	20.89	0.0001
	Soil	2.30	2	1.15	11.15	0.0001
	Interaction	1.27	2	0.63	6.16	0.0001

581
582 U= Utrecht soil; C= Clue field soil; M= Meijendel soil.

583

Figure 1

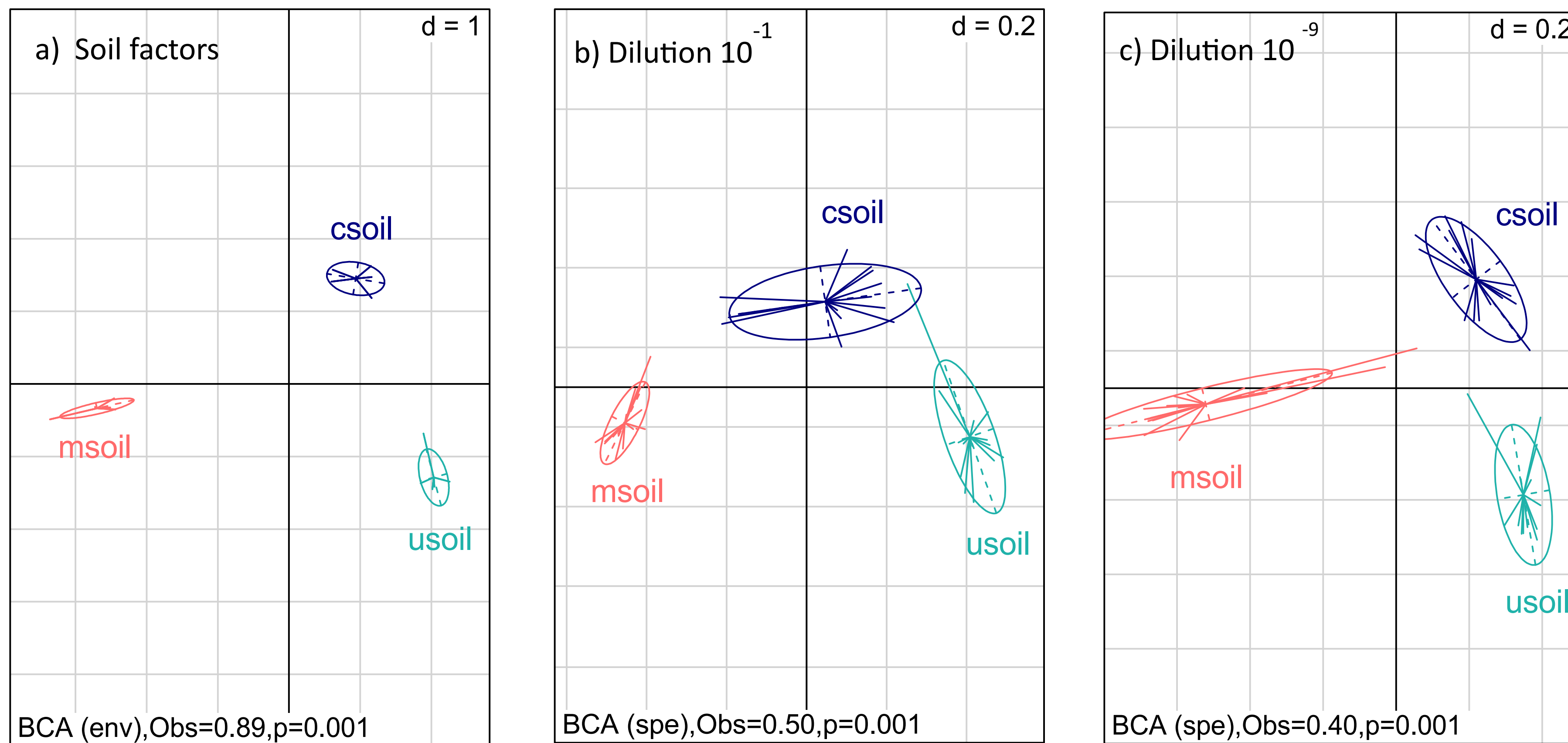


Figure 1

Figure 2

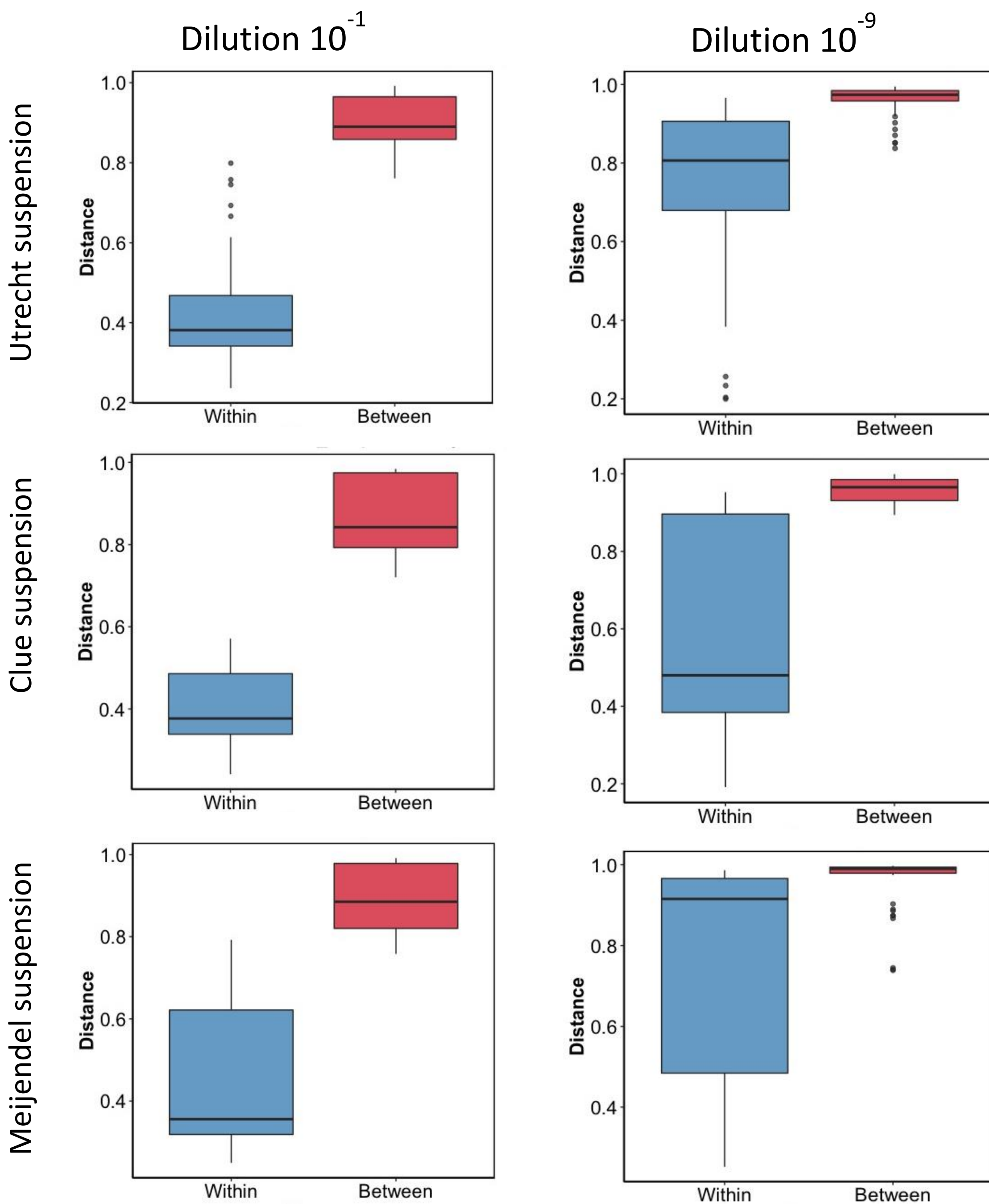
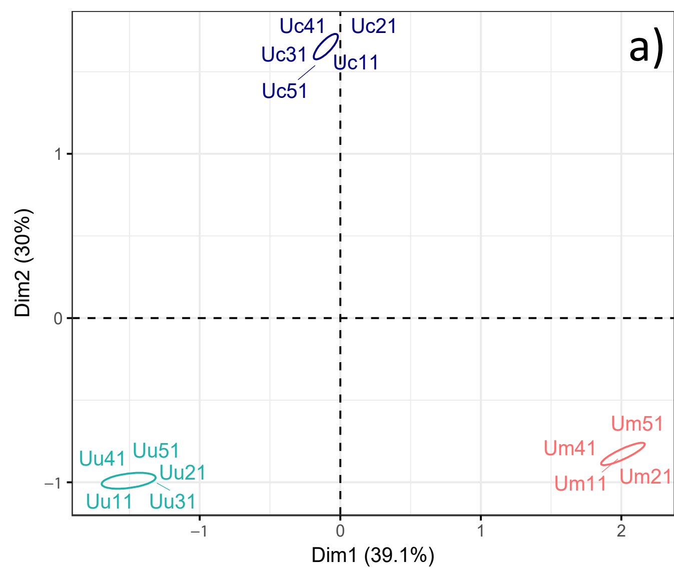
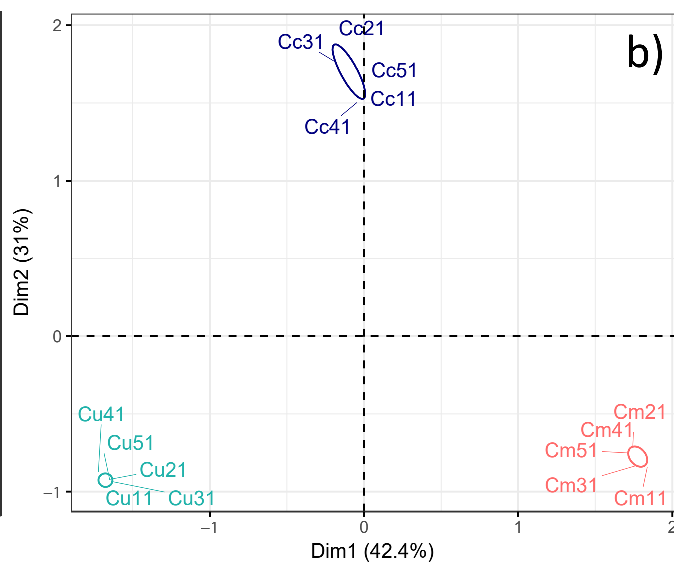
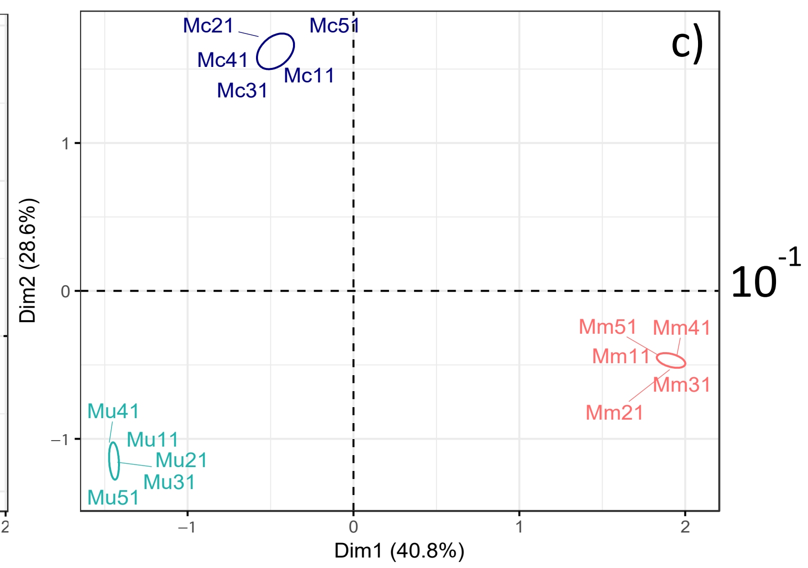


Figure 3

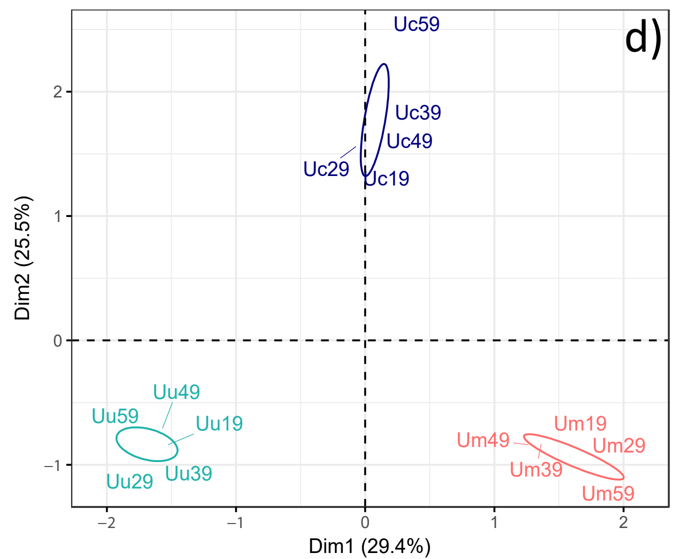
U suspension | C suspension | M suspension

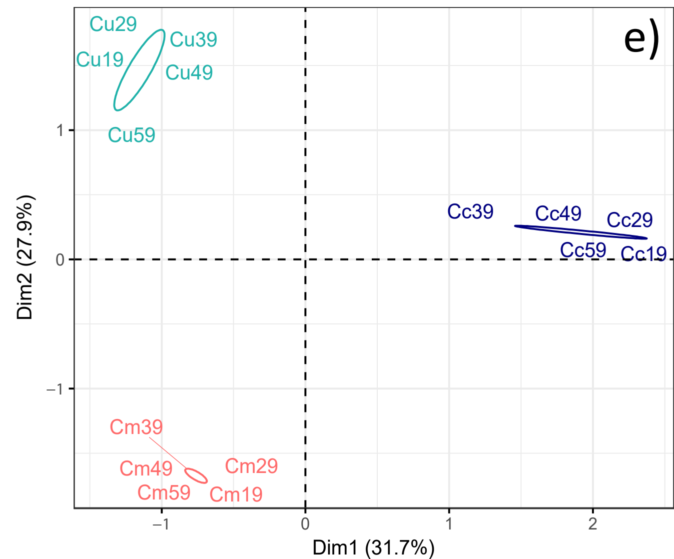


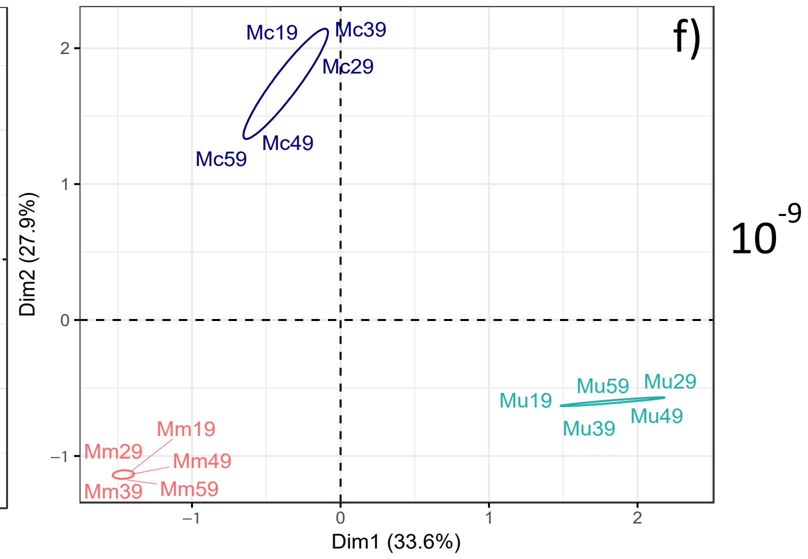




10^{-1}







10^{-9}

Supplementary Material for online publication only

Click here to download Supplementary Material for online publication only: R2-Supplementary_Material.docx