

**This is an electronic reprint of the original article.  
This reprint *may differ* from the original in pagination and typographic detail.**

**Author(s):** Gopala Krishnan, Manoj Kumar; Brader, Günter; Sessitsch, Angela; Mäki, Anita; Elsas, Jan Dirk Van; Nissinen, Riitta

**Title:** Plants Assemble Species Specific Bacterial Communities From Common Core Taxa in Three Arcto-Alpine Climate Zones

**Year:** 2017

**Version:**

**Please cite the original version:**

Gopala Krishnan, M. K., Brader, G., Sessitsch, A., Mäki, A., Elsas, J. D. V., & Nissinen, R. (2017). Plants Assemble Species Specific Bacterial Communities From Common Core Taxa in Three Arcto-Alpine Climate Zones. *Frontiers in Microbiology*, 8, 12.  
<https://doi.org/10.3389/fmicb.2017.00012>

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

# Plants Assemble Species Specific Bacterial Communities From Common Core Taxa in Three Arcto-Alpine Climate Zones

Manoj Kumar<sup>1, 2\*</sup>, Günter Brader<sup>3</sup>, Angela Sessitsch<sup>3</sup>, Anita Mäki<sup>2</sup>, Jan Dirk Van Elsas<sup>1</sup>, Riitta Nissinen<sup>2</sup>

<sup>1</sup>Department of Microbial Ecology, University of Groningen, Netherlands, <sup>2</sup>Department of Biological and Environmental Science, University of Jyväskylä, Finland, <sup>3</sup>Health & Environment Department, AIT Austrian Institute of Technology, Austria

*Submitted to Journal:*  
Frontiers in Microbiology

*Specialty Section:*  
Plant Biotic Interactions

*ISSN:*  
1664-302X

*Article type:*  
Original Research Article

*Received on:*  
21 May 2016

*Accepted on:*  
03 Jan 2017

*Provisional PDF published on:*  
03 Jan 2017

*Frontiers website link:*  
[www.frontiersin.org](http://www.frontiersin.org)

*Citation:*  
Kumar M, Brader G, Sessitsch A, Mäki A, Van\_elsas J and Nissinen R(2016) Plants Assemble Species Specific Bacterial Communities From Common Core Taxa in Three Arcto-Alpine Climate Zones. *Front. Microbiol.* 8:12. doi:10.3389/fmicb.2017.00012

*Copyright statement:*  
© 2017 Kumar, Brader, Sessitsch, Mäki, Van\_elsas and Nissinen. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Provisional

1 **Plants Assemble Species Specific Bacterial Communities**  
2 **From Common Core Taxa in Three Arcto-Alpine Climate**  
3 **Zones**

4  
5 Manoj Kumar <sup>1,2\*</sup>, Günter Brader<sup>3</sup>, Angela Sessitsch<sup>3</sup>, Anita Mäki<sup>2</sup>, Jan Dirk van  
6 Elsas<sup>1</sup>, Riitta Nissinen<sup>2</sup>

7  
8 Affiliations

9 1 Department of Microbial Ecology, University of Groningen, Groningen, the  
10 Netherlands

11 2 Department of Biological and Environmental Science, University of Jyväskylä,  
12 Jyväskylä, Finland

13 3 Health & Environment Department, AIT Austrian Institute of Technology, Tulln,  
14 Austria

15  
16 **Correspondence :**

17 Manoj Kumar, [makugopa@jyu.fi](mailto:makugopa@jyu.fi)

18  
19  
20  
21 **5918 words, 7 Figures**

22  
23 Keywords: endophytic bacteria, arcto-alpine plant, biogeographical diversity, core  
24 bacteriome, *Oxyria digyna*, *Saxifraga oppositifolia*.

25  
26 **Running title:** Bacterial communities in arcto-alpine plants.

27 **Abstract**

28 Evidence for the pivotal role of plant-associated bacteria to plant health and  
29 productivity has accumulated rapidly in the last years. However, key questions related  
30 to what drives plant bacteriomes remain unanswered, among which is the impact of  
31 climate zones on plant-associated microbiota. This is particularly true for wild plants  
32 in arcto-alpine biomes. Here, we hypothesized that the bacterial communities  
33 associated with pioneer plants in these regions have major roles in plant health  
34 support, and this is reflected in the formation of climate and host plant specific  
35 endophytic communities. We thus compared the bacteriomes associated with the  
36 native perennial plants *Oxyria digyna* and *Saxifraga oppositifolia* in three arcto-alpine  
37 regions (alpine, low Arctic and high Arctic) with those in the corresponding bulk  
38 soils. As expected, the bulk soil bacterial communities in the three regions were  
39 significantly different. The relative abundances of *Proteobacteria* decreased  
40 progressively from the alpine to the high-arctic soils, whereas those of Actinobacteria  
41 increased. The candidate division AD3 and *Acidobacteria* abounded in the low Arctic  
42 soils. Furthermore, plant species and geographic region were the major determinants  
43 of the structures of the endosphere communities. The plants in the alpine region had  
44 higher relative abundances of *Proteobacteria*, while plants from the low- and high-  
45 arctic regions were dominated by *Firmicutes*. A highly-conserved shared set of  
46 ubiquitous bacterial taxa (core bacteriome) was found to occur in the two plant  
47 species. *Burkholderiales*, *Actinomycetales* and *Rhizobiales* were the main taxa in this  
48 core, and they were also the main contributors to the differences in the endosphere  
49 bacterial community structures across compartments as well as regions. We postulate  
50 that the composition of this core is driven by selection by the two plants.

51

Provisional

52 **Introduction**

53 Among the terrestrial environments on Earth, arctic and alpine ecosystems cover  
54 about 8% of the global land area, which is more than the area covered by tropical  
55 forests (Chapin and Körner, 1996). These arctic and alpine ecosystems have the least  
56 biologically usable heat and the lowest diversity of plants (Billings and Mooney,  
57 1968). The plants in these systems are well adapted to cold and short growing seasons  
58 and low-nutrient soils. The typical plant species occurring in both biomes are  
59 collectively referred to as arcto-alpine vegetation. These plants are important in these  
60 soils, as they constitute the prime settlers that are at the basis of the local living  
61 ecosystem. It has been hypothesized that the local microbiota plays an important role  
62 in the ecological success of these pioneering plants (Borin et al., 2010; Mapelli et al.,  
63 2011). While arctic and alpine biomes share many characteristics, including short and  
64 cool growing seasons, cold winters and soils with low levels of nutrients, there are  
65 also distinct differences: the alpine biome is characterized by high annual and diurnal  
66 temperature fluctuation and high solar intensity in the summer and, in general, well-  
67 drained soils. The vegetation in the Arctic, on the other hand, experiences weeks to  
68 months long polar night in winter and 24-hour daylight during the growing season.  
69 Moreover, arctic soils are typically water-logged due to underlying permafrost  
70 (Körner, 2003). These differences have led to ‘climatic ecotypes’ within arcto-alpine  
71 vegetation, where the growth morphology and phenology of the same plant species in  
72 different biomes reflect adaptation to distinct climates.

74 Endophytic bacteria are ubiquitous across both cultivated and wild plants. They have  
75 been shown to contribute to major aspects of plant life, including regulation of growth  
76 and development, nutrient acquisition and protection from biotic and abiotic stressors  
77 (reviewed in Hardoim et al., 2015). Studies conducted mainly in agricultural or model  
78 plant systems have offered a rapidly growing insight into the assembly, structure and  
79 function of the endophytic communities in plants (Compant et al., 2010; Zhang et al.,  
80 2006). Soil type and plant species and genotype are both known to shape the  
81 rhizosphere (Berg and Smalla, 2009; Garbeva et al., 2004) and root endosphere  
82 communities (Bulgarelli et al., 2013). Rhizosphere soil is considered to be the main  
83 source of endophytes (Bulgarelli et al., 2013), but vertical transmission via seeds has  
84 also been reported (Puente et al., 2009; Hardoim et al., 2012). However, the factors  
85 governing the plant-associated microbiota of perennial wild plants in the  
86 aforementioned arcto-alpine soils may differ from those of model or well-fertilized  
87 crop plants. For plants in the low-arctic fell tundra, we have previously shown that  
88 plant species, rather than sampling site, determines the structure of the endophytic  
89 (Nissinen et al., 2012) and rhizospheric (Kumar et al., 2016) microbial communities.

91 Most bacterial species are considered to be cosmopolitan, as they have been found  
92 across biogeographic regions in habitats like soils, sediments, lakes and the sea  
93 (Hanson et al., 2012). Interestingly, the bacterial diversity in arctic soils has been  
94 shown not to differ from that of other biomes (Chu et al., 2010). With respect to  
95 community structure, endemism per region has been observed for bacteria, with some  
96 taxa reportedly being restricted to distinct geographical regions (Cho, 2000; Oakley et  
97 al., 2010).

99 The main goal of this study was to investigate the factors that shape the bacterial  
100 communities associated with two plant species in three geographic regions, from the  
101 high Arctic to the Alps. Our target plant species, *Oxyria digyna* and *Saxifraga*

102 *oppositifolia*, are arcto-alpine plant species with wide distribution from the high  
103 Arctic to the mid-latitude alpine tundra. Both are typical pioneer species that  
104 efficiently colonize low-nutrient tundra soils. *O. digyna* is a member of the  
105 Polygonaceae (order Caryophyllales), whereas *S. oppositifolia* belongs to the order  
106 Saxifragales, which diverged from other core eudicots 114-124 MYA (Soltis et al.,  
107 2000; Wikström et al., 2001). We focused on the root endophytic bacteria, and also  
108 examined the bacterial communities in the relevant rhizosphere and bulk soil samples.

109  
110 We hypothesized that (1) geographic region, related to climate zone, determines the  
111 diversity and community structure of the soil bacterial communities in the selected  
112 habitats, and (2) plants strongly shape the plant-associated communities, resulting in  
113 plant species specificity, regardless of the geographic region. We also hypothesized  
114 that (3) part of the plant-associated bacteria are consistently present in their hosts,  
115 constituting an endophytic core microbiome.

116  
117 To achieve our aims, we used community DNA based amplicon sequencing targeting  
118 the bacterial 16S rRNA gene region and subsequent analyses.

## 119 120 **Materials and Methods**

### 121 122 **Sampling locations and study sites**

123 Plant and soil samples were collected from eight sampling sites in three distinct  
124 regions representing different climate zones; Ny-Ålesund, high Arctic (3 sampling  
125 sites), Kilpisjärvi, low Arctic (3 sampling sites) and Mayrhofen, European Alps (2  
126 sampling sites) (Figure 1A). Kilpisjärvi is at the northwestern Finland and located  
127 along the Fenno-Scandinavian border. Its flora is dominated by mountain birch forest  
128 in the valleys and by fell tundra at higher elevations. The annual mean temperature is  
129 about -2.2°C with plant growth season of ca. 90-100 days. Ny-Ålesund (Svalbard,  
130 Norway) is located on an isolated archipelago in the high Arctic; the land cover is  
131 dominated by glaciers and permafrost layers, and the mean annual temperature is -  
132 4°C. The soil temperatures have been reported to be below zero for more than 250  
133 days per year ranging from -6°C to -25°C (Coulson and Hodkinson, 1995). Most  
134 biological activity is restricted to less than 10% of the total land mass coupled with  
135 about three months of plant growing season. The sampling location in the Mayrhofen  
136 is located above the tree line south of Mayrhofen over the snow-covered mountains  
137 (altitude ca. 2400 m above sea level) in the alpine tundra of the European Alps.  
138 Coordinates and details of sampling sites are listed in Supplemental Table S1.

### 139 140 **Sample collection and processing**

141 12 replicates of bulk soil samples (the top 5 cm soil was removed and soil samples  
142 from 5-10 cm and 10-15 cm, corresponding to major root mass of target plant species,  
143 were both used for analysis) and six samples of both *O. digyna* and *S. oppositifolia*  
144 (as whole plants with adhering rhizosphere soils) were collected from all sites, except  
145 in site “Saana” (Kilpisjärvi) where only *O. digyna* plants were sampled and site  
146 “Cliff” (Mayrhofen) where we sampled only 6 bulk soil samples. Sampling was  
147 performed during summer 2012. All harvested plants were flowering at the time of  
148 the sampling. Rhizosphere and bulk soil samples were processed and stored as  
149 specified by Kumar et al. (2016). After removing rhizosphere soils, plant roots were  
150 thoroughly washed with water and surface sterilized by immersing the plant material  
151 into 3% sodium hypochlorite for 3 minutes and then subsequently in sterile double

152 distilled water (3 x 90 s). 80-100 mg of root samples were weighed, snap frozen with  
153 liquid nitrogen and stored at -80°C for further DNA analysis.  
154 Soil pH and soil organic matter (SOM) content were measured as described in Kumar  
155 et al. (2016), while available phosphorous (P) was measured based on Bray No 1  
156 extraction method (Bray and Kurtz, 1945). All the soil chemical analyses were  
157 performed in duplicates (2 technical replicates) per sample, and with 4-8 biological  
158 replicates per site and sample type (Table 1).

159

#### 160 **DNA isolation**

161 Microbial DNA from soil samples were extracted following manufacturer's  
162 instruction using MoBio Power soil kit (MoBio, Carlsbad, CA USA). For soil  
163 samples 0.5 g of soil was used instead of 0.25 g because of low microbial counts in  
164 our soils (data not shown). For isolation of endophyte samples, Invisorb Spin Plant  
165 Mini Kit (STRATEC Biomedical AG, Germany) was used in order to ensure  
166 prolonged stability of endophytic DNA in the plant derived samples. Frozen plant  
167 tissues were homogenized by bead beating for 45 s with 0.1mm sterilized glass beads  
168 with FastPrep homogenizer (mpbio.com), followed by DNA extraction according to  
169 manufacturer's protocol.

170

#### 171 **16S rRNA gene library generation and sequencing**

172 After isolating DNA from all six plant replicates from both plant species, four (rhizo-  
173 and endosphere) or eight (bulk soil) samples technically best samples (good DNA  
174 yield, good PCR amplification) were included in the next generation sequencing  
175 library construction. 16S rRNA gene was amplified using primers 799f/1492r (Chelius  
176 and Triplett, 2001) and M13-1062f/1390r in a nested approach. The nested primers  
177 targeting the V6-V8 regions of 16s rRNA gene enable elimination of plant chloroplast  
178 16S rRNA gene amplicons as well as separation of endophyte amplicons from plant  
179 mitochondrial amplicons by size fractionation (799f-1492r, Chelius and Triplett  
180 (2001)) and produce an amplicon with high phylogenetic coverage and optimal size  
181 for IonTorrent sequencing (1062f-1390r). Primers 1062f (Ghyselinck et al., 2013) and  
182 1390r (Zheng et al., 1996) were tagged with M13 sequences to enable sample  
183 barcoding as described below and in Mäki et al. (2016). Both reactions had 1 µl of  
184 sample DNA, 1x PCR buffer, 1 mg/ml of BSA, 0.2 mM dNTP's, 0.3 µM of each  
185 primer and 1250 U/ml GoTaq DNA Polymerase (Promega, WI USA) in a 30µl  
186 reaction volume. 5-10 and 25-30 ng of soil and endophyte DNA, respectively, was  
187 used in the first PCR, and 1µl of 1:10 diluted amplicons (for bulk and rhizosphere soil  
188 samples) and 1 µl of amplicons (for endosphere samples) from the first PCR were  
189 used as a template for the second run. Amplifications for both PCR reactions were  
190 performed as follows: 3 mins denaturation at 95°C followed by 35 cycles of  
191 denaturing, annealing, and extension at 95°C for 45 secs, 54°C for 45 secs and 72°C  
192 for 1 min, respectively. Final extension was carried out at 72°C for 5 mins. Prior to  
193 library production, the PCR protocol was optimized with regard to several primer pair  
194 combinations, PCR protocols and test of PCR blockers to minimize the strong  
195 interference of mitochondrial rRNA in *O. digyna* and *S. oppositifolia*. The above  
196 described protocol, using high coverage, minimal bias primer pairs, was shown to  
197 produce enough eubacterial (endophytic) amplicons with no observable decline in  
198 diversity (as detected by T-RFLP) for sequencing, while most alternatives lead to very  
199 low amplification levels endophytes and strong mitochondrial signal.

200



201 Sequence libraries were prepared by running a third PCR to attach the M-13 barcode  
202 system developed by Mäki et al. (2016). Amplicons from second PCR were diluted  
203 1:5 and re-amplified using barcode attached M13 system as forward primer and  
204 1390r-P1 with adaptor A as a reverse primer. PCR mix and conditions were similar as  
205 described above, with an exception of using 8 cycles for amplification. Amplified  
206 libraries were purified using Agencourt AMPure XP PCR purification system  
207 (Beckman Coulter, CA USA). Purified samples were quantified with Qubit  
208 Fluorometer (Invitrogen, MA USA) and an equivalent DNA quantity of each sample  
209 was pooled together. The pooled samples were then size fractionated (size selection  
210 range of 350-550 bp) using Pippin Prep (Sage Science, MA USA) 2% Agarose gel  
211 cassette (Marker B) following the manufacturer's protocol. Size fractionated libraries  
212 were sequenced using Ion 314 chip kit V2 BC on Ion Torrent PGM (Life  
213 Technologies, CA USA) in Biocenter Oulu, Finland.

214

### 215 **Bioinformatics and statistical analysis**

216 The raw sequence reads were processed using QIIME (Caporaso et al., 2010) and  
217 UPARSE (Edgar, 2013) based on a 16S rRNA gene data analysis pipeline developed  
218 by Pylro et al. (2014) with slight modifications in quality filtering. Sequences were  
219 trimmed by removing sequences with low quality reads (Q score <25) and shorter  
220 base pair (<150) length. Furthermore, all the raw reads were trimmed (200 bp),  
221 aligned and clustered at 97% identity using USEARCH algorithm (Edgar, 2010).  
222 UCLUST algorithm along with Greengenes database (DeSantis et al., 2006) was used  
223 to assign taxonomies at 97% identity to the individual OTUs. In total, 426,135 high-  
224 quality reads (1468 reads - 5331 reads per sample) were clustered into 985 OTUs. For  
225 alpha diversity analysis all the samples were rarefied (subsampling) to 1400 reads per  
226 sample. Shannon index and species richness were obtained using Univariate Diversity  
227 Indices (DIVERSE, PRIMER 6 (PRIMER-E Ltd)). The differences in diversity  
228 indexes between the soil samples and their correlation with soil physico-chemical  
229 properties were determined using two-way ANOVA and Pearson correlation (SPSS  
230 Statistics, IBM). The significance of the differences between the soil samples were  
231 tested by Games-Howell post-hoc tests (two-way ANOVA).

232

233 To normalize the data for community structure and other analyses all the samples with  
234 more than the median reads were rarefied to the median (2780 reads), while the  
235 samples with less reads were used as such, as described in deCárcer et al. (2011). In  
236 addition, all the singletons and OTUs with less than 50 reads were removed before  
237 processing. The influence of sampling site, geographic region, plant compartment and  
238 plant species on bacterial community structures, based on Bray-Curtis distance  
239 matrixes of square root transformed abundance data, were analysed using  
240 permutational multivariate analysis of variance (PERMANOVA) and visualized by  
241 PCoA ordinations at the OTU level. Taxonomic groups (phyla or OTU) with strongest  
242 impact on significant differences between community structures were identified with  
243 SIMPER (Similarity Percentages - species contributions), all performed with  
244 PRIMER 6 software package with PERMANOVA+ add-on (primer-e.com).

245

246 All the Ternary plots were made by calculating the mean relative abundances of  
247 OTUs per geographic region/compartment and with the function 'ternaryplot' 'vcd'  
248 (Meyer et al., 2015) from the R package. All other graphs (bar and scatter plots), also  
249 based on the mean relative abundances of taxa, were constructed using the R package  
250 'graphics'.

251

## 252 **Picking endosphere core OTUs**

253 The **highly conserved OTUs (core OTUs)** were manually picked by selecting OTUs  
254 that were constantly observed (present in at least 3 out of 4 replicates per site) in the  
255 endosphere of either *O. digyna* and *S. oppositifolia* or both. To determine the  
256 distribution of core OTUs reads across different compartments, the averaged read  
257 count per compartment was calculated for each OTU, the averages were summed up  
258 and presented as the relative distribution of each of the 13 core OTUs per  
259 compartment.

260

## 261 **Results**

262

263 A total of 426,135 quality-filtered sequence reads was retrieved from the total of 174  
264 samples in our sample set, representing the endospheres and rhizospheres of the two  
265 plant species and the corresponding bulk soils from the three geographic regions  
266 (Figure 1A). These sequences were separated into 985 OTUs (defined at the 97% cut-  
267 off level) and subjected to downstream analyses. Of these, 933 OTUs were present in  
268 at least one of the samples from each of the three regions, 43 in two regions and nine  
269 were restricted to one region only. 778 of the 985 OTUs were found in all  
270 compartments (bulk soil, rhizosphere soil or endosphere), 190 in two compartments  
271 and 17 were compartment-specific (Figure 1B).

272

## 273 **Soil characteristics are different across three arcto-alpine regions**

274 Table 1 lists the soil characteristics in the three geographic regions: Mayrhofen  
275 (alpine), Kilpisjärvi (low-arctic) and Ny-Ålesund (high-arctic) (Figure 1A). The Ny-  
276 Ålesund [bulk] soils had significantly higher pH (two-way ANOVA,  $p < 0.05$ ) and soil  
277 organic matter (SOM) values (two-way ANOVA,  $p < 0.01$ ), and significantly lower  
278 levels of available phosphorus (two-way ANOVA,  $p < 0.05$ ) than the Kilpisjärvi and  
279 Mayrhofen soils. The Kilpisjärvi soils had the lowest average pH values, but there  
280 were no significant differences in the other physico-chemical properties between the  
281 Kilpisjärvi and Mayrhofen bulk soils.

282

## 283 **Geographical region and soil properties impact the diversity of the bulk soil, but not of the rhizosphere or endosphere bacterial communities**

284 The species richness (SR) and  $\alpha$ -diversity (Shannon index, SI) values of the bulk soil  
285 bacterial communities differed between the geographic regions. The Kilpisjärvi bulk  
286 soils (SR=33.95, SI=4.21) had significantly lower richness (two-way ANOVA,  
287  $p < 0.01$ ) and diversity (two-way ANOVA,  $p < 0.01$ ) values than the Mayrhofen  
288 (SR=41.04, SI=4.61) and Ny-Ålesund bulk soils (SR=42.55, SI=4.92) (Figure 2A). In  
289 contrast, there were no significant differences in the diversity levels of the rhizosphere  
290 soil or endosphere samples between the regions (Figure 2A).

291

292  
293 There was a significant positive relationship of both SR and SI with soil pH (Pearson  
294 correlation [2-tailed], SR  $p < 0.001$ , SI  $p < 0.001$ ), and a negative one with the levels of  
295 available phosphorus (P) (Pearson correlation [2-tailed], SR  $p < 0.014$ , SI  $p < 0.001$ ).  
296 There was a significant positive correlation between SOM and SI, but not between  
297 SOM and SR (Figure 2B).

298

## 299 **Bacterial community structures in samples from different regions differ at the** 300 **phylum level**

301  
302 Collectively, the OTUs from all our samples fell into 21 bacterial phyla. Eight of  
303 these, i.e. *Proteobacteria*, *Actinobacteria*, *Acidobacteria*, candidate division *AD3*,  
304 *Bacteroidetes*, *Firmicutes*, *Chloroflexi* and *Gemmatimonadetes*, were prominent,  
305 collectively making up about 97% of the total microbiome. The remaining 13 phyla  
306 were present at less than 1% relative abundance each.

307  
308 Bacterial community structures in the samples from the different regions were  
309 significantly different at phylum level (PERMANOVA  $F=8.1155$ ,  $P=0.001$ ). SIMPER  
310 analyses confirmed that *Proteobacteria*, *Acidobacteria*, *AD3* and *Actinobacteria* were  
311 the main phyla contributing to the overall dissimilarities between the regions (Table  
312 2). *Proteobacteria* were relatively more abundant in the alpine (Mayrhofen; average  
313 relative abundance 57%) than in the arctic regions (46% in Kilpisjärvi and 43% in  
314 Ny-Ålesund). The phylum *Acidobacteria* and the candidate division *AD3* were  
315 observed in higher average relative abundances in Kilpisjärvi than in the other two  
316 regions (Figure 3A, 3B). The *AD3* candidate division had reduced diversity (Figure  
317 3A), with a single abundant OTU (OTU 10) dominating the Kilpisjärvi bulk soil  
318 samples, representing about 25% of the total bulk soil community. The Ny-Ålesund  
319 samples were enriched with *Actinobacteria* (Figure 3A), with average relative  
320 abundances in Kilpisjärvi=14%, Mayrhofen=14% and Ny-Ålesund=23%.

321  
322 The increased relative abundances of *Proteobacteria* in Mayrhofen, *Acidobacteria* and  
323 *AD3* in Kilpisjärvi and *Actinobacteria* in Ny-Ålesund were also consistent in the  
324 communities in the different compartments (bulk soil, rhizosphere soil, endosphere)  
325 (Supplemental file S2), with the exception of *AD3*, which was present at very low  
326 abundances in the endosphere (<0.9%) in all three regions (Table 2). Additionally, the  
327 relative abundances of *Firmicutes* in the endosphere samples increased with  
328 increasing latitude, being lowest in Mayrhofen and highest in Ny-Ålesund.

329  
330 **Firmicutes, Proteobacteria and Bacteroidetes dominate endosphere communities**  
331 Bacterial community structures were clearly different in the different compartments at  
332 the phylum level (PERMANOVA pseudo- $F=64.371$ ,  $P=0.001$ ; Table 2). These  
333 differences were mainly driven by strong relative enrichment of *Firmicutes* in the  
334 endosphere-derived sequence data sets, compared to their very low abundances in the  
335 bulk and rhizosphere soils (Figure 4, Table 2). The relative abundances of  
336 *Proteobacteria* and *Bacteroidetes* increased progressively from bulk to rhizosphere  
337 soil to the endosphere, with a concomitant decrease in those of candidate division  
338 *AD3*, *Gemmatimonadetes* and *Chloroflexi*, which collectively constituted <4% of  
339 endosphere communities (Table 2, Figure 4B). This trend was similar in all three  
340 geographic regions.

341  
342 The divergence of the endosphere communities from the soil communities was also  
343 evident at the class level. For example, OTUs representing the actinobacterial class  
344 *Thermoleophilia* were abundant in the bulk and rhizosphere soil communities,  
345 whereas these were rare in the endosphere communities. The latter were dominated by  
346 the class *Actinobacteria* (Class, order and family level analyses in Supplementary  
347 data S2).

348  
349 **Compartment impacts bacterial diversity and community structures more than**  
350 **geographic region or sampling site**

351 The diversity values of the endosphere communities, analyzed at the OTU level, were  
352 significantly lower than those of the bulk or rhizosphere soil communities (Figure  
353 2A). The rhizosphere soils had the highest diversity values, but the differences  
354 between rhizosphere and bulk soils were not significant (two-way ANOVA ,  $p>0.05$ ,  
355 Figure 2A). However, we observed no such differences between the plant species, as  
356 *O. digyna* and *S. oppositifolia* had similar SR and SI indices in the rhizo- as well as  
357 the endosphere communities.

358  
359 Also, the community structures of the endosphere bacterial communities differed  
360 clearly from those of the bulk and rhizosphere soil ones across all three regions, as  
361 demonstrated by PCoA (Figure 5A). A separate analysis of the soil-derived samples  
362 revealed that the bulk soil communities diverged from the rhizosphere soil ones in the  
363 three regions (Figure 5B). This was supported by PERMANOVA, where  
364 compartment was identified as a significant and strong driver of the differences  
365 between bacterial community structures (Pseudo-F=30.962,  $P<0.001$ , Table 3). Pair-  
366 wise analyses of the community structures supported the PCoA analyses, with  
367 significant ( $P=0.001$ ) differences between the endosphere and bulk soil, endosphere  
368 and rhizosphere and rhizosphere and bulk soil communities, and t-values of 6.339,  
369 6.39 and 3.134, respectively.

370  
371 In addition to compartment, sampling site (pseudo-F=4.0646,  $P<0.001$ ) and region  
372 (Pseudo-F=6.5495,  $P<0.001$ ) both had significant effects on the bacterial community  
373 structures, although these factors had less impact than compartment (Table 3).  
374 PERMANOVA performed on each of the compartments separately revealed that  
375 region and sampling site had the greatest influence on the structure of bulk soil  
376 communities (Pseudo-F=9.1503,  $P<0.001$  and Pseudo-F=7.6707,  $P<0.001$ ,  
377 respectively) with their influence decreasing for the rhizosphere (Pseudo-F=5.9962,  
378  $P<0.001$  and Pseudo-F=5.0728,  $P<0.001$ , respectively) and endosphere (Pseudo-  
379 F=2.7877,  $P<0.001$  and Pseudo-F=2.1418,  $P<0.001$ , respectively) (Table 3).  
380 Interestingly, region shaped community structures more than sampling site for all  
381 compartments, indicating an impact of bioclimatic conditions (Table 3). Thus, in  
382 further analyses, we focused on comparing communities from the different regions  
383 and plant species.

### 384 **Plant species and region both impact endosphere bacterial community structures**

385 PERMANOVA identified both plant species and geographic region as significant  
386 drivers of the community structures of the rhizosphere soil communities  
387 (PERMANOVA  $p<0.01$ ), but region (Pseudo-F= 5.8857) had more impact on the  
388 differences than plant species (Pseudo-F=2.9879) (Table 3). In contrast, while plant  
389 species, region and their interaction all had significant impact on endosphere  
390 community structures (PERMANOVA,  $P<0.01$ ), plant species had stronger impact on  
391 the differences between the communities (Pseudo-F=4.0332) than region or  
392 interaction between these factors (Pseudo-F=2.9678 and Pseudo-F=1.6249,  
393 respectively) (Table 3). The endosphere communities from all three regions, being  
394 relatively similar to each other, tended to diverge based on plant species (*O. digyna* or  
395 *S. oppositifolia*) on the first two axes in the PCoA ordination (Figure 5C), while we  
396 did not observe plant species specific clustering in the PCoA of the corresponding  
397 rhizosphere communities (data not shown).

398  
399

400 On the basis of the above analyses, we found partial support for our hypothesis that  
401 plant species strongly shape the plant-associated bacterial communities, as this factor  
402 emerged as the major (albeit not the only) significant driver of the endosphere  
403 bacterial community structures over multiple sites and several regions (climate  
404 zones). Plant species also had a small, but significant impact on the rhizosphere  
405 community structures, but these were mainly determined by geographic factors.  
406

407 **Differences in the endosphere bacterial community structures between the two**  
408 **plant species are explained by differential acquisition of shared bacterial taxa**

409 Remarkably, the majority of the endosphere bacterial taxa was present in both plant  
410 species, but in different relative abundances. A total of 841 OTUs was found in the  
411 endosphere samples, comprising 152,050 reads. A vast majority, i.e. 612 OTUs  
412 (149,422 reads, 98.3% of all endosphere reads), was shared between the two plant  
413 species (Figure 6A), and many of these OTUs were consistently enriched along plant  
414 species. For example, OTUs representing *Sphingobacteriales* (*Sphingobacteriia*,  
415 *Bacteroidetes*), *Burkholderiales* ( $\beta$ -proteobacteria) and *Bradyrhizobiaceae* were  
416 enriched in the *O. digyna* samples, while OTUs in the *Clostridiales*, along with  
417 *Actinobacteria*, and Acidimicrobiia as well as several OTUs representing  
418 *Myxococcales* and *Saprospirales* were relatively more abundant in *S. oppositifolia*  
419 across the three climate zones (Figure 6B, 6C, Supplemental data S2). These were  
420 also identified as the main OTUs responsible for plant species specific community  
421 structures in the SIMPER analysis (Table 4). In addition to the shared bacterial taxa,  
422 162 OTUs (1,749 reads, 1.1% of the total endosphere reads) and 57 OTUs (879 reads,  
423 0.6%) were observed only in *O. digyna* and *S. oppositifolia*, respectively (Figure 6A).  
424

425 **Thirteen bacterial taxa are highly conserved in the *O. digyna* and *S. oppositifolia***  
426 **endosphere communities in all three regions, constituting a major portion of**  
427 **these**

428 We examined the bacterial taxa that were highly conserved (belonging to the ‘tight’  
429 core) in the *O. digyna* or *S. oppositifolia* endospheres using as a criterion ‘OTUs  
430 present in at least three out of four endosphere samples per plant species across all  
431 sampling sites and regions’. Thirteen such OTUs were found, of which five,  
432 representing *Bradyrhizobium* (2 OTUs), *Rhodoplanes* ( $\alpha$ -Proteobacteria),  
433 *Janthinobacterium* ( $\beta$ -Proteobacteria) and *Planococcaceae* (*Firmicutes*), were  
434 consistently present in both plant species (Table 5). Additionally, eight OTUs were  
435 consistently present in just one of the plant species. Thus *O. digyna* specific core  
436 OTUs belonged to *Comamonadaceae* ( $\beta$ -Proteobacteria) and *Enterobacteriaceae* ( $\gamma$ -  
437 Proteobacteria), whereas *S. oppositifolia* specific core OTUs belonged to  
438 *Micromonosporaceae*, *Micrococcaceae* (*Actinobacteria*), *Bradyrhizobiaceae* ( $\alpha$ -  
439 Proteobacteria) and unidentified  $\beta$ -Proteobacteria (Table 5). Collectively, these  
440 (highly conserved) core OTUs accounted for 38% of the total reads in the endosphere  
441 communities. Significantly, eleven of these core OTUs (all except OTUs 171 and 429,  
442 Table 5) were among the main drivers of the divergence of the endosphere  
443 communities of the two plant species (Table 4). They also explained the differences  
444 between the endosphere and the soil bacterial communities, and those between the  
445 endosphere communities in the different geographical regions (Table 4). Of the 13  
446 core OTUs, 11 were predominantly present in the plant associated compartments, as  
447 over 75% of their reads were detected in the endosphere, and over 80% in the endo-  
448 or rhizosphere (Figure 7).  
449

## 450 Discussion

451

### 452 Factors shaping the bacterial diversity in soils across three climatic regions

453 In this study, we examined the bacterial communities in three regions spanning over  
454 3000 km in distance, i.e. Mayrhofen (alpine), Kilpisjärvi (low-arctic) and Ny-Ålesund  
455 (high-arctic). In these three regions, the climatic conditions are clearly different. The  
456 highest bacterial species richness and diversity values in the bulk soils were found in  
457 the Ny-Ålesund samples, which was consistent with data by Chu et al. (2011) and  
458 Neufeld and Mohn (2005) who also detected highest bacterial diversities in high  
459 northern latitudes. However, our data stand in contrast to those from Yergeau et al.  
460 (2007), who reported decreasing bacterial diversities in Antarctic soils with increasing  
461 latitude towards the south pole. We found a clear positive correlation of bacterial  
462 diversity with soil pH and SOM, and a negative correlation with the level of available  
463 P, agreeing with studies that put forth soil pH as a major driver of bacterial diversity  
464 (Fierer and Jackson, 2006; Fierer and Lennon, 2011; Lauber et al., 2009; Rousk et al.,  
465 2010; Shi et al., 2015). Soil nutritional status and available P have also been shown to  
466 significantly impact bacterial diversity (Siciliano et al., 2014).

467

468 With respect to compartment, the endosphere bacterial communities were  
469 significantly less diverse than those in the corresponding soils. However, in contrast  
470 with studies from other soils (İnceoğlu et al., 2011; Kowalchuk et al., 2002; Smalla et  
471 al., 2001), where rhizosphere soil communities have been reported to be less diverse  
472 and rich than bulk soil ones, we observed a trend towards higher richness and  
473 diversity in the rhizosphere than in the corresponding bulk soils, although these  
474 differences were not statistically significant. This trend was similar to findings in a  
475 previous study from the Kilpisjärvi site, where the rhizosphere samples had highest  
476 richness and diversity (Kumar et al., 2016). Miniaci et al. (2007) and Coleman-Derr et  
477 al. (2016), studying low-SOM glacier forefield or desert soils, respectively, also  
478 observed higher bacterial diversity and richness values in the rhizospheres than in the  
479 corresponding bulk soils. Further, Yergeau et al. (2007) found that, although soil  
480 bacterial diversities in unvegetated Antarctic fell-field soils decreased with increasing  
481 (southern) latitude, those from vegetated sites did not. This suggests that a plant-  
482 incited “protective or nutritional” effect on bacterial communities becomes  
483 increasingly more important in soils in which conditions are challenging.

484

### 485 Specific OTUs determine the divergence of the soil bacteriomes across three 486 regions

487 In this study, we detected only few ‘endemic’ bacterial OTUs, as the great majority of  
488 the bacterial taxa was found in all three, geographically distant, regions. However,  
489 these taxa were present in very different relative abundances, leading to region-driven  
490 community structures. Roughly, proteobacterial taxa decreased and Gram-positive  
491 ones increased towards the north, with *Acidobacteria* and candidate division *AD3*  
492 being enriched in the Kilpisjärvi samples. This clear progressive change in bacterial  
493 community structures hints at specific effects of the shifting local conditions on the  
494 aforementioned taxa. Thus, habitat filtering rather than [long-distance] dispersal  
495 impacts the bacterial community compositions across the three cold climate sites.  
496 The dominance of *Proteobacteria* in the bulk soil samples from Mayrhofen was  
497 consistent with findings by Margesin et al. (2009) in alpine soils. Moreover,  
498 corroborating earlier studies (Männistö et al., 2007, 2013), the high abundance of  
499 *Acidobacteria* was likely linked to the low pH in the Kilpisjärvi soils (Chu et al.,

500 2010; Griffiths et al., 2011). Also, the high abundance of candidate division *AD3* in  
501 Kilpisjärvi (Figure 5b) was consistent with similar findings for the Mitchell peninsula  
502 in Antarctica (Ji et al., 2015) and low-nutrient sandy soils (Zhou et al., 2003).  
503 However, earlier studies by Männistö et al. (2007, 2013) have not detected candidate  
504 division *AD3* in high-SOM Kilpisjärvi soils. We here assume that the candidate  
505 division *AD3* members that were found are well adapted to the [low SOM/ low  
506 nutrient] soils. Alternatively, their absence from the previous data sets might be due to  
507 different 16S rRNA targeting primers used in the different studies.

508

### 509 **Compartment is the primary driver of bacterial community structures**

510 A striking observation was that both *O. digyna* and *S. oppositifolia* sampled in any of  
511 the three regions exhibited quite similar endosphere bacterial communities. We  
512 previously observed compartmental influence between bulk and rhizosphere soils of  
513 *O. digyna* and *S. oppositifolia* (Kumar et al., 2016), and so extended this to the  
514 endospheres that were addressed in the current study. Clearly, even though the bulk  
515 soil bacterial communities were influenced by region and sampling site, which may  
516 relate to soil edaphic factors, the plant endospheres shared similar bacterial  
517 endophytes across the three regions. This points to a strong and specific filtering  
518 effect of the two pioneering plants that were studied, allowing similar bacteria to  
519 colonize plants from the widely divergent soils in different regions.

520

521 As a token of the plant-incited filtering effect, members of the *Proteobacteria*,  
522 *Actinobacteria*, *Bacteroidetes* and *Firmicutes* dominated the endosphere bacterial  
523 communities. Several other studies, performed with both agricultural and wild plants,  
524 also reported these four taxa to be dominant in several endospheres, with  
525 *Proteobacteria* being the most dominant one (Coleman-Derr et al., 2016; Santoyo et  
526 al., 2016; Zhao et al., 2016). Other taxa, including *Acidobacteria*, candidate division  
527 *AD3*, *Chloroflexi* and *Gemmatimonadetes*, were virtually absent from the endosphere.  
528 A general underrepresentation of *Acidobacteria* in the endosphere has also been  
529 observed in other systems (Coleman-Derr et al., 2016; Edwards et al., 2015;  
530 Zarraonaindia et al., 2015). The enrichment of *Firmicutes* in the endosphere samples  
531 in this study was mainly ascribed to the raised abundance of OTUs belonging to the  
532 *Clostridia* (in particular OTU 21; genus *Clostridium*). Possibly, such organisms might  
533 have been selected for their capacities to fix nitrogen in the cold and often water-  
534 logged soils (Rosenblueth and Martínez-Romero, 2006) in the permafrost-impacted  
535 Arctic sites. This hypothesis is supported by our [unpublished] observations, that *nifH*  
536 gene libraries prepared from the same plants as used in the current study are  
537 dominated by *Clostridium*-type genes in the (high) Arctic. Although *Clostridium* has  
538 been described as a strictly anaerobic genus, members of this genus have been shown  
539 to fix nitrogen in rice roots (Minamisawa et al., 2004), and survive in the potato  
540 endosphere in aerobic conditions (Shabuer et al., 2015). Interestingly, *nifH* genes of  
541 *Clostridium* spp. have been reported to be frequent in soil samples from the Canadian  
542 high Arctic (Deslippe and Egger, 2006). Similarly, in our study, plants were sampled  
543 in early growing season, when these started flowering and snow was melting in most  
544 sampling sites.

545

### 546 **A small set of highly conserved OTUs shapes the endosphere bacterial** 547 **communities in two arcto-alpine plant species**

548 *O. digyna* and *S. oppositifolia*, the target plant species in this study, are both perennial  
549 herbs with similar habitat requirements, producing tap root systems of similar size and

550 depth; the plants often grow at close proximity to each other. However, they are  
551 taxonomically quite distant (Soltis et al., 2000; Wikström et al., 2001) and have  
552 differing mycorrhizal associations. *O. digyna* is non-mycorrhizal, whereas *S.*  
553 *oppositifolia* is endomycorrhizal, which is likely to have strong impact on its nutrient  
554 acquisition efficiency.

555  
556 Despite these differences, the endosphere communities of these two plants were  
557 strikingly similar. While we did find an effect of plant species on the endosphere  
558 community structures (Table 3, Figure 6c), the plants shared a core microbiome,  
559 dominated by *Burkholderiales*, *Actinomycetales* and *Rhizobiales*, across plants in the  
560 three arcto-alpine climatic regions. Of these, *Actinomycetales* and *Burkholderiales*  
561 have been reported as components of the core root microbiome of, e.g., *A. thaliana*  
562 (Schlaeppli et al., 2014). *Rhizobiales* are known plant symbionts with nitrogen fixing  
563 abilities, while *Burkholderiales* are well known for their biodegradative capacities and  
564 antagonistic properties towards multiple soil-borne fungal pathogens (Benítez and  
565 McSpadden Gardener, 2009; Chebotar et al., 2015). In our study, the core microbiome  
566 OTUs representing *Burkholderiales*, especially *Comamonadaceae* and  
567 *Oxalobacteraceae*, were relatively more abundant in *O. digyna*. We have repeatedly  
568 isolated bacteria from *O. digyna* vegetative tissues with very high sequence  
569 homologies to the above core OTUs (Nissinen et al., 2012; unpublished). Further, we  
570 have isolated or detected (in clone libraries) bacteria in *O. digyna* seeds with 100%  
571 (16S rRNA gene based) identity to six of the core OTUs (OTUs 2 and 16 representing  
572 *Rhizobiales*, OTUs 8, 13 and 35 (*Burkholderiales*) and OTU 15 (*Actinomycetales*)  
573 (unpublished data). Core OTUs related to similar strains from seeds were highly  
574 enriched (Figure 7) in the endosphere or rhizosphere soils. Part of these core  
575 organisms could thus be seed-transmitted and colonize the rhizo- and endosphere of  
576 developing seedlings, as previously described by Puente et al. (2009) in desert cacti.  
577 This indicates the potential importance of such seed-transmitted endophytes in  
578 pioneer plants. Horizontal transmission of a set of endophytes has also been observed  
579 by Hardoim et al. (2012) and Johnston-Monje and Raizada (2011).

580  
581 In addition, these core OTUs were among the primary drivers of region, compartment  
582 or host plant species differences among the bacterial communities. The higher relative  
583 abundances of *Clostridia* in Ny-Ålesund and *Rhizobia* in Mayrhofen in the  
584 endosphere communities is one such example, as discussed above.

585  
586 In summary, we here report that, on the basis of data obtained with two plant species,  
587 host plant-specific endophytic communities can be acquired despite a distance of over  
588 3000 km and differences in climate and chemistry between soils. These plant species-  
589 specific assemblages are formed from a shared core set of bacteria, most of which are  
590 strongly enriched in the endosphere. We surmised that plant-driven selection  
591 processes play a role, possibly concomitant with a highly efficient adaptation and  
592 fitness of these bacteria in the plant environment. Some of the core OTUs could even  
593 be seed-inherited, explaining their tight association with the host plant. Very closely-  
594 related endophytic taxa have previously been found to be shared by plants from other  
595 cold climates (Carrell and Frank, 2015; Nissinen et al., 2012; Poosakkannu et al.,  
596 2015), indicating the ecological tightness of [efficient] establishment of specific  
597 bacteria in arcto-alpine plants.

598  
599 **Conflict of interest**



600 The authors declare that the research was conducted in the absence of any commercial  
601 or financial relationships that could be construed as a potential conflict of interest.

602

### 603 **Author contributions**

604 Study was conceptualized and designed by RN and MK. Field work was performed  
605 by MK, RN and GB. Sample processing was done by MK and RN. Supporting soil  
606 analysis was done by MK while library preparation for sequence analysis was done by  
607 MK with assistance of AM. Bioinformatics analysis was performed by MK and the  
608 data analysis was done by MK and RN. Manuscript draft was prepared by MK, RN  
609 and JE and revisions was done by MK, RN, AM, GB, AS and JE. Final version for  
610 the submission was prepared by MK and RN.

611

### 612 **Nucleotide sequence data**

613 Nucleotide sequence data has been submitted to the ENA database and with  
614 accession number PRJEB17695.

615

### 616 **Funding**

617 This research was funded by NWO project 821.01.005 (for JE), Finnish cultural  
618 foundation Lapland regional fund (for RN) and by Finnish Academy, project 259180  
619 (for RN)

620

621

### 622 **Acknowledgements**

623 Authors acknowledge Maarten Loonen and other members of Netherlands Arctic  
624 Station, Ny-Ålesund, Svalbard and Kilpisjärvi Biological Station of the University of  
625 Helsinki for assistance in sampling. We also thank Andrés López-Sépulcre and Jenni  
626 Niku for their valuable suggestions with statistical analysis and Jolanda Brons for her  
627 assistance in soil physico-chemical analysis.

628

629

### 630 **Supplementary material**

631 S1 Site coordinates

632 S2 Phylum, class, order and family level analyses of the data.

633

### 634 **References**

635 Benítez, M. S., and McSpadden Gardener, B. B. (2009). Linking sequence to function  
636 in soil bacteria: Sequence-directed isolation of novel bacteria contributing to  
637 soilborne plant disease suppression. *Applied and Environmental Microbiology*  
638 75, 915–924. doi:10.1128/AEM.01296-08.

639 Berg, G., and Smalla, K. (2009). Plant species and soil type cooperatively shape the  
640 structure and function of microbial communities in the rhizosphere. *FEMS*  
641 *microbiology ecology* 68, 1–13. doi:10.1111/j.1574-6941.2009.00654.x.

642 Billings, W., and Mooney, H. (1968). The ecology of arctic and alpine plants.  
643 *Biological Reviews* 43, 481–529.

644 Borin, S., Ventura, S., Tambone, F., Mapelli, F., Schubotz, F., Brusetti, L., et al.  
645 (2010). Rock weathering creates oases of life in a high Arctic desert.  
646 *Environmental Microbiology* 12, 293–303. doi:10.1111/j.1462-  
647 2920.2009.02059.x.

648 Bray, R. H., and Kurtz, L. T. (1945). Determination of Total, Organic, and Available  
649 Forms of Phosphorus in Soils. *Soil Science* 59, 39–46. doi:10.1097/00010694-

194501000-00006.

651 Bulgarelli, D., Schlaeppli, K., Spaepen, S., Ver Loren van Themaat, E., and Schulze-  
652 Lefert, P. (2013). Structure and Functions of the Bacterial Microbiota of Plants.  
653 *Annual review of plant biology*. doi:10.1146/annurev-arplant-050312-120106.

654 Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D.,  
655 Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput  
656 community sequencing data. *Nature methods* 7, 335–6. doi:10.1038/nmeth.f.303.

657 Carrell, A. A., and Frank, A. C. (2015). Bacterial endophyte communities in the  
658 foliage of coast redwood and giant sequoia. *Frontiers in Microbiology* 6, 1008.  
659 doi:10.3389/fmicb.2015.01008.

660 Chapin, S. F., and Körner, C. (1996). “Arctic and Alpine Biodiversity: Its Patterns,  
661 Causes and Ecosystem consequences,” in *Functional Roles of Biodiversity: A*  
662 *Global perspective*, 7–32.

663 Chebotar, V. K., Malfanova, N. V., Shcherbakov, A. V., Ahtemova, G. A., Borisov,  
664 A. Y., Lugtenberg, B., et al. (2015). Endophytic bacteria in microbial  
665 preparations that improve plant development. *Applied Biochemistry and*  
666 *Microbiology* 51, 271–277. doi:10.1134/S0003683815030059.

667 Chelius, M. K., and Triplett, E. W. (2001). The Diversity of Archaea and Bacteria in  
668 Association with the Roots of Zea mays L. *Microbial ecology* 41, 252–263.  
669 doi:10.1007/s002480000087.

670 Cho, J. (2000). Biogeography and Degree of Endemicity of Fluorescent Pseudomonas  
671 Strains in Soil. 66, 5448–5456. doi:10.1128/AEM.66.12.5448-5456.2000.

672 Chu, H., Fierer, N., Lauber, C. L., Caporaso, J. G., Knight, R., and Grogan, P. (2010).  
673 Soil bacterial diversity in the Arctic is not fundamentally different from that  
674 found in other biomes. *Environmental microbiology* 12, 2998–3006.  
675 doi:10.1111/j.1462-2920.2010.02277.x.

676 Chu, H., Neufeld, J. D., Walker, V. K., and Grogan, P. (2011). The Influence of  
677 Vegetation Type on the Dominant Soil Bacteria, Archaea, and Fungi in a Low  
678 Arctic Tundra Landscape. *Soil Science Society of America Journal* 75, 1756.  
679 doi:10.2136/sssaj2011.0057.

680 Coleman-Derr, D., Desgarenes, D., Fonseca-Garcia, C., Gross, S., Clingenpeel, S.,  
681 Woyke, T., et al. (2016). Plant compartment and biogeography affect  
682 microbiome composition in cultivated and native *Agave* species. *New*  
683 *Phytologist* 209, 798–811. doi:10.1111/nph.13697.

684 Compant, S., Clément, C., and Sessitsch, A. (2010). Plant growth-promoting bacteria  
685 in the rhizo- and endosphere of plants: Their role, colonization, mechanisms  
686 involved and prospects for utilization. *Soil Biology and Biochemistry* 42, 669–  
687 678. doi:10.1016/j.soilbio.2009.11.024.

688 Coulson, S., and Hodkinson, I. (1995). Thermal environments of Arctic soil  
689 organisms during winter. *Arctic and Alpine Research* 27, 364–370.  
690 doi:10.2307/1552029.

691 deCárcer, D. A., Denman, S. E., McSweeney, C., and Morrison, M. (2011).  
692 Evaluation of subsampling-based normalization strategies for tagged high-  
693 throughput sequencing data sets from gut microbiomes. *Applied and*  
694 *Environmental Microbiology* 77, 8795–8798. doi:10.1128/AEM.05491-11.

695 DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al.  
696 (2006). Greengenes, a chimera-checked 16S rRNA gene database and  
697 workbench compatible with ARB. *Applied and environmental microbiology* 72,  
698 5069–72. doi:10.1128/AEM.03006-05.

699 Deslippe, J. R., and Egger, K. N. (2006). Molecular diversity of nifH genes from

700 bacteria associated with high arctic dwarf shrubs. *Microbial Ecology* 51, 516–  
701 525. doi:10.1007/s00248-006-9070-8.

702 Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST.  
703 *Bioinformatics* 26, 2460–2461. doi:10.1093/bioinformatics/btq461.

704 Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial  
705 amplicon reads. *Nature Methods* 10, 996–998. doi:10.1038/nmeth.2604.

706 Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N. K.,  
707 Bhatnagar, S., et al. (2015). Structure, variation, and assembly of the root-  
708 associated microbiomes of rice. *Proceedings of the National Academy of*  
709 *Sciences of the United States of America* 112, E911–20.  
710 doi:10.1073/pnas.1414592112.

711 Fierer, N., and Jackson, R. B. (2006). The diversity and biogeography of soil bacterial  
712 communities. *Proceedings of the National Academy of Sciences of the United*  
713 *States of America* 103, 626–31. doi:10.1073/pnas.0507535103.

714 Fierer, N., and Lennon, J. T. (2011). The generation and maintenance of diversity in  
715 microbial communities. *American journal of botany* 98, 439–48.  
716 doi:10.3732/ajb.1000498.

717 Garbeva, P., van Veen, J. A., and van Elsas, J. D. (2004). MICROBIAL DIVERSITY  
718 IN SOIL: Selection of Microbial Populations by Plant and Soil Type and  
719 Implications for Disease Suppressiveness. *Annual Review of Phytopathology* 42,  
720 243–270. doi:10.1146/annurev.phyto.42.012604.135455.

721 Ghyselinck, J., Pfeiffer, S., Heylen, K., Sessitsch, A., De Vos, P., Larsen, P., et al.  
722 (2013). The Effect of Primer Choice and Short Read Sequences on the Outcome  
723 of 16S rRNA Gene Based Diversity Studies. *PLoS ONE* 8, e71360.  
724 doi:10.1371/journal.pone.0071360.

725 Griffiths, R. I., Thomson, B. C., James, P., Bell, T., Bailey, M., and Whiteley, A. S.  
726 (2011). The bacterial biogeography of British soils. *Environmental microbiology*  
727 13, 1642–54. doi:10.1111/j.1462-2920.2011.02480.x.

728 Hanson, C. a., Fuhrman, J. a., Horner-Devine, M. C., and Martiny, J. B. H. (2012).  
729 Beyond biogeographic patterns: processes shaping the microbial landscape.  
730 *Nature Reviews Microbiology* 10, 1–10. doi:10.1038/nrmicro2795.

731 Hardoim, P. R., Hardoim, C. C. P., van Overbeek, L. S., and van Elsas, J. D. (2012).  
732 Dynamics of seed-borne rice endophytes on early plant growth stages. *PloS one*  
733 7, e30438. doi:10.1371/journal.pone.0030438.

734 Hardoim, P. R., van Overbeek, L. S., Berg, G., Pirttilä, A. M., Compant, S.,  
735 Campisano, A., et al. (2015). The Hidden World within Plants: Ecological and  
736 Evolutionary Considerations for Defining Functioning of Microbial Endophytes.  
737 *Microbiology and Molecular Biology Reviews* 79, 293–320.  
738 doi:10.1128/MMBR.00050-14.

739 Inceoğlu, Ö., Al-Soud, W. A., Salles, J. F., Semenov, A. V., and van Elsas, J. D.  
740 (2011). Comparative Analysis of Bacterial Communities in a Potato Field as  
741 Determined by Pyrosequencing. *PLoS ONE* 6, e23321.  
742 doi:10.1371/journal.pone.0023321.

743 Ji, M., van Dorst, J., Bissett, A., Brown, M. V., Palmer, A. S., Snape, I., et al. (2015).  
744 Microbial diversity at Mitchell Peninsula, Eastern Antarctica: a potential  
745 biodiversity “hotspot.” *Polar Biology* 39, 237–249. doi:10.1007/s00300-015-  
746 1776-y.

747 Johnston-Monje, D., and Raizada, M. N. (2011). Conservation and Diversity of Seed  
748 Associated Endophytes in *Zea* across Boundaries of Evolution, Ethnography and  
749 Ecology. *PLoS ONE* 6, e20396. doi:10.1371/journal.pone.0020396.

750 Körner, C. (2003). *Alpine Plant Life - Functional Plant Ecology of High Mountain*  
751 *Ecosystems*. Second Edi. Springer.

752 Kowalchuk, G. A., Buma, D. S., de Boer, W., Klinkhamer, P. G. L., and van Veen, J.  
753 A. (2002). Effects of above-ground plant species composition and diversity on  
754 the diversity of soil-borne microorganisms. *International Journal of General and*  
755 *Molecular Microbiology* 81, 509–520. doi:10.1023/A:1020565523615.

756 Kumar, M., Männistö, M. K., van Elsas, J. D., and Nissinen, R. M. (2016). Plants  
757 impact structure and function of bacterial communities in Arctic soils. *Plant and*  
758 *Soil* 399, 319–332. doi:10.1007/s11104-015-2702-3.

759 Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. (2009). Pyrosequencing-based  
760 assessment of soil pH as a predictor of soil bacterial community structure at the  
761 continental scale. *Applied and Environmental Microbiology* 75, 5111–20.  
762 doi:10.1128/AEM.00335-09.

763 Mäki, A., Rissanen, J. A., and Tirola, M. (2016). A practical method for barcoding  
764 and size-trimming PCR templates for amplicon sequencing. *Biotechniques* 60,  
765 88–90. doi:10.2144/000114380.

766 Männistö, M. K., Kurhela, E., Tirola, M., and Häggblom, M. M. (2013).  
767 Acidobacteria dominate the active bacterial communities of Arctic tundra with  
768 widely divergent winter-time snow accumulation and soil temperatures. *FEMS*  
769 *microbiology ecology* 84, 47–59. doi:10.1111/1574-6941.12035.

770 Männistö, M. K., Tirola, M., and Häggblom, M. M. (2007). Bacterial communities in  
771 Arctic fields of Finnish Lapland are stable but highly pH-dependent. *FEMS*  
772 *Microbiology Ecology* 59, 452–465. doi:10.1111/j.1574-6941.2006.00232.x.

773 Mapelli, F., Marasco, R., Rizzi, A., Baldi, F., Ventura, S., Daffonchio, D., et al.  
774 (2011). Bacterial communities involved in soil formation and plant establishment  
775 triggered by pyrite bioweathering on arctic moraines. *Microbial ecology* 61,  
776 438–47. doi:10.1007/s00248-010-9758-7.

777 Margesin, R., Jud, M., Tscherko, D., and Schinner, F. (2009). Microbial communities  
778 and activities in alpine and subalpine soils. *FEMS Microbiology Ecology* 67,  
779 208–218. doi:10.1111/j.1574-6941.2008.00620.x.

780 Meyer, D., Zeileis, A., and Hornik, K. (2015). vcd: Visualizing Categorical Data. R  
781 package version. 1.4–1. Available at: [https://cran.r-](https://cran.r-project.org/web/packages/vcd/citation.html)  
782 [project.org/web/packages/vcd/citation.html](https://cran.r-project.org/web/packages/vcd/citation.html).

783 Minamisawa, K., Nishioka, K., Miyaki, T., Ye, B., Miyamoto, T., You, M., et al.  
784 (2004). Anaerobic nitrogen-fixing consortia consisting of clostridia isolated from  
785 gramineous plants. *Applied and environmental microbiology* 70, 3096–102.  
786 doi:10.1128/AEM.70.5.3096-3102.2004.

787 Miniaci, C., Bunge, M., Duc, L., Edwards, I., Bürgmann, H., and Zeyer, J. (2007).  
788 Effects of pioneering plants on microbial structures and functions in a glacier  
789 forefield. *Biology and Fertility of Soils* 44, 289–297. doi:10.1007/s00374-007-  
790 0203-0.

791 Neufeld, J., and Mohn, W. (2005). Unexpectedly high bacterial diversity in arctic  
792 tundra relative to boreal forest soils, revealed by serial analysis of ribosomal  
793 sequence tags. *Applied and environmental microbiology* 71, 5710–5718.  
794 doi:10.1128/AEM.71.10.5710.

795 Nissinen, R. M., Männistö, M. K., and van Elsas, J. D. (2012). Endophytic bacterial  
796 communities in three arctic plants from low arctic fell tundra are cold-adapted  
797 and host-plant specific. *FEMS microbiology ecology* 82, 510–22.  
798 doi:10.1111/j.1574-6941.2012.01464.x.

799 Oakley, B. B., Carbonero, F., van der Gast, C. J., Hawkins, R. J., and Purdy, K. J.

800 (2010). Evolutionary divergence and biogeography of sympatric niche-  
801 differentiated bacterial populations. *The ISME journal* 4, 488–97.  
802 doi:10.1038/ismej.2009.146.

803 Poosakkannu, A., Nissinen, R., and Kytöviita, M. M. (2015). Culturable endophytic  
804 microbial communities in the circumpolar grass, *Deschampsia flexuosa* in a sub-  
805 Arctic inland primary succession are habitat and growth stage specific.  
806 *Environmental Microbiology Reports* 7, 111–122. doi:10.1111/1758-  
807 2229.12195.

808 Puente, M. E., Li, C. Y., and Bashan, Y. (2009). Rock-degrading endophytic bacteria  
809 in cacti. *Environmental and Experimental Botany* 66, 389–401.  
810 doi:10.1016/j.envexpbot.2009.04.010.

811 Pylro, V. S., Roesch, L. F. W., Morais, D. K., Clark, I. M., Hirsch, P. R., and Tótolá,  
812 M. R. (2014). Data Analysis for 16S Microbial Profiling from Different  
813 Benchtop Sequencing Platforms. *Journal of microbiological methods* 107, 30–  
814 37. doi:10.1016/j.mimet.2014.08.018.

815 Rosenblueth, M., and Martínez-Romero, E. (2006). Bacterial endophytes and their  
816 interactions with hosts. *Molecular plant-microbe interactions : MPMI* 19, 827–  
817 37. doi:10.1094/MPMI-19-0827.

818 Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., et  
819 al. (2010). Soil bacterial and fungal communities across a pH gradient in an  
820 arable soil. *The ISME journal* 4, 1340–51. doi:10.1038/ismej.2010.58.

821 Santoyo, G., Moreno-Hagelsieb, G., del Carmen Orozco-Mosqueda, M., and Glick, B.  
822 R. (2016). Plant growth-promoting bacterial endophytes. *Microbiological*  
823 *Research* 183, 92–99. doi:10.1016/j.micres.2015.11.008.

824 Schlaeppli, K., Dombrowski, N., Oter, R. G., Ver Loren van Themaat, E., and  
825 Schulze-Lefert, P. (2014). Quantitative divergence of the bacterial root  
826 microbiota in *Arabidopsis thaliana* relatives. *Proceedings of the National*  
827 *Academy of Sciences of the United States of America* 111, 585–92.  
828 doi:10.1073/pnas.1321597111.

829 Shabuer, G., Ishida, K., Pidot, S. J., Roth, M., Dahse, H.-M., and Hertweck, C.  
830 (2015). Plant pathogenic anaerobic bacteria use aromatic polyketides to access  
831 aerobic territory. *Science* 350.

832 Shi, Y., Xiang, X., Shen, C., Chu, H., Neufeld, J. D., Walker, V. K., et al. (2015).  
833 Vegetation-Associated Impacts on Arctic Tundra Bacterial and Microeukaryotic  
834 Communities. *Applied and Environmental Microbiology* 81, 492–501.  
835 doi:10.1128/AEM.03229-14.

836 Siciliano, S. D., Palmer, A. S., Winsley, T., Lamb, E., Bissett, A., Brown, M. V., et al.  
837 (2014). Soil fertility is associated with fungal and bacterial richness, whereas pH  
838 is associated with community composition in polar soil microbial communities.  
839 *Soil Biology and Biochemistry* 78, 10–20. doi:10.1016/j.soilbio.2014.07.005.

840 Smalla, K., Wieland, G., Buchner, A., Zock, A., Parzy, J., Roskot, N., et al. (2001).  
841 Bulk and Rhizosphere Soil Bacterial Communities Studied by Denaturing  
842 Gradient Gel Electrophoresis : Plant-Dependent Enrichment and Seasonal Shifts  
843 Revealed. *Applied and environmental microbiology* 67, 4742–4751.  
844 doi:10.1128/AEM.67.10.4742.

845 Soltis, D. E., Soltis, P. S., Chase, M. W., Mort, M. E., Albach, D. C., Zanis, M., et al.  
846 (2000). Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB  
847 sequences. *Botanical Journal of the Linnean Society* 133, 381–461.  
848 doi:10.1006/boj.2000.0380.

849 Wikström, N., Savolainen, V., and Chase, M. W. (2001). Evolution of the

850 angiosperms: calibrating the family tree. *Proceedings. Biological sciences / The*  
851 *Royal Society* 268, 2211–20. doi:10.1098/rspb.2001.1782.  
852 Yergeau, E., Newsham, K. K., Pearce, D. A., and Kowalchuk, G. A. (2007). Patterns  
853 of bacterial diversity across a range of Antarctic terrestrial habitats.  
854 *Environmental microbiology* 9, 2670–82. doi:10.1111/j.1462-  
855 2920.2007.01379.x.  
856 Zarraonaindia, I., Owens, S. M., Weisenhorn, P., West, K., Hampton-Marcell, J., Lax,  
857 S., et al. (2015). The soil microbiome influences grapevine-associated  
858 microbiota. *mBio* 6, 1–10. doi:10.1128/mBio.02527-14.  
859 Zhang, H. W., Song, Y. C., and Tan, R. X. (2006). Biology and chemistry of  
860 endophytes. *Natural product reports* 23, 753–71. doi:10.1039/b609472b.  
861 Zhao, S., Zhou, N., Zhao, Z.-Y., Zhang, K., and Tian, C.-Y. (2016). High-Throughput  
862 Sequencing Analysis of the Endophytic Bacterial Diversity and Dynamics in  
863 Roots of the Halophyte *Salicornia europaea*. *Current Microbiology* 72, 557–562.  
864 doi:10.1007/s00284-016-0990-3.  
865 Zheng, D., Alm, E. W., Stahl, D. A., and Raskin, L. (1996). Characterization of  
866 universal small-subunit rRNA hybridization probes for quantitative molecular  
867 microbial ecology studies. *Applied and environmental microbiology* 62, 4504–  
868 13.  
869 Zhou, J., Xia, B., Huang, H., Treves, D. S., Hauser, L. J., Mural, R. J., et al. (2003).  
870 Bacterial phylogenetic diversity and a novel candidate division of two humid  
871 region, sandy surface soils. *Soil Biology and Biochemistry* 35, 915–924.  
872 doi:10.1016/S0038-0717(03)00124-X.  
873  
874  
875  
876  
877

Provisional

878 **Figure legends**

879

880 Figure 1. Sampling sites and OTU distribution. (A) Map of Europe depicting our three  
881 sampling locations Mayrhofen from Austrian Alps, Kilpisjärvi from low-arctic  
882 Finnish Lapland and Ny-Ålesund from high-arctic Svalbard archipelago. (B) Venn  
883 diagrams of shared OTUs (number of reads of respective OTUs) across three regions  
884 and (C) compartments.

885

886 Figure 2 Estimated Shannon diversity (A) in bulk soil, rhizosphere soil and  
887 endophytic bacterial communities from three climatic regions Mayrhofen, Kilpisjärvi  
888 and Ny-Ålesund (B) Scatter plots of bulk soil communities explaining the correlation  
889 (Pearson correlation) between Shannon diversity with soil-physico chemical  
890 properties from three climatic regions.

891

892 Figure 3 Distribution of OTUs and phyla across regions (A) Ternary plot of OTU  
893 distribution across three climatic regions. Each circle represents one OTU, and the  
894 size, color and position of the circle represent its relative abundance, bacterial phylum  
895 and affiliation of the OTU with the different regions, respectively. (B-D) Average  
896 relative abundances of bacterial phyla distributed across different regions in (B) Bulk  
897 soil samples, (C) Rhizosphere soil samples, (D) Endosphere samples. Major phyla  
898 (average relative abundance above 1%) with significantly differential distribution (as  
899 detected by Kruskal-Wallis analysis) are marked with asterisks.

900

901 Figure 4 Distribution of OTUs and phyla across different compartments (A) Ternary  
902 plot of all OTUs plotted based on the compartment specificity. Each circle represents  
903 one OTU. The size, color and position of each OTU represents its relative abundance,  
904 bacterial phyla and contribution of the OTU to the nearby compartments respectively.  
905 (B) Distribution of average relative abundance of selected major bacterial phyla from  
906 all three regions across the compartments. Major phyla (average relative abundance  
907 above 1%) with significantly differential distribution (detected by Kruskal-Wallis  
908 analysis) are marked with asterisks.

909

910 Figure 5 Principal Coordinate Analysis (PCoA) plots of bacterial communities from  
911 bulk soils, rhizosphere soils and endospheres of *O. digyna* and *S. oppositifolia* from  
912 three climatic regions Mayrhofen, Kilpisjärvi and Ny-Ålesund. (A) All samples, (B)  
913 Bulk soils and rhizosphere soils, (C) Endospheres from *O. digyna* and *S. oppositifolia*.  
914 The symbol colors correspond to compartment (A and B) or plant species (C) and the  
915 shapes of the symbols correspond to the geographic regions. Compartment, region  
916 and plant species all had significant impact on community structures in global as well  
917 as in pair wise analyses (PERMANOVA  $P=0.001$ ). All ordinations are based on Bray-  
918 Curtis distance matrixes.

919

920 Figure 6 (A) Venn diagram of common shared OTUs and plant species specific OTUs  
921 (number of reads of the respective OTUs) between *O. digyna* and *S. oppositifolia*  
922 from all the endosphere samples. Average relative abundance of endophytic bacterial  
923 communities associated with *O. digyna* and *S. oppositifolia* endosphere samples at  
924 different taxonomical level. (B) bacterial class, (c) bacterial order. Only selected  
925 major bacterial orders and classes were classified and shown. Major bacterial classes  
926 or orders (with average relative abundance above 1% in endosphere) with

927 significantly different distribution (detected by Kruskal-Wallis analysis) are marked  
928 with asterisks.

929

930 Figure 7 Relative distribution of core OTUs' reads across different compartments.

931 The graph is based on average read count of each OTU in different compartments.

932

933

Provisional



934 Table 1. Soil physico-chemical properties

Region	Sampling Site []	SOM (%)	pH	Available Phosphorous mg/kg
Mayrhofen	Alps (A) [8]	0.01 (0.002)	7.03 (0.9)	1.84 (1.1)
	Cliff (C) [4]	0.02 (0.002)	4.60 (0.1)	1.48 (0.4)
	Average	0.01 (0.002)	5.81 (0.5)	1.66 (0.8)
Kilpisjärvi	Jehkas New (JN) [8]	0.02 (0.002)	5.55 (0.2)	1.31 (0.4)
	Jehkas Old (JO) [8]	0.02 (0.008)	6.36 (0.4)	0.76 (0.4)
	Saana (S) [8]	0.02 (0.01)	5.49 (0.6)	2.45 (1.5)
	Average	0.02 (0.01)	5.80 (0.5)	1.51 (0.8)
Ny- Ålesund	Knudsenheia (K) [8]	0.03 (0.01)	7.4 (0.9)	0.83 (0.5)
	Midtre Lovénbreen (M) [8]	0.03 (0.03)	6.4 (1.2)	0.63 (0.1)
	Red River (RR) [8]	0.04 (0.01)	7.78 (0.5)	0.34 (0.1)
	Average	0.04 (0.02)	7.20 (0.9)	0.60 (0.3)

935 ( ) – Standard deviation values

936 [] – number of biological replicates/sampling site

Provisional

937 Table 2. Contributions of variables to similarity (SIMPER) analysis based on Bray-  
 938 Curtis dissimilarity indexes at phylum level identifying the major phyla driving the  
 939 dissimilarities between different regions or compartments. Pairwise Permutational  
 940 multivariate analyses (PERMANOVA) were performed prior to SIMPER to test for  
 941 significant differences between the tested groups. Pseudo-F and p-values, or t and p-  
 942 values from PERMANOVA are given for each factor and group pair, respectively.

Bacterial phylum level - regions (Pseudo-F: 8.116 p:0.001)			
t: 2.087, p: 0.016_	Mayrhofen	Kilpisjärvi	
Phylum	Average Abundance	Average Abundance	Contribution to Dissimilarity %
<i>Proteobacteria</i>	56.93	46.14	27.77
<i>Acidobacteria</i>	7.68	13.74	17.35
<i>AD3</i>	4.48	10.04	15.66
<i>Actinobacteria</i>	13.64	13.88	13.36
<i>Firmicutes</i>	2.36	4.06	7.17
<i>Bacteroidetes</i>	5.82	4.09	5.72
<i>Gemmatimonadetes</i>	2.5	2.69	3.9
t: 2.578, p: 0.001	Mayrhofen	Ny-Ålesund	
Phylum	Average Abundance	Average Abundance	Contribution to Dissimilarity %
<i>Proteobacteria</i>	56.93	43.3	25.91
<i>Actinobacteria</i>	13.64	23.38	18.14
<i>Acidobacteria</i>	7.68	6.18	12.7
<i>Firmicutes</i>	2.36	5.48	9.62
<i>AD3</i>	4.48	2.23	8.19
<i>Bacteroidetes</i>	5.82	6.99	7.48
<i>Chloroflexi</i>	2.17	4.83	5.79
<i>Gemmatimonadetes</i>	2.5	3.68	5
t: 3.495, p:0.001	Kilpisjärvi	Ny-Ålesund	
Phylum	Average Abundance	Average Abundance	Contribution to Dissimilarity %
<i>Proteobacteria</i>	46.14	43.3	19.47
<i>Actinobacteria</i>	13.88	23.38	18.25
<i>Acidobacteria</i>	13.74	6.18	16.32
<i>AD3</i>	10.04	2.23	13.64
<i>Firmicutes</i>	4.06	5.48	10.04
<i>Bacteroidetes</i>	4.09	6.99	6.92
<i>Chloroflexi</i>	1.99	4.83	5.16
<i>Gemmatimonadetes</i>	2.69	3.68	4.52
Bacterial phylum level– compartment (Pseudo-F: 64.371, p: 0.001)			
t: 3.915, p: 0.001	Bulk Soil	Rhizosphere	
Phylum	Average Abundance	Average Abundance	Contribution to Dissimilarity %
<i>AD3</i>	13.59	2.89	20.44
<i>Proteobacteria</i>	36.92	47.47	19.58
<i>Actinobacteria</i>	17.7	19.71	18.99
<i>Acidobacteria</i>	13.84	11.77	18.92
<i>Gemmatimonadetes</i>	5.88	2.88	5.97
<i>Bacteroidetes</i>	3.33	5.47	4.78

<i>Chloroflexi</i>	4.39	4.38	4.46
t: 9.353, p: 0.001	Bulk Soil	Endosphere	
Phylum	Average Abundance	Average Abundance	Contribution to Dissimilarity %
<i>Proteobacteria</i>	36.92	58.06	24.91
<i>AD3</i>	13.59	0.55	14.37
<i>Acidobacteria</i>	13.84	2.16	13.47
<i>Actinobacteria</i>	17.7	15.45	13.32
<i>Firmicutes</i>	0.56	11.81	12.16
<i>Bacteroidetes</i>	3.33	8.16	6.69
<i>Gemmatimonadetes</i>	5.88	0.48	5.8
<i>Chloroflexi</i>	4.39	0.76	4.12
t: 8.569, p: 0.001	Rhizosphere	Endosphere	
Phylum	Average Abundance	Average Abundance	Contribution to Dissimilarity %
<i>Proteobacteria</i>	47.47	58.06	23.81
<i>Firmicutes</i>	0.29	11.81	16.5
<i>Acidobacteria</i>	11.77	2.16	15.09
<i>Actinobacteria</i>	19.71	15.45	14.71
<i>Bacteroidetes</i>	5.47	8.16	8.24
<i>Chloroflexi</i>	4.38	0.76	5.51
<i>AD3</i>	2.89	0.55	4.44
<i>Gemmatimonadetes</i>	2.88	0.48	3.52

943

944

Provisional

945 Table 3. Permutational multivariate analysis (PERMANOVA) of factors impacting  
 946 differences between community structures of bacteria at OTU level from different  
 947 climatic regions, sampling sites, compartments or plant species.

<b>PERMANOVA, Single Factor Analysis</b>					
Factor	df	SS	MS	Pseudo-F	p-value(perm)
<b>All Samples</b>					
Compartment	2	1.02E+05	51.132	30.96	0.001
Site	7	56.283	8.040	4.064	0.001
Region	2	27.369	13.685	6.549	0.001
<b>Bulk Soil</b>					
Site	7	43.653	6.236	7.671	0.001
Region	2	21.221	10.610	9.150	0.001
<b>Rhizosphere Soil</b>					
Site	7	28.835	4.119	5.073	0.001
Region	2	12.352	6.176	5.996	0.001
<b>Endosphere</b>					
Site	7	29.679	4.239	2.142	0.001
Region	2	11.842	5.921	2.788	0.001
<b>PERMANOVA, Two-Factor Analysis</b>					
<b>Rhizosphere Soil</b>					
Region	2	11.466	5.732	5.886	0.001
Plant species	1	2.910	2.910	2.988	0.007
Region X Plant species	2	2.983	1.491	1.531	0.067
Residuals	54	52.599	974		
<b>Endosphere</b>					
Region	2	11.659	5.829	2.968	0.001
Plant species	1	7.922	7.922	4.033	0.001
Region X Plant species	2	6.383	3.191	1.625	0.003
Residuals	52	1.02E+05	1.964		

948 Table 4. 20 key OTUs shaping the endosphere communities in *O. digyna* and *S. oppositifolia* in the three regions identified by SIMPER  
 949 (Contributions of variables to similarity analysis). Numerical values indicate % contribution of the respective OTUs in determining the  
 950 difference in community composition between endosphere and rhizosphere, between *O. digyna* and *S. oppositifolia* and between the three  
 951 regions. \* indicate the top 20 OTUs strongly contributing to the differences in community structures between the compartments, plant species  
 952 and geographic regions. OTUs which are also part of tightly associated OTUs were highlighted by **bold letters** in the OTU # column.

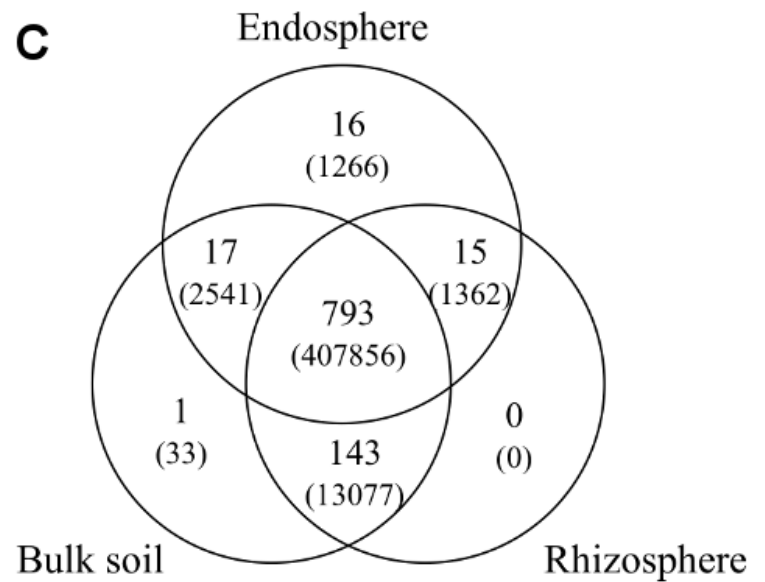
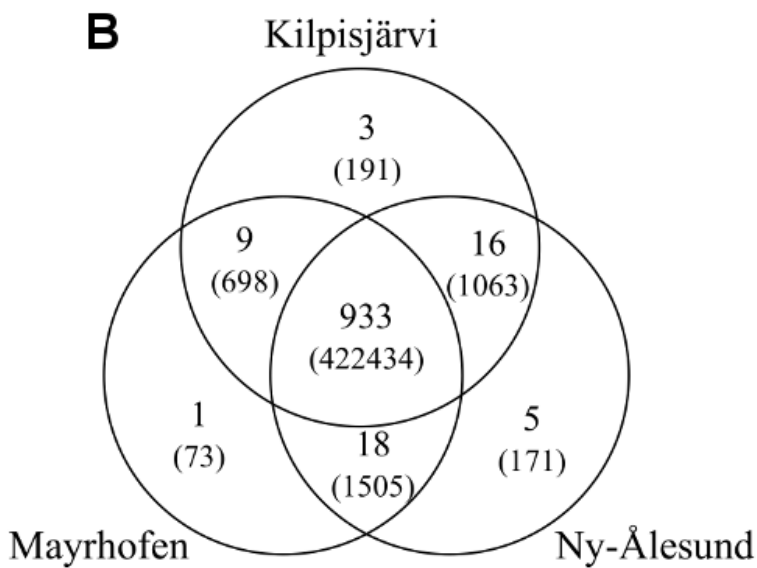
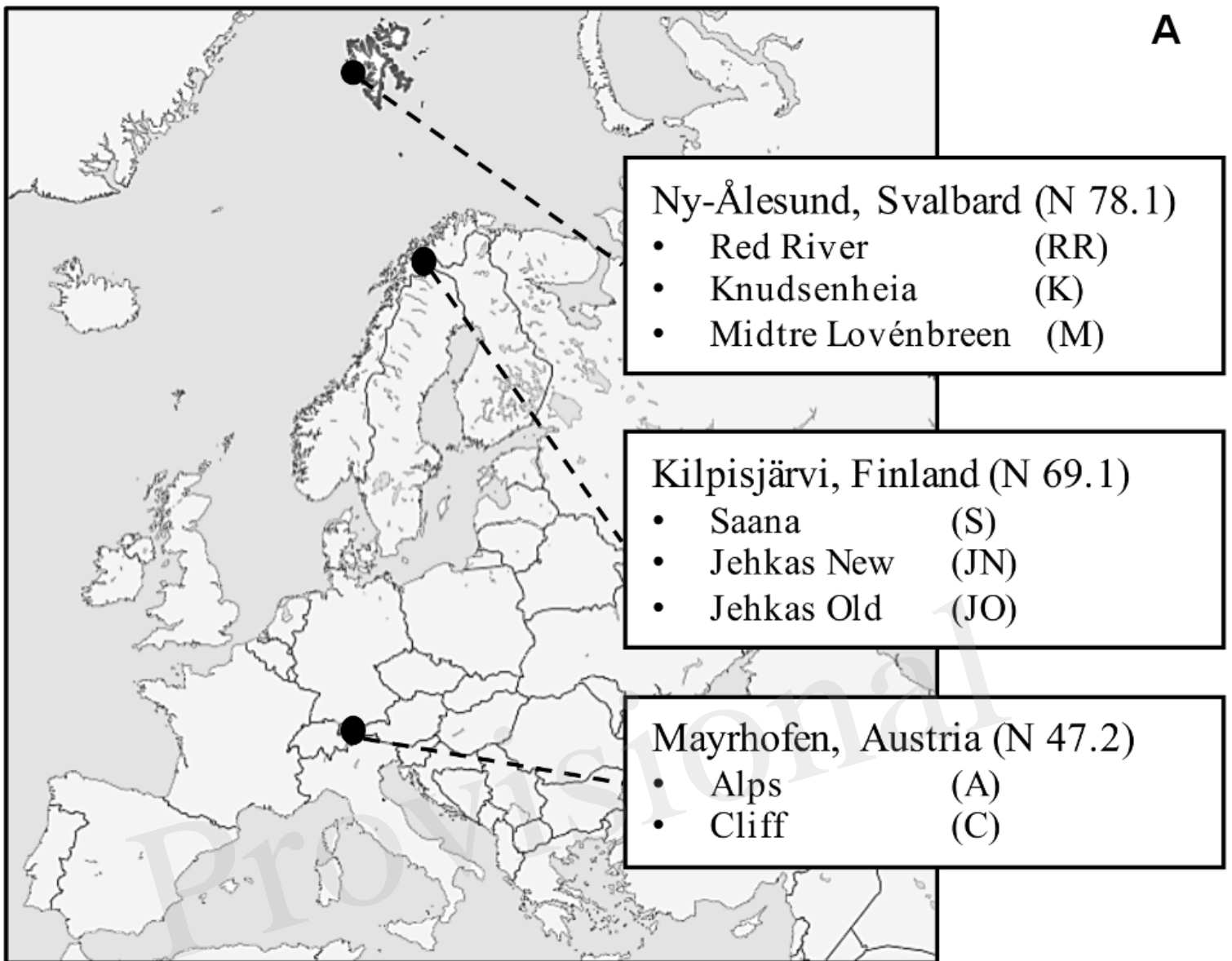
OTU #	Endo vs Rhizo	Plant species	Region	Phyla	Class	Order	Family	Genus	Species
OTU_21	0.97*	1.71*	1.69*	<i>Firmicutes</i>	<i>Clostridia</i>	<i>Clostridiales</i>	<i>Clostridiaceae</i>	<i>Clostridium</i>	
<b>OTU_5</b>	0.95*	1.01*	1.04*	<i>Firmicutes</i>	<i>Bacilli</i>	<i>Bacillales</i>	<i>Planococcaceae</i>		
<b>OTU_2</b>	0.93*	1.73*	1.87*	<i>Proteobacteria</i>	$\alpha$ -proteobacteria	<i>Rhizobiales</i>	<i>Bradyrhizobiaceae</i>	<i>Bradyrhizobium</i>	
OTU_3	0.81*	1.59*	1.68*	<i>Proteobacteria</i>	$\delta$ -proteobacteria	<i>Myxococcales</i>	<i>Haliangiaceae</i>		
<b>OTU_15</b>	0.73*	1.28*	1.20*	<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Actinomycetales</i>	<i>Micrococcaceae</i>	<i>Kocuria</i>	
<b>OTU_8</b>	0.71*	1.52*	1.79*	<i>Proteobacteria</i>	$\beta$ -proteobacteria	<i>Burkholderiales</i>	<i>Oxalobacteraceae</i>	<i>Janthinobacterium</i>	<i>lividum</i>
<b>OTU_33</b>	0.65*	0.99*	0.97*	<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Actinomycetales</i>	<i>Micromonosporaceae</i>		
<b>OTU_4</b>	0.62*	1.43*	1.24*	<i>Proteobacteria</i>	$\beta$ -proteobacteria	<i>Burkholderiales</i>	<i>Comamonadaceae</i>		
OTU_706	0.62*	1.43*	1.25*	<i>Bacteroidetes</i>	<i>Saprospirae</i>	<i>Saprospirales</i>	<i>Chitinophagaceae</i>		
<b>OTU_13</b>	0.53*	1.12*	1.24*	<i>Proteobacteria</i>	$\beta$ -proteobacteria	<i>Burkholderiales</i>	<i>Comamonadaceae</i>	<i>Methylibium</i>	
OTU_84	0.5*	0.99*	1.17*	<i>Proteobacteria</i>	$\beta$ -proteobacteria	<i>Burkholderiales</i>	<i>Comamonadaceae</i>	<i>Limnohabitans</i>	
OTU_48	0.4	0.69*	0.78*	<i>Proteobacteria</i>	$\gamma$ -proteobacteria	<i>Pseudomonadales</i>	<i>Pseudomonadaceae</i>	<i>Pseudomonas</i>	
<b>OTU_16</b>	0.37	0.65*	0.63	<i>Proteobacteria</i>	$\alpha$ -proteobacteria	<i>Rhizobiales</i>	<i>Hyphomicrobiaceae</i>	<i>Rhodoplanes</i>	
OTU_22	0.37	0.74*	0.79*	<i>Proteobacteria</i>	$\beta$ -proteobacteria	<i>Burkholderiales</i>	<i>Comamonadaceae</i>		
OTU_36	0.37	0.92*	0.72*	<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Actinomycetales</i>			
OTU_23	0.29	0.71*	0.59	<i>Proteobacteria</i>	$\gamma$ -proteobacteria	<i>Xanthomonadales</i>	<i>Sinobacteraceae</i>	<i>Steroidobacter</i>	
OTU_26	0.28	0.64*	0.62	<i>Proteobacteria</i>	<i>Un_Proteobacteria</i>				
OTU_41	0.27	0.56*	0.51	<i>Proteobacteria</i>	$\gamma$ -proteobacteria	<i>Xanthomonadales</i>	<i>Sinobacteraceae</i>		
OTU_37	0.26	0.71*	0.54	<i>Proteobacteria</i>	$\alpha$ -proteobacteria	<i>Rhizobiales</i>	<i>Hyphomicrobiaceae</i>		
OTU_83	0.26	0.59*	0.47	<i>Actinobacteria</i>	<i>Acidimicrobiia</i>	<i>Acidimicrobiales</i>			

953 Table 5. Highly conserved core OTUs of *O. digyna* and *S. oppositifolia* endospheres. OTUs present in minimum of three endosphere samples out  
 954 of four in all sampling sites in all regions per plant species are included.

OTU #	Phyla	Class	Order	Family	Genus	Species
Core OTUs of both plant species						
OTU_16	<i>Proteobacteria</i>	<i>α-proteobacteria</i>	<i>Rhizobiales</i>	<i>Hyphomicrobiaceae</i>	<i>Rhodoplanes</i>	
OTU_2	<i>Proteobacteria</i>	<i>α-proteobacteria</i>	<i>Rhizobiales</i>	<i>Bradyrhizobiaceae</i>	<i>Bradyrhizobium</i>	
OTU_429	<i>Proteobacteria</i>	<i>α-proteobacteria</i>	<i>Rhizobiales</i>	<i>Bradyrhizobiaceae</i>		
OTU_5	<i>Firmicutes</i>	<i>Bacilli</i>	<i>Bacillales</i>	<i>Planococcaceae</i>		
OTU_8	<i>Proteobacteria</i>	<i>β-proteobacteria</i>	<i>Burkholderiales</i>	<i>Oxalobacteraceae</i>	<i>Janthinobacterium</i>	<i>lividum</i>
Additional core OTUs of <i>S. oppositifolia</i>						
OTU_15	<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Actinomycetales</i>	<i>Micrococcaceae</i>	<i>Kocuria</i>	
OTU_171	<i>Proteobacteria</i>	<i>α-proteobacteria</i>	<i>Rhizobiales</i>	<i>Bradyrhizobiaceae</i>		
OTU_33	<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Actinomycetales</i>	<i>Micromonosporaceae</i>		
OTU_7	<i>Proteobacteria</i>	<i>β-proteobacteria</i>	<i>Ellin6067</i>	<i>Un_Ellin6067</i>		
Additional core OTUs of <i>O. digyna</i>						
OTU_13	<i>Proteobacteria</i>	<i>β-proteobacteria</i>	<i>Burkholderiales</i>	<i>Comamonadaceae</i>	<i>Methylibium</i>	
OTU_32	<i>Proteobacteria</i>	<i>γ-proteobacteria</i>	<i>Enterobacteriales</i>	<i>Enterobacteriaceae</i>		
OTU_35	<i>Proteobacteria</i>	<i>β-proteobacteria</i>	<i>Burkholderiales</i>	<i>Comamonadaceae</i>		
OTU_4	<i>Proteobacteria</i>	<i>β-proteobacteria</i>	<i>Burkholderiales</i>	<i>Comamonadaceae</i>		

955

Figure 01.TIF



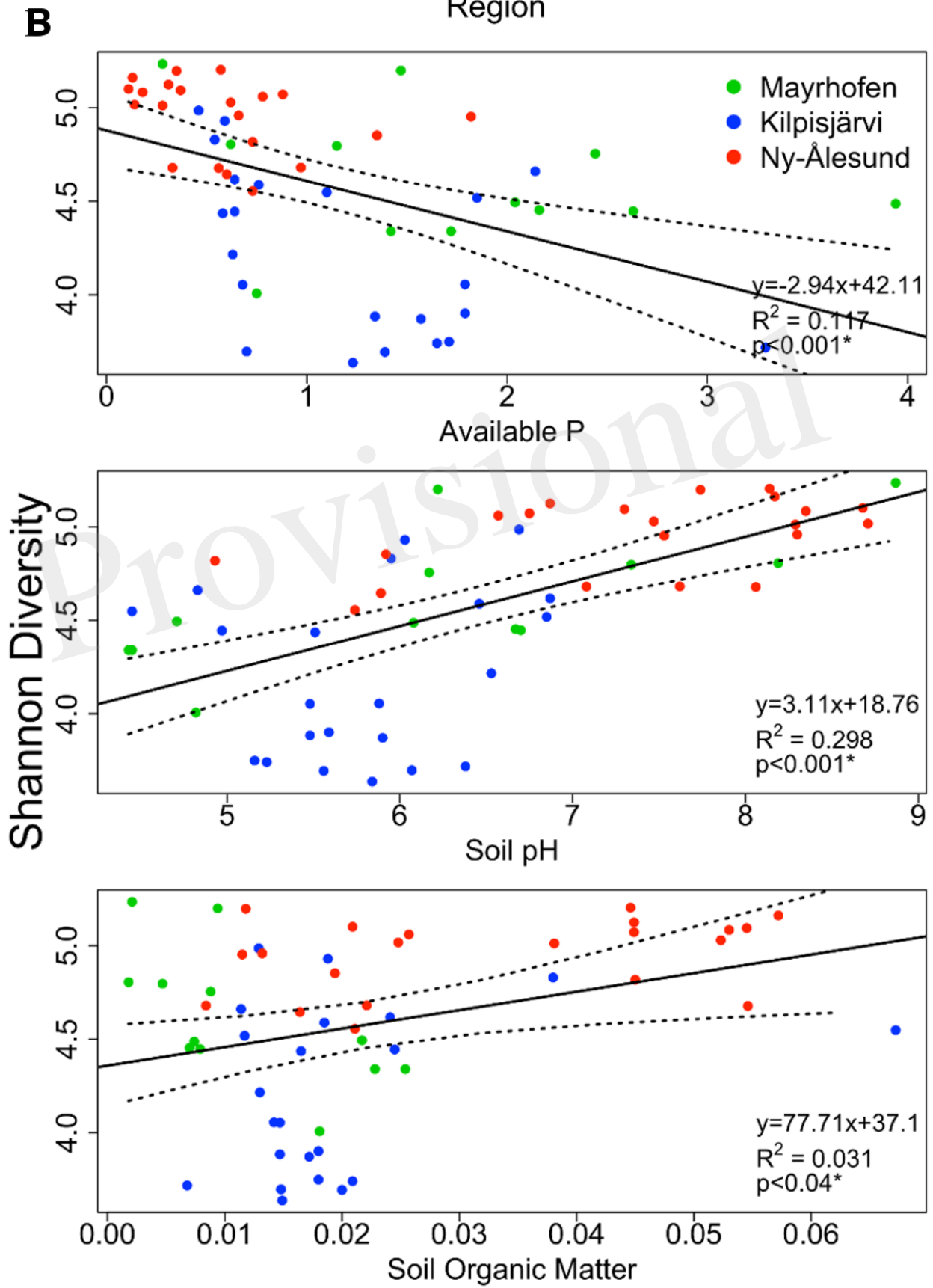
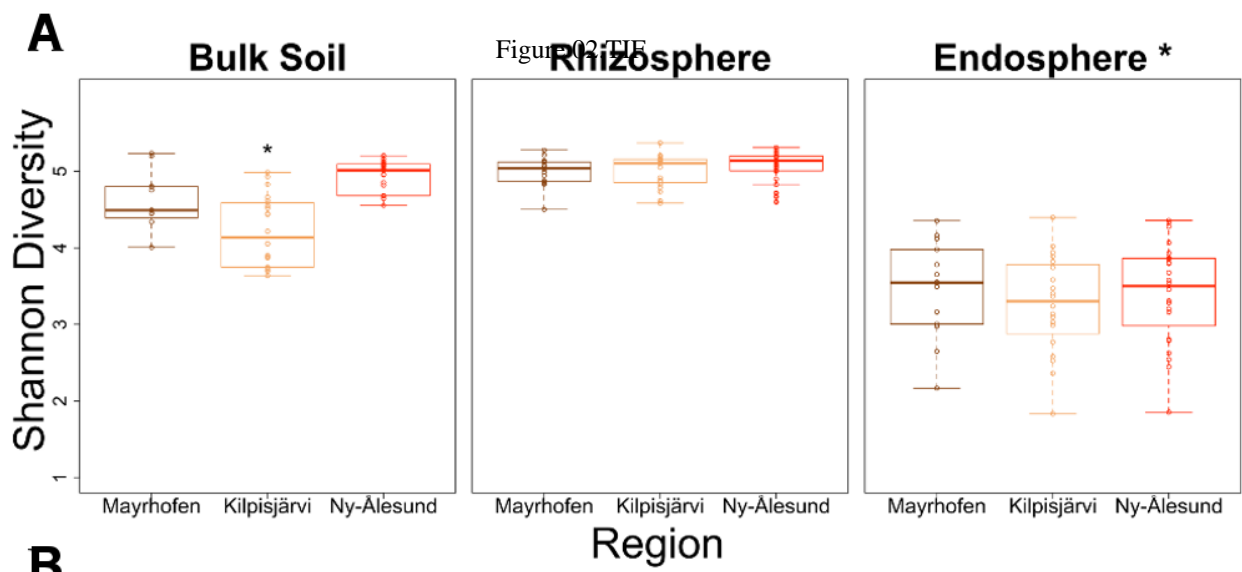




Figure 03.TIFF

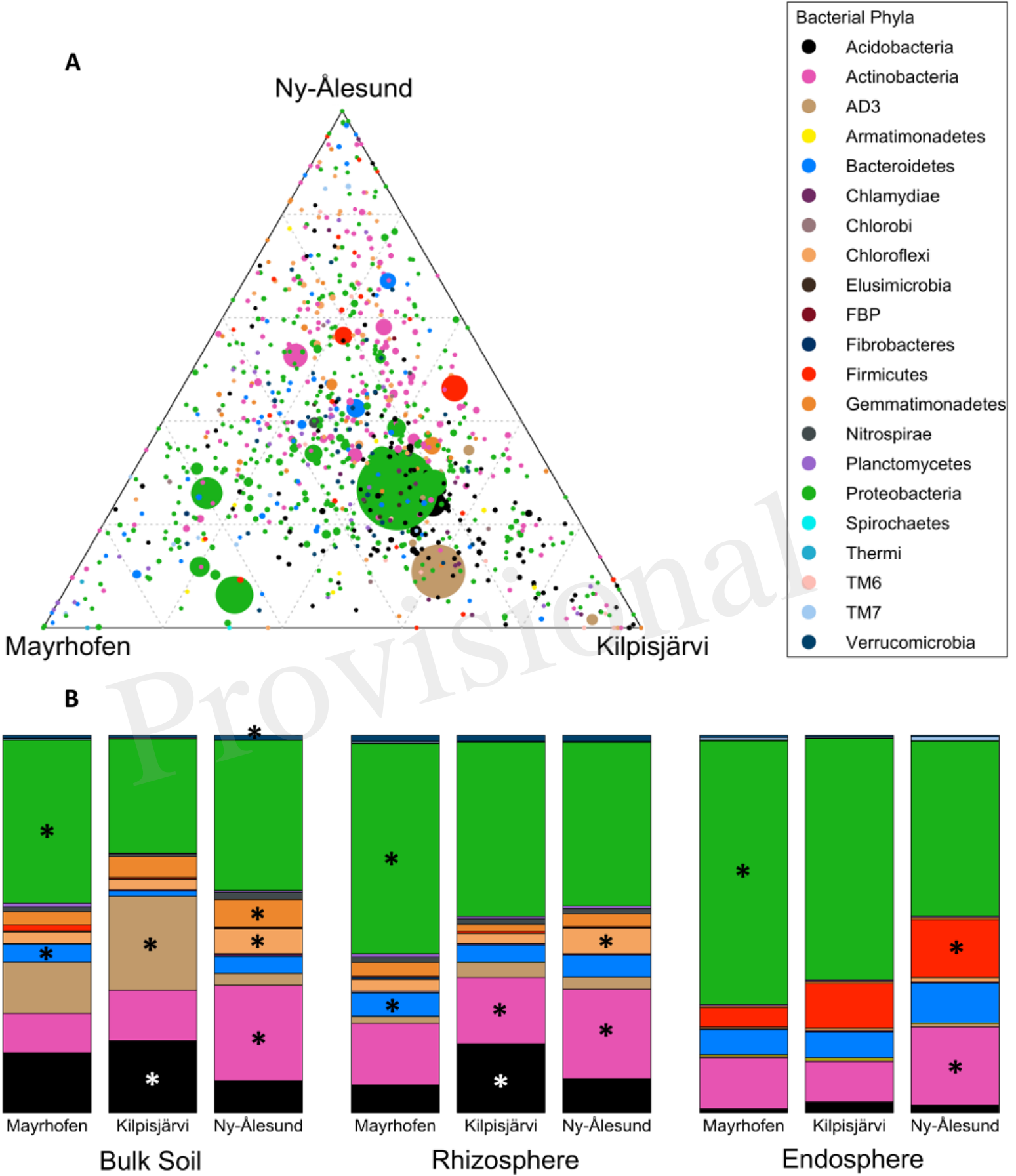
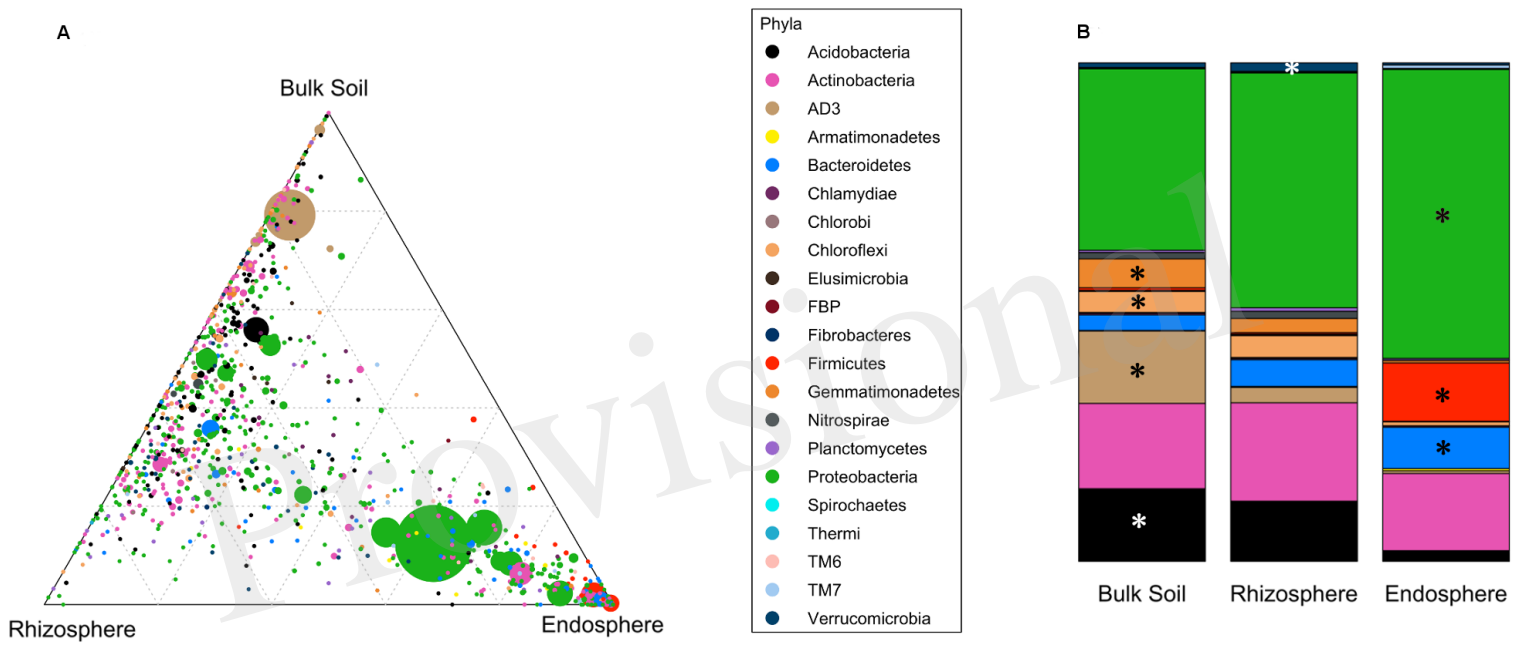


Figure 04.TIFF



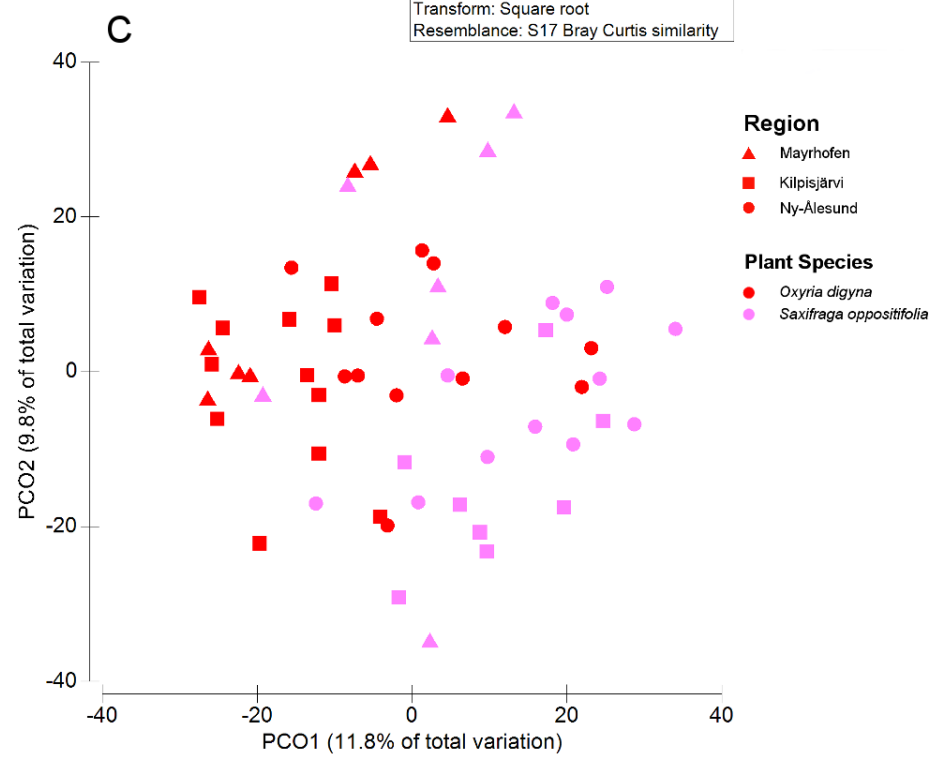
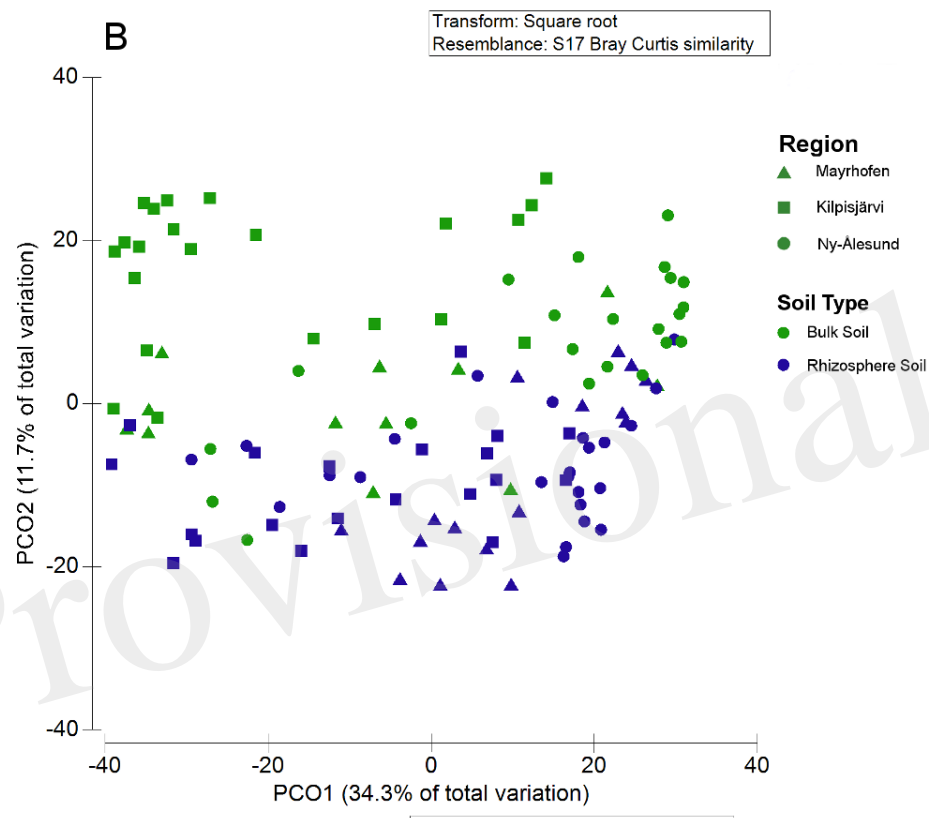
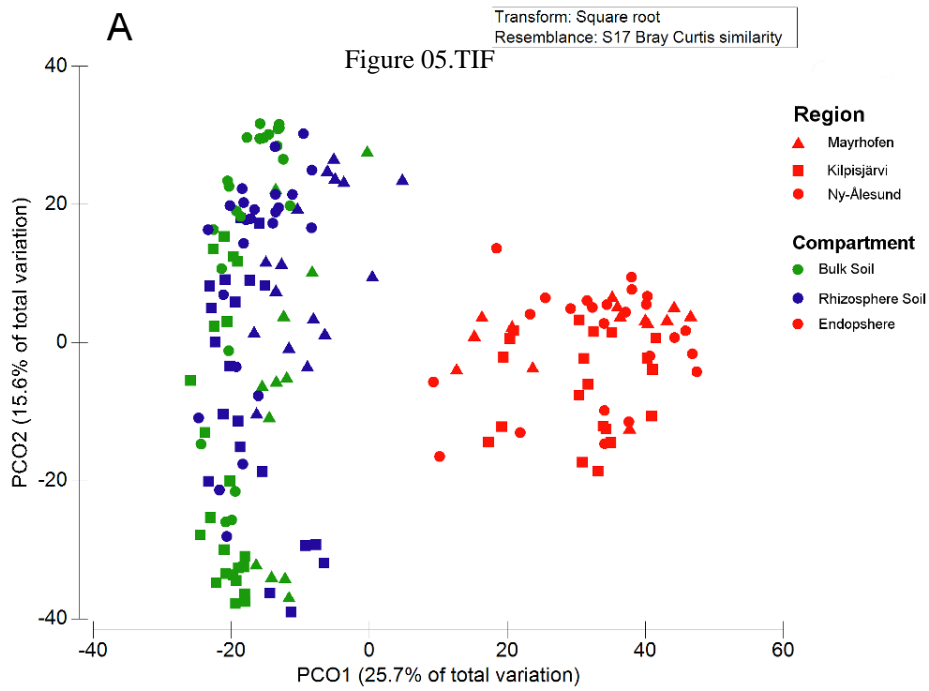


Figure 06.TIFF

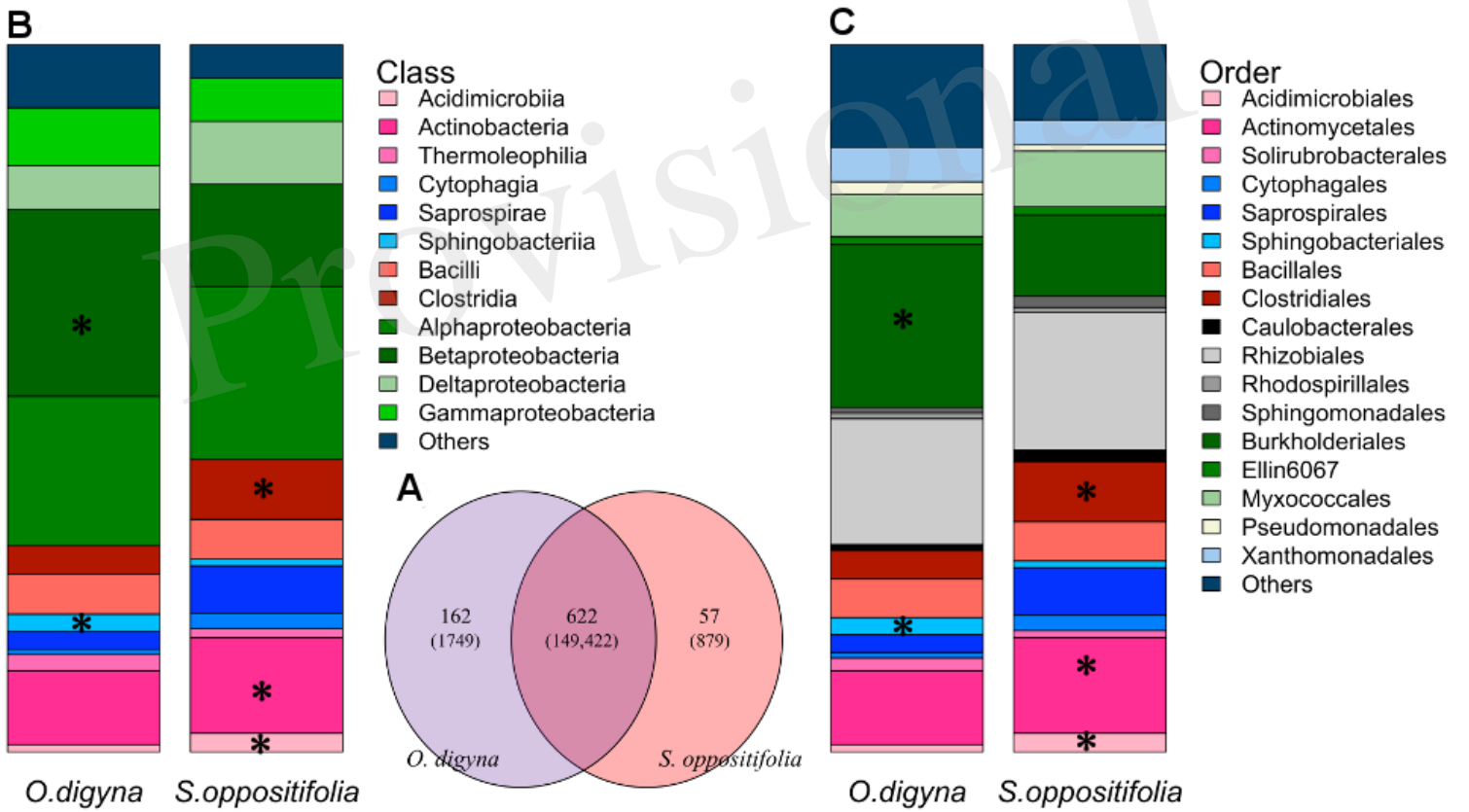


Figure 07.TIF

