Metagenomic exploration of soils microbial communities associated to Antarctic vascular plants

Marco A. Molina-Montenegro [Corresp., 1, 2] , **Gabriel I. Ballesteros** [1] , **Eduardo Castro-Nallar** [3] , **Claudio Meneses** [4, 5] , **Cristian Torres-Díaz** [6] , **Jorge Gallardo-Cerda** [1]

[1] Centro de Estudios Avanzados en Ecología Molecular y Funcional, Instituto de Ciencias Biológicas, Universidad de Talca, Talca, Chile

[2] Centro de Estudios Avanzados en Zonas Áridas (CEAZA), Facultad de Ciencias del Mar, Universidad Católica del Norte, Coquimbo, Chile

[3] Centro de Bioinformática y Biología Integrativa, Facultad de Ciencias Biológicas, Universidad Andrés Bello, Santiago, Chile

[4] Centro de Biotecnología Vegetal, Facultad de Ciencias Biológicas, Universidad Andrés Bello, Santiago, Chile

[5] FONDAP, Center for Genome Regulation, Santiago, Chile

[6] Grupo de Biodiversidad y Cambio Global (GBCG), Departamento de Ciencias Básicas, Universidad del Bío-Bío, Santiago, Chile

Corresponding Author: Marco A. Molina-Montenegro
Email address: marco.molina@utalca.cl

Antarctica is one of the most stressful ecosystems worldwide with few vascular plants, which are limited by abiotic conditions. Here, plants such as *Deschampsia antarctica* (*Da*) could generate more suitable micro-environmental conditions for the establishment of other plants as *Colobanthus quitensis* (*Cq*). Although, plant-plant interaction is known to determine the plant performance, little is known about how microorganisms might modulate the ability of plants to cope with stressful environmental conditions. Several reports have focused on the possible ecological roles of microorganism with vascular plants, but if the rizospheric microorganisms can modulate the positive interactions among vascular Antarctic plants has been seldom assessed. In this study, we compared the rhizosphere microbiomes associated with *Cq*, either growing alone or associated with *Da*, using a shotgun metagenomic DNA sequencing approach and using eggNOG for comparative and functional metagenomics. Overall, results show higher diversity of taxonomic and functional groups in rhizospheric soil from *Cq+Da* than *Cq*. On the other hand, functional annotation shows that microorganisms from rhizospheric soil from *Cq+Da* have a significantly higher representation of genes associated to metabolic functions related with environmental stress tolerance than *Cq* soils. Additional research is needed to explore both the biological impact of these higher activities in terms of gene transfer on plant performance and in turn help to explain the still unsolved enigma about the strategy deployed by *Cq* to inhabit and cope with harsh conditions prevailing in Antarctica.

1   **Metagenomic exploration of soils microbial communities associated to**

2   **Antarctic vascular plants**

3

4   Marco A. Molina-Montenegro[1,2,]*, Gabriel I. Ballesteros[1], Eduardo Castro-Nallar[3], Claudio

5   Meneses[4,5], Cristian Torres-Díaz[6] and Jorge Gallardo-Cerda[1]

8   [1]Centro de Ecología Molecular y Funcional, Instituto de Ciencias Biológicas, Universidad de

9   Talca, Avda. Lircay s/n, Talca, Chile.

10   [2]Centro de Estudios Avanzados en Zonas Áridas (CEAZA), Facultad de Ciencias del Mar,

11   Universidad Católica del Norte, Coquimbo, Chile.

12   [3]Centro de Bioinformática y Biología Integrativa, Facultad de Ciencias Biológicas, Universidad

13   Andrés Bello, Santiago, Chile.

14   [4]Centro de Biotecnología Vegetal, Facultad de Ciencias Biológicas, Universidad Andrés Bello,

15   Santiago, Chile

16   [5]FONDAP, Center for Genome Regulation, Santiago, Chile

17   [6]Grupo de Biodiversidad y Cambio Global (GBCG), Departamento de Ciencias Básicas,

18   Universidad del Bío-Bío, Chillán, Chile.

19

20

21

22   *Corresponding author: Marco A. Molina-Montenegro, Universidad de Talca, Avda. Lircay s/n,

23   Talca, Chile; E-mail: marco.molina@utalca.cl

ABSTRACT

24

25 Antarctica is one of the most stressful ecosystems worldwide with few vascular plants, which are

26 limited by abiotic conditions. Here, plants such as *Deschampsia antarctica* (*Da*) could generate

27 more suitable micro-environmental conditions for the establishment of other plants as

28 *Colobanthus quitensis* (*Cq*). Although, plant-plant interaction is known to determine the plant

29 performance, little is known about how microorganisms might modulate the ability of plants to

30 cope with stressful environmental conditions. Several reports have focused on the possible

31 ecological roles of microorganism with vascular plants, but if the rizospheric microorganisms can

32 modulate the positive interactions among vascular Antarctic plants has been seldom assessed. In

33 this study, we compared the rhizosphere microbiomes associated with *Cq*, either growing alone or

34 associated with *Da*, using a shotgun metagenomic DNA sequencing approach and using eggNOG

35 for comparative and functional metagenomics. Overall, results show higher diversity of

36 taxonomic and functional groups in rhizospheric soil from *Cq+Da* than *Cq*. On the other hand,

37 functional annotation shows that microorganisms from rhizospheric soil from *Cq+Da* have a

38 significantly higher representation of genes associated to metabolic functions related with

39 environmental stress tolerance than *Cq* soils. Additional research is needed to explore both the

40 biological impact of these higher activities in terms of gene transfer on plant performance and in

41 turn help to explain the still unsolved enigma about the strategy deployed by *Cq* to inhabit and

42 cope with harsh conditions prevailing in Antarctica.

INTRODUCTION

43

44 Diverse mutualistic bacteria and fungi thrive on plant surfaces and inhabit most plant tissues.

45 Many of these microorganisms interact with their plant hosts intimately; they can influence plant

46 metabolism and hormonal pathways in addition to providing novel nutritional or biosynthetic

47 capacities stimulating plant growth and conferring enhanced resistance to different stressors

48 (Lugtenberg & Kamilova, 2009; de Zelicourt *et al.* 2013).

49 Several studies have shown that microorganisms can have a direct effect on the plant

50 capacity to resist biotic and abiotic stress such as herbivory, drought, extreme temperatures and

51 high salinity (Redman *et al.,* 2002; Marquez *et al.,* 2007; Giauque & Hawkes, 2013). Many

52 bacteria and fungi have been found in association with plant roots, facilitating the establishment,

53 spread and/or plant fitness in stressful environments (Frey-Klett *et al.,* 2007; Bano & Fatima,

54 2009; Hoffman & Arnold, 2010). On the other hand, it has been documented that some

55 microorganisms can modulate the interaction between plants or filtering the establishment of new

56 species in a determinate community (Amsellem *et al.* 2017). Therefore, microorganisms have

57 been shown to have great impact on plant-plant interactions, thus studying the diversity and

58 composition of microbial communities is key to understand how vascular plants survive under the

59 stresfull environmental conditions of Antarctic habitats.

60 The Antarctic ecosystem is one of the most stressful natural habitats, especially for

61 terrestrial plants (Convey *et al.* 2014; Pointing *et al.* 2015). Likewise, only two vascular plants

62 have colonized the Antarctic environment, *Colobanthus quitensis* (Caryophyllaceae) and

63 *Deschampsia antarctica* (Poaceae) (Moore, 1970). Although both plants colonize Antarctica, *C.*

64 *quitensis* is mainly found growing in association with *D. antarctica* in more stressful areas and

65 alone in low-stress area, while *D. antarctica* is capable of growing alone in areas with higher

66 abiotic stress (Alberdi *et al.* 2002). *D. antarctica* is a grass that form tussocks where micro-

67 environmental conditions above and below their canopy could be milder than outside, acting like

68   a "nurse species" for other less tolerant species (e.g., *C. quitensis*) in Antarctica (see Molina-

69   Montenegro *et al.* 2013). In fact, some native and invasive species increase its physiological

70   performance and fitness-related traits when growing in association with *D. antarctica* compared

71   with those growing outside of them (Molina-Montenegro *et al*., under review). Although it is

72   clear that positive interactions can determine the performance and survival for some less tolerant

73   species, the underlying mechanisms are not clear, and/or whether microorganisms mediate this

74   positive interaction remains unknown.

75   Positive inter-specific interactions play a pivotal role in the structure and functioning of

76   several plant communities. Despite their known potential to drive the ability of plants to cope

77   with stressful environmental conditions, little is known about how microorganisms might affect

78   plant performance in this environment (but see, Torres-Díaz *et al.* 2016). Symbiotic interactions

79   between fungi and/or bacteria, and higher Antarctic plants in the Antarctic environment have

80   been demonstrated (Rosas *et al.* 2009). In fact, several reports have focused on the occurrence,

81   type of association, diversity and possible ecological roles of microorganism interactions with

82   vascular plants (Upson *et al.* 2009, Torres-Díaz *et al.* 2016). Nevertheless, to the best of our

83   knowledge, this work is the first study assessing if the rizospheric microorganisms could

84   modulate the positive interactions among vascular Antarctic plants.

85   In this study, we compare the rhizosphere microbiomes associated with *C. quitensis,* either

86   growing alone or associated with *D. antarctica,* using shotgun metagenomic DNA sequencing

87   technology for comparative metagenomics. This approach allows us to gain insight into the

88   rhizospheric microbial community structure associated with *C. quitensis* and *C. quitensis + D.*

89   *antarctica*, through the study of soil's microbial taxonomic diversity, including non-culturable

90   organisms. Such analysis could also provide valuable information regarding microbial functional

91   diversity (Nesme *et al*., 2016). This functional diversity might be playing important roles in

92   conferring different degrees of tolerance to Antarctica's harsh environmental conditions such as

low temperatures, desiccation, and low water and nutrient availability, which could help to explain the "enigma" of the success of these plant species in such harsh environments (sensu, Smith, 2003).

MATERIALS AND METHODS

Site description and soil sample processing

Rhizospheric soil samples were collected from Devils Point, Byers Peninsula, Livingstone Island, Antarctica (62°40'11.8"S; 61°10'20.7"W) during Summer growing season (February 2016; Figure 1). *Colobanthus quitensis* rhizospheric soil (*Cq*) and rhizospheric soil of *C. quitensis* growing associated with *Deschampsia antarctica* (*Cq+Da*) were sampled at sea level (Figure 1). Plants were dug out using a sterilized shovel and were transferred to sterilized polyethylene bags to avoid excessive desiccation during transport and were stored at 4 °C. Bulk soil was discarded by vigorously shaking the plants by hand until non-adhering particles were completely removed. Rhizosphere soil was collected by hand shaking the roots in 1 L of a sterile 0.9% NaCl solution to remove adhering soil and soil suspensions were centrifuged in 200 mL sterile tubes to concentrate soil particles in pellet. Supernatants were removed by filtration on 1 mm sieves to eliminate residuals in suspension before DNA extraction and processing. All plant and soil samples were collected under permission of the Chilean Antarctic Institute (INACH; authorization number: 1060/2014).

DNA Extraction and sequencing

117   For a total of six rhizospheric soil samples (three replicates per condition, *Cq* or *Cq+Da*), total

118   DNA was extracted using the PowerSoil DNA Isolation Kit (MoBio Laboratories, Inc). DNA

119   integrity was checked with capillary electrophoresis using a Fragment Analyzer (AATI) and DNA

120   quantification was performed using fluorometry (Qubit 2.0; Qubit DNA Broad Range Assay Kit,

121   Invitrogen). After QC, all samples were subjected to library construction for Illumina sequencing.

122   Briefly, DNA was fragmented by Covaris ultrasonicator (average fragment size of 550bp), and

123   size-selected using AMPure XP purification beads. Libraries were constructed using the TruSeq

124   LT kit following manufacturer instructions (Illumina), ligated to indexed adapters for cluster

125   generation and sequenced using the Illumina MiSeq reagent kit (v3) in an Illumina MiSeq

126   sequencer (600 cycle; 300bp, Paired-End sequencing) (Illumina, San Diego, CA).

127   Demultiplexing and fastq generation were performed automatically using Illumina's built-in

128   software.

129

130   *Bioinformatics and statistical analysis*

131

132   For each sequenced library ($n = 6$), raw sequences were filtered using Trimmomatic 0.36 (Bolger

133   et al., 2014) to remove specific Illumina sequencing adapters, low quality bases and all sequences

134   shorter than 100 bp. Filtered, un-assembled libraries were interleaved and aligned using

135   DIAMOND BLASTX algorithm ver. 0.8.30.92 using default parameters (Buchfink *et al*., 2015)

136   to the NCBI-NR database (June 2017). Only alignments with an e-value of 10^{-3} or lower were

137   included in our analysis. Alignment result files were imported in MEGAN6, which parsed results

138   using the Lowest Common Ancestor algorithm (LCA) and NCBI's taxonomy under default

139   values. All reads aligning to non-bacterial species were not considered for further analysis.

140   Remaining bacterial reads were normalized between samples by using MEGAN6's built-in

141   normalization tool. Functional profile analysis and annotation was performed using the eggNOG

142 (Evolutionary genealogy of genes: Non-supervised Orthologous Groups) orthologous groups and

143 functional annotation database (Huerta-Cepas *et a*l., 2016) included in MEGAN6. Finally, the

144 relative abundance of reads binned to eggNOG clusters of orthologous groups were counted

145 using MEGAN6 for all samples. Comparison of binned reads between *Cq* and *Cq+Da*

146 rhizhospheric soil samples was performed using both a two-sided Welch's t-test with correction

147 for multiple comparisons (Benjamini-Hochberg false discovery rate correction approach, q-value

148 < 0.05) with the software STAMP (Statistical Analysis of Metagenomic Profiles, version 2.3.1;

149 Parks *et a*l., 2014). Relative abundances of binned reads to either functional or taxonomical

150 categories between both samples were compared using a multiple *t*-test with FDR correction (q-

151 value < 0.05). Taxonomic profiles and diversity analysis were carried out in R using the packages

152 ggplot2, phyloseq, vegan, and DESeq2 (Dixon, 2003; Love *et a*l., 2014; McMurdie & Holmes,

153 2013; Wickham, 2009). Raw metagenomic data was deposited in NCBI's Sequence Read Archive

154 database under BioProject ID: PRJNA419970. All supplementary material (both figures and

155 tables) are available at https://figshare.com/s/5d7961c1859f33067dab.

156

157 **RESULTS**

158 *Sequencing results*

159 For the *Cq* metagenome, a total of 20,081,770 reads were obtained, while for *Cq+Da* samples,

160 19,348,212 reads were obtained. Read sizes ranged from 35bp to 300bp, although the majority of

161 reads (~97% of total sequenced reads per library) had a length equal or over 290bp. Filtering

162 steps removed sequencing adapters, low-quality reads and reads shorter than 100bp, reducing

163 library sizes in approximately 18%. These filtered libraries were in turn used for Taxonomical

164 and Functional analysis with DIAMOND and MEGAN6 as described previously. Sequencing

165 statistics are shown in Supplementary table 1.

166

167 *Taxonomic analysis and Rhizospheric soil diversity*

168 Metagenomic analyses were conducted by importing DIAMOND BLASTX results to MEGAN6.

169 DIAMOND found one or more significant alignments for 49% (15.7 million reads) out of the 32

170 million filtered reads used as input. Taxonomic analysis showed that Bacteria, followed by

171 Eukaryota and Archaea dominated both samples (98%, 0.2% and 1.7%, respectively). Our

172 analysis also indicated that *Cq* samples had significantly higher relative abundances for

173 Eukaryota compared to *Cq+Da* (mean difference between samples: 60,744.3; p-value = $7.6e^{-4}$).

174 Specifically, a high number of sequences were aligned to *Viridiplantae* in *Cq* compared to

175 *Cq+Da* (mean difference between samples: 58542,3; p-value = $1.4e^{-3}$), hence we suspect that

176 some plant tissue was also taken during rhizospheric soil sampling in both cases, although it

177 seems that a higher amount of plant tissue was present in the *Cq* samples. Therefore, all non-

178 bacterial aligned reads were filtered and removed from the following analyses. Only bacterial-

179 mapped reads were both taxonomically and functionally analyzed (~15 million aligned reads)

180 using MEGAN6 and eggNOG database. Microbiome analyses at the Phylum level from all soil

181 samples showed that Proteobacteria was the most abundant Phylum (33.7%), followed by

182 Actinobacteria (23.5%) and Bacteroidetes (16.2%) (Figure 2; Core Microbiome is shown in

183 Supplementary figure 1). Interestingly, comparative taxonomic analysis showed differences

184 regarding bacterial species present in our samples; only 46.1% of bacterial species was found in

185 both rhizospheric soil samples, while 25.44% was exclusively found in *Cq* rhizospheric soil and

186 28.5% was found in *Cq+Da* rhizospheric soil (Supplementary figure 2).

187   Regarding alpha diversity, all *Cq+Da* samples were consistently found to be more diverse

188 than *Cq* samples, both in richness (observed and corrected [Chao1]) and evenness (ACE;

189 Shannon; Simpson; Inverse Simpson) (Figure 3). While interquartile ranges did not overlap under

190 any metric, suggesting significant differences, these are sensitive to low sample sizes for which

191 more samples need to be collected and evaluated to obtain precise alpha diversity estimates.

192　However, median values between Observed and Chao1 estimates are in agreement (~755 taxa),

193　which suggests sequencing depth is adequate for sampling these rhizosphere communities.

194

195　*Functional Rhizospheric soil sample comparison*

196　Functional categorization and annotation of sequenced microorganisms from *Cq* and *Cq-Da* was

197　performed by analyzing mapped reads to the NR and eggNOG databases using MEGAN6. 6.8

198　million reads (37.7%) were annotated and associated with at least one cluster of orthologous

199　genes of the following top hierarchies: cellular processes and signaling, information storage and

200　processing, and metabolism. The most represented functional categories at eggNOG level 2 (i.e.,

201　the categories with the higher proportion of reads assigned to it) are aminoacid transport and

202　metabolism (12%), replication, recombination and repair (10.9%) and cell

203　wall/membrane/envelope biogenesis (7.2%) (Supplementary table 2). We also found differences

204　in the microbial communities' functional profiles between the two rhizospheric soils at eggNOG

205　level-2, where 14 functional genes categories (out of 22) had significantly higher relative

206　abundances between rhizospheric soil samples (p-value < 0.05; Figure 4). Of these 14 categories,

207　10 were highly represented in *Cq* (Figure 4, blue dots) while the other 4 were highly represented

208　in *Cq+Da* (Figure 4, orange dots). In addition, mean differences ranged between 0.06% to 1.1%

209　(Supplementary table 4; Figure 4). The differences in terms of relative abundance (1.1%) were

210　found for the eggNOG category "Replication, recombination and repair" which is more

211　represented in the *Cq+Da* condition. A global, deep comparison of reads binned to different

212　eggNOG terms (N=11,732 categories) showed eight terms with significative representation

213　differences between both samples. Interestingly, we found a significantly higher representation of

214　Serine/Threonine protein kinases (COG0515; 54% more represented) in *Cq+Da* samples, which

215　in bacteria have been linked to phosphorylation of serine or threonine residues in proteins, being

216　a key mechanism in regulation of protein activity and control cellular functions such as stress

217 response (Pereira et al., 2011). Other over-represented terms in *Cq+Da* are DNA ligases

218 (COG1793; 86% more represented) and transcriptional regulators (COG3903; 71.9% more

219 represented). These results suggest a higher transcriptional potential activity and/or DNA repair

220 in microorganisms from this sample (possible due to the existence of transposable insertion

221 sequences), which could be a consequence of microorganisms being exposed to a more

222 challenging environment compared to *Cq* growing alone.

223

224 **DISCUSSION**

225 In this study, we explored the species composition and functional genomics of rhizosphere soil

226 samples associated with *C. quitensis* growing alone (*Cq*) and *C. quitensis* growing in association

227 with *D. antartica* (*Cq+Da*). Shotgun metagenomics has proved to be useful for assessing

228 microbial community structure, composition and abundance (Castañeda & Barbosa, 2017), while

229 also providing a foundational sequence dataset allowing genomic functional analysis. Our results

230 showed that bacterial species had the highest relative abundance in both habitats (98%) compared

231 to Archaea (0.22%) and Eukaryota (1.77%), whereas the most abundant bacterial Phyla are

232 Proteobacteria, Actinobacteria, Bacteroidetes, Acidobacteria and Firmicutes, accounting for

233 approximately 85% of the sequences in rhizosphere soil samples. These Phyla have been

234 described as widespread and often abundant in other different soil samples (Aislabie *et al*., 2013;

235 Imchen *et a*l., 2017; Lauber *et a*l., 2009; Yang, 2015), including Antarctic soils (Bottos *et a*l.,

236 2014; da Silva *et a*l., 2017; Teixeira *et al*., 2013) and may constitute a core root microbiome

237 playing a pivotal role in plant growth promotion, nutrient acquisition and abiotic stress tolerance

238 (Chen *et al*., 2017). Proteobacteria is a metabolically diverse group that has been related to

239 nutrient cycling, including carbon, nitrogen and sulfur, by degrading soluble organic molecules

240 such as organic acids, aminoacids and sugars (Eilers *et al*., 2010). The relative abundance of

241 Proteobacteria found in our study is similar to the reported in other soil studies (Castañeda &

242   Barbosa, 2017). Actinobacteria also participates actively in the terrestrial carbon cycling through

243   organic matter turnover, breakdown of recalcitrant molecules and production of secondary

244   metabolites (de Menezes *et a*l., 2015). Interestingly, a higher relative abundance of Bacteroidetes

245   was found in *Cq* compared to *Cq+Da* samples (19.08% and 13.53%, respectively); Bacteroidetes

246   are involved in degradation of plant material and related organic molecules such as starch and

247   cellulose (Aislabie *et al*., 2013) and their abundance correlates with soil pH, available

248   nitrogen/phosphorus content and water content (Zhang *et al*., 2014). A possible explanation for

249   the differences observed in terms of relative abundance of certain Phyla may be that plants are

250   thought to be involved in shaping the rhizosphere by differentially altering bacterial species

251   across these sites, and these altered bacterial communities may provide beneficial services to the

252   host plants (Mahoney *et al*., 2017). Furthermore, the observed differences in terms of unique and

253   shared bacterial species present in both rhizospheres (Supplementary figure 2) could be a

254   consequence of *C. quitensis* plants differentially shaping the rhizosphere composition compared

255   to *C. quitensis* growing associated with *D. antartica*. It is thought that rhizospheric effects depend

256   specifically on the plant species, as plants significantly influence and shape soil microbial

257   communities through exudation of unique, species-specific root compounds, which in turn

258   decrease microbial diversity in the rhizosphere compared to bulk soil (Kielak *et al*., 2008).

259   However, a previous study did not find any clear difference between microbial communities in

260   the rhizosphere of *C. quitensis* and *D. antartica* growing separately in Admiralty Bay (Teixeira *et*

261   *al*., 2010), while finding significant differences between microbial communities in bulk soil and

262   rhizospheric samples from the same plant species (Teixeira *et al*., 2013). Similar results have

263   been found in other studies; plant species composition appeared to have a reduced effect on the

264   diversity, structure or size of associated rhizosphere bacterial communities (Kielak *et al*., 2008).

265   Although plant roots generally influence microbial community composition and diversity in the

266 rhizosphere, dominant soil physicochemical factors could be a determinant factor in shaping

267 microbial communities rather than plant species (Nunan *et al*., 2005; Singh *et al*., 2007).

268 Metagenomic shotgun sequencing and functional annotation using eggNOG functional

269 categories revealed that "metabolism", was the highest represented category, followed by

270 "cellular process and signaling", and "information storage and processing". In the metabolism

271 category, the highest represented terms were those related to "aminoacid transport and

272 metabolism", "energy production and conversion", "carbohydrate transport and metabolism" and

273 "inorganic ion transport and metabolism". This suggests that microbial communities present in

274 these soils could be playing an important role in both enhancing nitrogen cycling in limited

275 nitrogen ecosystems and in carbohydrate degradation through organic matter decomposition,

276 indicating a higher carbon cycling activity in these soils which in turn might help to provide the

277 necessary energy and precursor materials for defense responses and secondary metabolism under

278 abiotic stress (Lin *et al*., 2013). The high number of reads annotated in these three categories

279 (30.4%) suggests that rhizosphere microorganisms have high metabolic capabilities (da Silva *et*

280 *al*., 2017), which may be relevant for plant-microorganism interaction in other rhizospheric

281 communities (Yan *et al*., 2016). We also found that relative abundance of sequences assigned to

282 eggNOG subclasses were similar among samples, although some small but significant differences

283 were detected (0.06% to 1.1%). Similar trends, in terms of relative differences, have been found

284 for other soil samples (Castañeda & Barbosa, 2017). The higher differences in terms of relative

285 abundance (1.1%) were found for the eggNOG category "Replication, recombination and repair",

286 which is more represented in the rhizosphere of *C. quitensis* growing with *D. antarctica*. A

287 detailed analysis for this category showed that the top five sub-categories with higher abundance

288 in *Cq+Da* compared to *Cq* are transposases, resolvases and RNA-directed DNA polymerases (or

289 reverse transcriptases). Transposases are enzymes that catalyze movement of transposons

290 ("jumping genes") to other parts of the genome, and require both DNA polymerases and ligases

291 (which were also over-represented in *Cq+Da*) to fill gaps and complete the recombination

292 process (Zhang *et al.*, 2009). Insertion of a transposon in a certain genomic position may change

293 underlying DNA sequences, subsequently activating or inactivating gene expression which

294 ultimately may have an impact on their hosts' fitness (Aziz *et al.*, 2010). Recently, transposases

295 have been linked to horizontal/lateral gene transfer events in bacteria, leading to potential gene

296 sequence and functionality being shared between microorganisms in order to thrive under similar

297 environmental conditions (Cuecas *et al.*, 2017). Resolvases (also known as recombinases) are

298 enzymes involved in site-specific DNA recombination, and may play a role in the spread of

299 chromosomal genes in the plant rhizosphere through retro-transference of DNA, and occurs in

300 bacteria of different genera (Ronchel *et al.*, 2000). These results suggest that *Cq+Da* rhizospheric

301 soil has an enrichment of several functional categories linked to regulation of protein activity

302 (through Serine/Threonine protein kinase), gene transcriptional activation/inactivation through

303 transposons and related enzymes (Transposases, DNA ligases), regulation of gene expression

304 through transcriptional regulators, and replication of retroid elements (such as retrotransposons)

305 which could be involved in spreading genes between organisms through genetic exchange. These

306 over-represented bacterial mechanisms could be contributing to increase plant fitness by

307 stimulating plant growth and/or by conferring enhanced resistance to abiotic stress, hence

308 allowing plants to grow in extreme environmental conditions (De Zelicourt *et al.*, 2013).

309

310 **CONCLUSION**

311 We explored the taxonomic and functional diversity of microbial communities in rhizospheric

312 soils of the vascular species from Antarctica using shotgun sequencing. Our metagenomics

313 analyses revealed that bacterial communities are similar in terms of taxonomic composition

314 independently of the plant species upon which rhizospheric soil was collected. However, these

315 similarities could be limited by taxonomic assignment based on BLASTx alignments using the

316   NR non-redundant protein database as reference and MEGAN6 accession parsing tool. However,

317   the results from these comparisons are valid since all samples were subjected to the same bias.

318   Conversely, some functional categories had shown significant relative differences in terms of

319   abundance between rhizospheres, suggesting that these microbial communities could have a

320   higher activity in terms of gene transfer, which ultimately could have an effect on plant's growth

321   and colonization. Additional research is needed to explore both the biological impact of these

322   higher activities in terms of gene transfer on plant performance and to explain the still unsolved

323   enigma about the strategy deployed by *C. quitensis* to inhabit and cope with the harsh abiotic

324   conditions prevailing in Antarctica (see Smith, 2003).

325

326

327   **Acknowledgements**

333   **REFERENCES**

334   Aislabie J, Deslippe JR, Dymond JR (2013) Soil microbes and their contribution to soil services.

335   In: Dymond JR (ed) Ecosystem Services in New Zealand - Condition and trends, pp 143–

336   161.

337   Alberdi M, Bravo LA, Gutierrez A, Gidekel M, Corcuera LJ (2002) Ecophysiology of Antarctic

338   vascular plants. Physiol Plant 115:479-486.

339   Amsellem L, Brouat C, Duron O, Porter SS, Vilcinskas A, Facon B (2017) Advances in

340   Ecological Research 57:99-146.

341   Aziz RK, Breitbart M, Edwards RA (2010) Transposases are the most abundant, most ubiquitous

342   genes in nature. Nucleic Acids Res 38:4207-4217.

343   Bano A, Fatima M (2009) Salt tolerance in *Zea mays* (L). following inoculation

344   with *Rhizobium* and *Pseudomonas*. Biology and Fertility of Soils 45:405-413.

345   Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence

346   data. Bioinformatics 30:2114–2120.

347   Bottos EM, Woo AC, Zawar-Reza P, Pointing SB, Cary SC (2014) Airborne Bacterial Populations

348   Above Desert Soils of the McMurdo Dry Valleys, Antarctica. Microb Ecol 67:120–128.

349   Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND.

350   Nat Methods 12: 59–60.

351   Castañeda LE, Barbosa O (2017) Metagenomic analysis exploring taxonomic and functional

352   diversity of soil microbial communities in Chilean vineyards and surrounding native forests.

353   PeerJ 5:e3098.

354   Chen L, Xin X, Zhang J, Redmile-Gordon M, Nie G (2017) Soil Characteristics Overwhelm

355   Cultivar Effects on the Structure and Assembly of Root-Associated Microbiomes of Modern

356   Maize. Pedosphere (in press).

357   Convey P, Chown SL, Clarke A, Barnes DKA, Bokhorst S, Cummings V, Ducklow HW, et al

358   (2014) The spatial structure of Antarctic biodiversity. Ecol Monogr 80:203-244.

359 Cuecas A, Kanoksilapatham W, González JM (2017) Evidence of horizontal gene transfer by

360   transposase gene analyses in Fervidobacterium species. PLoS ONE 12:e0173961.

361 da Silva AC, da Costa-Rachid CTC, de Jesus HE, Rosado AS, Peixoto RS (2017) Predicting the

362   biotechnological potential of bacteria isolated from Antarctic soils, including the rhizosphere

363   of vascular plants. Polar Biol 40:1393–1407.

364 de Menezes AB, Prendergast-Miller MT, Poonpatana P, Farrell M, Bissett A, Macdonald LM,

365   Toscas P, et al. (2015) C/N Ratio Drives Soil Actinobacterial Cellobiohydrolase Gene

366   Diversity. Appl Environ Microbiol 81:3016–3028.

367 De Zelicourt A, Al-Yousif M, Hirt H (2013) Rhizosphere microbes as essential partners for plant

368   stress tolerance. Mol Plant 6:242–245.

369 Dixon P (2003) VEGAN, a package of R functions for community ecology. J Veg Sci 14:927–

370   930.

371 Eilers KG, Lauber CL, Knight R, Fierer N (2010) Shifts in bacterial community structure

372   associated with inputs of low molecular weight carbon compounds to soil. Soil Biol

373   Biochem 42:896–903.

374 Frey-Klett P, Tarkka GM (2007) The mycorrhiza helper bacteria revisited. New Phytol 176:22-36.

375 Giauque H, Hawques CV (2013) Climate affects symbiotic fungal endophyte diversity and

376   performance. Am J Bot 7:1435-1444.

377 Hoffman MT, Arnold AE (2010) Diverse Bacteria Inhabit Living Hyphae of Phylogenetically

378   Diverse Fungal Endophytes. Appl Environ Microbiol 12:4063-4075.

379 Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, et al. (2016)

380   eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for

381   eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res 44:D286–D293.

382 Imchen M, Kumavath R, Barh D, Avezedo V, Ghosh P, Viana M, Wattam AR (2017) Searching

383   for signatures across microbial communities: Metagenomic analysis of soil samples from

384   mangrove and other ecosystems. Sci Rep 7:1–13.

385 Kielak A, Pijl AS, Van Veen JA, Kowalchuk GA (2008) Differences in vegetation composition

386   and plant species identity lead to only minor changes in soil-borne microbial communities in

387   a former arable field. FEMS Microbiol Ecol 63:372–382.

388 Lauber CL, Hamady M, Knight R, Fierer N (2009) Pyrosequencing-based assessment of soil pH

389   as a predictor of soil bacterial community structure at the continental scale. Appl Environ

390   Microbiol 75:5111–5120.

391 Lin W, Wu L, Lin S, Zhang A, Zhou M, Lin R, Wang H, et al. (2013) Metaproteomic analysis of

392   ratoon sugarcane rhizospheric soil. BMC Microbiol 13:135.

393 Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for

394   RNA-seq data with DESeq2. Genome Biol 15:550.

395 Lugtenberg B, Kamilova F (2009) Plant-Growth-Promoting Rhizobacteria. Annu Rev Microbiol

396   63: 541-546.

397 Mahoney AK, Yin C, Hulbert SH (2017) Community Structure, Species Variation, and Potential

398   Functions of Rhizosphere-Associated Bacteria of Different Winter Wheat (*Triticum*

399   *aestivum*) Cultivars. Front Plant Sci 217 8:132.

400 Márquez LM, Redman RS, Rodriguez RJ, Roossinck MJ (2007) A virus in a fungus in a plant:

401   three-way symbiosis required for thermal tolerance. Science 315:513-515.

402 McMurdie PJ, Holmes S (2013) phyloseq: An R Package for Reproducible Interactive Analysis

403   and Graphics of Microbiome Census Data. PLoS One 8:e61217.

404 Molina-Montenegro MA, Ricote-Martínez N, Muñoz-Ramírez C, Torres-Díaz C, Gómez-

405   González S, Gianoli E (2013) Positive interactions between the lichen *Usnea antarctica*

406   (Parmeliaceae) and the native flora in maritime antarctica. J Veg Sci 24:463-472.

407 Moore DM (1970) Studies in Colobanthus quitensis (Kunth) Bartl. and Deschampsia antarctica

408   Desv. II. Taxonomy, distribution and relationships. British Antarctic Survey Bulletin 23:63-

409   80.

410 Nesne J, Achouak W, Agathos SN, Bailey M, Baldrian P, Brunel D, et al. (2016) Back to the

411   Future of Soil Metagenomics. Front Microbiol 7:73.

412 Nunan N, Daniell TJ, Singh BK, Papert A, McNicol JW, Prosser JI (2005) Links between Plant

413   and Rhizoplane Bacterial Communities in Grassland Soils, Characterized Using Molecular

414   Techniques. Appl Environ Microbiol 71:6784–6792.

415 Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic

416   and functional profiles. Bioinformatics 30:3123–3124.

417 Pointing SB, Büdel B, Convey P, Gillman LN, Körner C, Leuzinger S, Vincent WF (2015)

418   Biogeography of photoautotrophs in the high polar biome. Front Plant Sci 6:692.

419 Redman RS, Sheehan KB, Stout RG, Rodriguez RJ, Henson JM (2002) Thermotolarance

420   generated by plant/fungal symbiosis. Science 298:1581

421 Ronchel MC, Ramos-Díaz MA, Ramos JL (2000) Retrotransfer of DNA in the rhizosphere.

422   Environ Microbiol 2:319–323.

423 Rosa LH, Vaz ABM, Caligiorne RB, Campolina S, Rosa CA (2009) Endophytic fungi associated

424   with the Antarctic grass *Deschampsia antarctica* Desv. (Poaceae). Polar Biol 32:161-167.

425 Singh BK, Munro S, Potts JM, Millard P (2007) Influence of grass species and soil type on

426   rhizosphere microbial community structure in grassland soils. Appl Soil Ecol 36:147–155.

427 Smith RIL (2003) The enigma of *Colobanthus quitensis* and *Deschampsia antarctica* in

428   Antarctica. In: Huiskes AHL, Gieskes WWC, Rozema J, Schorno RML, van der Vies SM,

429   Wolff WJ (eds). Antarctic biology in a global context. Leiden: Backhuys Publishers, pp 234–

430   239.

431 Teixeira LCRS, Peixoto RS, Cury JC, Sul WJ, Pellizari VH, Tiedje J, Rosado AS (2010) Bacterial

432   diversity in rhizosphere soil from Antarctic vascular plants of Admiralty Bay, maritime

433   Antarctica. Isme J 4:989.

434   Teixeira LCRS, Yeargeau E, Balieiro FC, Piccolo MC, Peixoto RS, Greer CW, Rosado AS (2013)

435   Plant and Bird Presence Strongly Influences the Microbial Communities in Soils of

436   Admiralty Bay, Maritime Antarctica. PLoS One 8:e66109.

437   Torres-Díaz C, Gallardo-Cerda J, Lavín P, Oses R. et al (2016) Biological interactions and

438   simulated climate change modulates the ecophysiological performance of *Colobanthus*

439   *quitensis* in the Antarctic ecosystem. PLoS ONE 11:e0164844.

440   Upson R, Newsham KK, Bridge PD, Pearce DA, DJ Read DL (2009) Taxonomic affinities of

441   dark septate root endophytes of *Colobanthus quitensis* and *Deschampsia antarctica*, the two

442   native Antarctic vascular plant species. Fungal Ecol 2:184-196.

443   Wickham H (2009) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

444   Yan Y, Kuramae EE, de Hollander M, Klinkhamer PGL, van Veen JA (2016) Functional traits

445   dominate the diversity-related selection of bacterial communities in the rhizosphere. Isme J

446   11:56.

447   Yang J (2015) Encyclopedia of Metagenomics 1–6.

448   Zhang K, He J, Yang M, Yen M, Yin J (2009) Identifying Natural Product Biosynthetic Genes

449   from a Soil Metagenome by Using T7 Phage Selection. Chem Bio Chem 10:2599–2606.

450   Zhang X, Chen Q, Han X (2014) Soil Bacterial Communities Respond to Mowing and Nutrient

451   Addition in a Steppe Ecosystem. PLoS One 8:e84210.

452

Figure captions

Figure 1. (A) Study site (Devils Point, Byers Peninsula, South Shetland Islands) in Antarctica where individuals of *Colobanthus quitensis* were sampled, growing alone on bare ground (B) or associated to *Deschampsia antarctica* tussocks (C).

Figure 2. Percentage distribution of bacterial phylum from rhizospheric soil samples of *C. quitensis* (*Cq*) and *C. quitensis + D. antarctica* (*Cq+Da*).

Figure 3. Alpha diversity measure for both *C. quitensis* (*Cq*) and *C. quitensis + D. antarctica* (*Cq+Da*).

Figure 4. EggNOG functional categories found in rhizospheric soil bacteria communities. Bar plot shows mean proportion (%) of functional categories found in rhizospheric bacterial communities based on the EggNOG database, level 2 categories. Points indicate the differences between *C. quitensis* and *C. quitensis + D. antarctica* soils (blue and orange bars, respectively). Corrected p-values (q-values) were derived from a Welch's t-test with Benjamini-Hochberg correction for false discovery rate.