Biological soil crusts from different soil substrates harbor distinct bacterial groups with the potential to produce exopolysaccharides and lipopolysaccharides

revised version June 19, 2019

Barbara Cania[1], Gisle Vestergaard[1, 2], Susanne Kublik[1], John Maximilian Köhne[3], Thomas Fischer[4], Andreas Albert[5], Barbro Winkler[5], Michael Schloter[1, 6], Stefanie Schulz[1]

[1]Research Unit Comparative Microbiome Analysis, Helmholtz Zentrum München Research Center for Environmental Health (GmbH), Ingolstädter Landstraße 1, 85764 Neuherberg

[2]Section for Bioinformatics, Department of Health Technology, Technical University of Denmark, 2800 Lyngby, Denmark

[3]Department of Soil System Science, Helmholtz Centre for Environmental Research (UFZ), Theodor-Lieser-Straße 4, 06120 Halle, Germany

[4]Central Analytical Laboratory, Brandenburg Technical University, Konrad-Wachsmann-Allee 6, 03046 Cottbus, Germany

[5]Research Unit Environmental Simulation, Helmholtz Zentrum München Research Center for Environmental Health (GmbH), Ingolstädter Landstraße 1, 85764 Neuherberg

[6]Chair for Soil Science, Technical University of Munich, Emil-Ramann-Straße 2; 85354 Freising, Germany

Corresponding author: Stefanie Schulz, stefanie.schulz@helmholtz-muenchen.de, +49 89 3187-3054

Acknowledgements

The authors wish to thank Gudrun Hufnagel for measuring the biochemical parameters, Christoph Schmidt and Abilash Chakravarthy Durai Raj for bioinformatical advice, as well as Viviane Radl and Antonios Michas for constructive feedback on the previous version of the manuscript. This study was performed as part of the Transregional Collaborative Research Centre 38 (SFB/TRR 38), which is financially supported by the Deutsche Forschungsgemeinschaft (DFG, Bonn) and the Brandenburg Ministry of Science, Research and Culture (MWFK,Potsdam), and the project "The influence of agricultural management practices on microbial functions and networks in biological soil crusts" funded by the DFG in frame of the DFG-Nachwuchsakademie "Agrarökosystemforschung: Bodenressourcen und Pflanzenproduktion". The authors also gratefully acknowledge the funding provided by the German Federal Office for Agriculture and Food (BLE).

Abstract

Biological soil crusts (BSCs) play an important role in improving soil stability and resistance to erosion by promoting aggregation of soil particles. During initial development, BSCs are dominated by bacteria. Some bacterial members of the crusts can contribute to the formation of soil aggregates by producing exopolysaccharides and lipopolysaccharides that act as "glue" for soil particles. However, little is known about the dynamics of "soil glue" producers during the initial development of BSCs. We hypothesized that different types of initial BSCs harbor distinct producers of adhesive polysaccharides. To investigate this, we performed a microcosm experiment, cultivating BSCs on two soil substrates. High-throughput shotgun

sequencing was used to obtain metagenomic information on microbiomes of bulk soils from the beginning of the experiment, and BSCs sampled after four and ten months of incubation. We discovered that the relative abundance of genes involved in the biosynthesis of exopolysaccharides and lipopolysaccharides increased in BSCs compared to bulk soils. At the same time, communities of potential "soil glue" producers that were highly similar in bulk soils underwent differentiation once BSCs started to develop. In the bulk soils, the investigated genes were harbored mainly by *Betaproteobacteria*, whereas in the BSCs, the major potential producers of adhesive polysaccharides were, aside from *Alphaproteobacteria*, either *Cyanobacteria* or *Chloroflexi* and *Acidobacteria*. Overall, our results indicate that the potential to form exopolysaccharides and lipopolysaccharides is an important bacterial trait for initial BSCs, and is maintained despite the shifts in bacterial community composition during BSC development.

Keywords

Biological soil crusts, exopolysaccharides, lipopolysaccharides, microbiome, metagenomics

Introduction

Biological soil crusts (BSCs) are important biotic components of many terrestrial ecosystems [1, 2]. They consist of highly specialized and complex communities of algae, mosses, lichens, fungi, cyanobacteria and other prokaryotes [3]. These organisms live in a close association with soil particles, forming a coherent layer within the uppermost few millimeters of the topsoil, or directly on the soil surface [1]. An important structural element of BSCs is the extracellular polymeric matrix (EPM) which is composed mostly of polysaccharides, and connects organisms and soil particles [4]. EPM ensures BSC integrity, provides protection from external harmful agents, and alters moisture content as well as nutrient availability [4].

EPM also fosters the stabilization of soil aggregates, and protects soils from erosion by wind or water [5-8]. Among organisms forming BSCs, the best-studied producers of polysaccharides are cyanobacteria and algae [1]. However, although not as thoroughly studied, also non-photosynthetic microbial members of BSCs, including fungi, proteobacteria and actinobacteria, are prominent producers of these compounds [9-11].

The composition and chemical properties of polysaccharides in EPM strongly depend on the community of organisms forming BSCs. For example, it has been demonstrated that non-photosynthetic bacteria primarily produce simple polysaccharides, composed mainly of mannose, galactose and glucose [12], while cyanobacteria, algae and fungi produce more complex polysaccharides, which may contain high amounts of non-neutral sugars [13-15]. As it was shown that even slight differences in the sugar composition can result in completely different physical traits of the polysaccharide [16], the properties of EPM could be influenced by any factor that changes the structure of polysaccharide-producing communities. It is known that the composition of organisms forming BSCs changes depending on (i) the developmental stage of BSCs [17-19], (ii) environmental factors like radiation, humidity, elevation, temperature [17-24], and (iii) edaphic factors like soil pH, texture and nutrient content [1, 17, 20, 21]. However, not all members of BSCs have the ability to produce polysaccharides, and little is known about the dynamics of polysaccharide-producing organisms during the development of different types of BSCs.

In this respect, bacterial polysaccharides, specifically exopolysaccharides (EPSs) and lipopolysaccharides (LPSs), are of great interest, as cyanobacteria and non-phototrophic bacteria form BSCs in the initial stage of BSC development [25]. EPSs are either synthesized intracellularly, and excreted by one of three different pathways: the Wzy-dependent pathway, the ABC transporter-dependent pathway and the synthase-dependent pathway, or synthesized extracellularly [26, 27]. In contrast, parts of LPSs are initially synthesized inside a cell, then ligated together at the inner membrane and transported to the cell surface as mature molecules [28, 29]. While LPSs are present in most Gram-negative bacteria [30], EPSs are

exuded by a wide range of taxa [16]. Among the most-recognized producers of EPSs are cyanobacterial members of *Oscillatoria*, *Nostoc*, *Lyngbya* and *Schizothrix*, as well as bacterial members of *Microbacterium*, *Pseudomonas*, *Bacillus*, *Paenibacillus* and *Streptomyces* [31]. These microorganisms are the first colonizers of bare soils, and their EPSs as well as LPSs are considered as essential for the initial consolidation of soil particles and the preparation of conditions for the establishment of cryptogamic surface cover in the later stages of BSC development [32]. Thus, a better understanding of the dynamics of polysaccharide-producing organisms during the initial development of BSCs requires more in-depth knowledge on cyanobacteria and other bacteria that initialize BSC establishment [4].

Many researchers studied polysaccharide-producing bacterial strains that were isolated from BSCs at different stages of development [33-36]. However, data on the community dynamics of bacterial EPS and LPS producers under natural conditions is missing. Thus, our aim was to investigate polysaccharide-producing bacterial communities during the initial stage of BSC development. We assumed that the relative abundance of genes related to EPS and LPS formation would increase once BSC development starts. Moreover, we hypothesized that different types of initial BSCs would harbor different communities of potential EPS and LPS producers. To test our hypotheses, we performed a microcosm experiment cultivating BSCs on two different soil substrates. As the soil substrates came from sites with different types of naturally occurring BSCs, we expected that the BSCs cultivated in the microcosm experiment would also consist of distinct microbial communities. To address our research questions, we used a high-throughput shotgun sequencing of DNA extracted from bulk soils from the beginning of the experiment, as well as initial BSCs sampled after four and ten months of incubation. We employed a bioinformatics pipeline described by Cania et al. [37] targeting genes specific for EPS and LPS production to obtain information on bacteria potentially involved in the production of adhesive polysaccharides.

Materials and Methods

Sites description

Soils for the incubation experiment were collected in 2011 from two sites at the initial stages of ecosystem development. One was the artificial catchment Chicken Creek (51°36'18'' N, 14°15'58'' E) and the other was an initial moving sand dune close to Lieberose (51°55'49'' N, 14°22'22'' E). Both sites are located in the state of Brandenburg in eastern Germany, approximately 37 km apart. The climate of the region is temperate continental with a mean air temperature of 8.9 °C and mean rainfall of 569 mm a^{-1} [19].

The Chicken Creek catchment was constructed in 2005 in an opencast mine near Cottbus by dumping and contouring sand and loamy sand material originating from Pleistocene sediments. Details on the construction works and site conditions are given by Gerwin et al. [38] and Russell et al. [39]. After construction, no restoration was undertaken and the area was allowed to undergo natural succession. BSC development at the site was heterogeneous depending on the appearance of vascular vegetation, which still was dynamic at the time of sampling [40]. For the Lusatian post-mining sites, the cyanobacterial species *Microcoleus vaginatus*, *Nostoc* spec., *Phormidium* spec., *Schizothrix* spec., *Tolypothrix* spec., the green algal species *Bracteacoccus* minor, *Chlorococcum* spec., *Cylindrocytis* spec., *Elliptochloris* spec., *Gloeocysis* spec., *Klebsormidium*, *Chlorella* spec., *Zygogonium* spec., *Ulothrix* spec., *Haematococcus* spec., the lichens *Placynthiella oligotropha* and *Cladonia subulata*, as well as the mosses *Polytrichum piliferum* and *Ceratodon purpureus* were reported [18, 41, 42]. The Chicken Creek site heterogeneity was also reflected by high variability of moss coverages, which were recorded on 107 vegetation monitoring plots each having a size of 5 x 5 m^2, which ranged from 0.1 to 95 % with a median coverage of 30 ±25 % [40]. Terminal successional stages of cryptogamic surface cover development could not be identified, mainly due to BSC extinction caused by vascular plant overgrowth.

The moving sand dune occurs near Lieberose as a result of extensive disturbances of the land surface by former military activities (until approximately 1992). The dune is composed of Pleistocene aeolian sand. A detailed description of the site is provided by Dümig et al. [43], and Fischer and Veste [19]. Depending on their position downslope an inland dune catena, three stages of BSC development could be identified. In microdepressions and at the lee side of tussocks consisting of *Corynephorus canescens* located in the center of the dune slope, dominating sand grains were physically stabilized in their contact zones by accumulated organic matter and by few filamentous algae (BSC stage 1, surface coverage 20 % [44]). At surface patches of the lower dune slope, filamentous algae enmeshed the sand grains and partially filled in the soil pores (BSC stage 2, surface coverage 40 % [44]). BSC stage 3 was characterized by full cover with filamentous and coccoid algae, and by few mosses, the latter covering less than 5 % of the surface. The dominating green algal and moss species were *Zygogonium ericetorum* and *Polytrichum piliferum*, respectively [45]. Cyanobacteria were a minor component within the *Zygogonium* crust, which did not form individual patches, whereas lichens could not be observed at the sampling site. The terminal successional stage of cryptogamic surface cover development, which was found in the vicinity of a less disturbed neighboring Scots pine forest (distance to the sampling site of around 500 m), was characterized by co-appearance of *Cladonia* spec. and *P. piliferum*, which formed dense surface covers.

Sampling and incubation experiment

Bulk soil from the Chicken Creek catchment and the Lieberose sand dune was used to establish a microcosm experiment. A total amount of 100 kg of soil was taken per site from the top 20 cm. At Lieberose, the soil was collected from five spots on the top of the dune, where no plants were growing, while at Chicken Creek, five plant-free spots were used for the soil sampling. The soil was transported and afterwards stored in the dark at room temperature for approximately six months before the incubation

157 experiment. During that time, pre-experiments to adjust the incubation conditions for BSC growth were
158 performed.

159 The soil was mixed and passed through a 2 mm sieve, then packed into plastic pots (10 cm x 10 cm x 10
160 cm) and compacted to the natural soil density of approximately 1.6 g cm^{-3} [43]. In total, the microcosm
161 experiment consisted of 18 pots (9 per site). The water content was set to 50 % of the maximum water
162 holding capacity of the soil samples, and adjusted weekly from the bottom, which ensured very low
163 disturbance for BSC development. Realistic climatic and light conditions were simulated in the sun
164 simulator facility of the Helmholtz Zentrum München (Neuherberg, Germany) by generating the entire
165 spectrum from the ultraviolet (UV, 280-400 nm; UV-B, 280-315 nm; UV-A, 315-400 nm) to the near
166 infrared (NIR) light with a combination of four types of lamps: metal halide lamps (Osram Powerstar HQI-
167 TS 400W/D), quartz halogen lamps (Osram Haloline 400), blue fluorescent tubes (Philips TL-D 36W/BLUE)
168 and UV-B fluorescent tubes (Philips TL 40W/12). The lamps were arranged in several groups to obtain the
169 natural diurnal variations of solar irradiance by switching appropriate groups of lamps on and off. The
170 short-wave cut-off was achieved by selected borosilicate and soda-lime glass filters as previously
171 described [46, 47]. The pots were exposed to radiation for 16 h per day. Maximum radiation was reached
172 in the middle of the day for 8 h at PAR (photosynthetic active radiation, 400-700 nm) of 940 µmol m^{-2} s^{-1},
173 UV-A of 17.7 W m^{-2} and UV-B of 0.37 W m^{-2}. The climatic conditions were adjusted to a night-day cycle
174 from 18 °C to 25 °C, and a relative air humidity of 95 – 90 %, respectively.

175 BSCs were sampled after four (T1) and ten (T2) months of incubation from three independent pots per
176 soil substrate and sampling time point. Only the upper 2 mm were considered as BSC. In addition, samples
177 of bulk soil without BSC development were taken at the beginning of the experiment (T0). In total, 18
178 samples were collected (3 sampling times x 2 sites x 3 replicates). Samples for DNA analyses were directly
179 frozen at -80 °C, while samples for biochemical analyses were stored at 4 °C until further processing. For

180 the determination of water repellency and the computed tomography analysis, undisturbed samples from
181 the end of the experiment were taken using Petri dishes and stored at 4 °C until further analysis.

182

183 *Physicochemical measurements*

184 For the analysis of dissolved organic carbon (DOC) and nitrogen (DON), bulk soils and BSC samples were
185 suspended with a 0.01 M $CaCl_2$ solution in a 1:3 ratio (w/v), and shaken horizontally for 45 min. After
186 passing through a Millex-HV 0.45 µm filter (Merck Millipore, USA), extracts were analyzed for DOC by
187 means of a DIMA-TOC 100 analyzer (Dimatec Analysentechnik GmbH, DE), and for DON – using a Skalar
188 Continuous Flow Analyzer SA5100 (Skalar Analytical B.V., NL) [48]. Soil pH of bulk soil samples was
189 measured in 0.01 M $CaCl_2$ solution with a soil:solution ration of 1:5 (w/v) after 3 hours of incubation time.
190 Water repellency of BSCs was measured as a dimensionless „repellency index“ using the ethanol/water
191 microinfiltrometric sorptivity method according to Fischer et al. [49], where a theoretical value of 1
192 characterizes totally non-repellent soils [50], and may exceed 50 for highly repellent soils [51].

193 Pre-experiments indicated that only BSCs from T2 grown on substrate from Chicken Creek developed a
194 thickness sufficient for visualization by computed tomography (CT). Thus, only these samples were used
195 to determine connectivity of the three-dimensional pore system of the BSCs and the underlying soil as
196 described previously [52]. The structure of the undisturbed samples was analyzed using a micro-computed
197 tomography scanner (X-Tek HMX 225, Nikon Metrology, BE) equipped with a fine-focus X-ray tube (spot
198 size of 5 µm) and a digital flat panel detector with a resolution of 512 by 512 pixels (width by height). The
199 resulting X-ray computed microtomography (XCMT) images were used to calculate Euler characteristics
200 for 26 nearest neighbors of each voxel. So defined Euler numbers were computed as a function of pore
201 size in the range between 15 and 291 µm [53].

202

DNA extraction, library preparation and sequencing

DNA was extracted from bulk soil and BSC samples using the 'Genomic DNA from soil' NucleoSpin Soil Kit (Macherey-Nagel, DE) according to the manufacturer's manual. Based on a pretest performance (data not shown), Buffer SL1 was chosen for sample lysis. DNA purity was verified by means of a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, USA). The quantity was also measured using a SpectraMax Gemini EM microplate reader (Molecular Devices, USA) together with a Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies, USA), and is presented in Table 2. DNA was sheared using an E220 Focused-ultrasonicator (Covaris, USA) with the following conditions: peak incident power = 175 W, duty factor = 10 %, cycles per burst = 200, treatment time = 100 s, temperature = 7 °C, water level = 6, sample volume = 50 µl, intensifier = yes. Library preparation was performed using the NEBNext Ultra DNA Library Prep Kit for Illumina and the NEBNext Multiplex Oligos for Illumina (both New England Biolabs, UK) as described in the protocol of the producer. Due to lower DNA concentrations (Table 2), samples from T0 underwent different molecular manipulations during library preparation than samples from T1 and T2. The NEBNext adaptor from Illumina was diluted 10-fold for samples from T1 and T2, and 50-fold for samples from T0, to prevent the occurrence of dimers. Size selection for samples from T1 and T2 was performed with Agencourt AMPure XP beads (Beckman Coulter, USA), using the volumes selecting for libraries with 500-700 bp inserts. No size selection was applied for samples from T0 due to low DNA concentrations of the libraries. PCR amplification was performed with 15 cycles for samples from T1 and T2, and 18 cycles for samples from T0. Primers used for samples from T1 and T2 were diluted 2-fold. Primers used for samples from T0 were not diluted. Libraries were pooled equimolarily, and 15 pM of the mixture was spiked with 1 % PhiX. Sequencing was carried out on a MiSeq sequencer using a MiSeq Reagent Kit v3 for 600 cycles (Illumina, USA). Raw sequencing data obtained from the MiSeq is available at the sequencing read archive (SRA) under the accession number PRJNA509545.

Bioinformatical analysis of sequencing data

The raw sequencing data was processed as described by Vestergaard et al. [54]. Removal of remnant adaptor sequences, trimming of terminal nucleotides with Phred quality scores less than 15, and removal of reads shorter than 50 bp was carried out using AdapterRemoval [55]. Reads containing more than 1 % ambiguous bases (N) were removed by means of PRINSEQ-lite (version 0.20.4) [56]. DeconSeq (version 0.4.3) [57] was used to remove PhiX contamination. Sufficient coverage of the metagenomic datasets was confirmed by means of Nonpareil (version 2.4) [58] with default settings (Supplementary material 1: Fig. S1).

Metagenomes obtained from bulk soils (T0) comprised reads on average 106 bp shorter than metagenomes created from BSCs (T1 and T2). To test whether the difference in read length affects the accuracy of annotations, T1 and T2 reads were trimmed *in-silico* in a randomized manner to resemble the length distributions of T0 reads. A comparison of the length distributions of exemplary "short" and "long reads" metagenomes, before and after trimming, is presented in Supplementary material 1: Fig. S2. The metagenomes with trimmed sequences were analyzed taxonomically together with the original metagenomes. Principal coordinates analysis (PCoA) ordination plots (Supplementary material 1: Fig. S3) showed that the taxonomic annotations were not notably biased by the difference in read length. Consequently, further analyses were performed on the metagenomes with original read lengths.

For taxonomic classification, metagenomic reads were aligned against the National Center for Biotechnology Information Non-Redundant (NCBI-NR) protein sequences database (January 2017) using Kaiju (version 1.4.4) [59] in Greedy mode with 5 allowed mismatches. Additionally, bacterial 16S rRNA gene sequences were extracted from the metagenomic datasets and annotated using SortMeRNA (version 2.0) [60] with the SILVA SSU database (release 132).

Subsequent functional annotations were performed for bacterial reads identified by Kaiju only. COG (Clusters of Orthologous Groups) functional categories were assigned based on the eggNOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups) database (version 4.5) [61]. Assignment of genes specific for EPS and LPS biosynthesis and excretion, which were the focus of the current study, was carried out according to Cania et al. [37] by hidden Markov model (HMM) searches combined with blasts against protein sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (October 2016). Briefly, HMMs were obtained from the TIGRFAMs (version 15) [62] and Pfam (version 30) [63] databases. FragGeneScan (version 1.19) [64] was used to predict open-reading frames, which were subsequently scanned with HMMER (version 3) (hmmer.org). Matching reads (E-value threshold = 10^{-5}) were mapped to KEGG Orthology (KO) numbers. A KO number was assigned to those reads for which the top 25 blast results were consistent. Blasting was carried out using Diamond (version 0.8.38) [65] with more-sensitive parameters. HMMs and KO numbers used for the analysis are listed in Table 1. Genes *algE*, *epsA* and *epsG* were not included in the analysis due to very low relative abundances (< 5 x 10^{7}). As most reads (> 50 %) assigned to the genes of interest using the HMM-KEGG pipeline were classified into the COG category "Function unknown", this study was based mainly on the targeted approach proposed by Cania et al. [37]. The eggNOG pipeline was employed only for a general overview of the data.

Statistical analysis and data visualization

Analyses of the sequencing data were based on relative abundances of reads. These were obtained by dividing the number of reads assigned to a gene, COG functional category or bacterial family, by the total number of bacterial reads per sample, and multiplying by 100.

Statistical analyses and data visualization were conducted using R (version 3.4.4) [66]. Effects of soil substrate, incubation time, and their possible interaction, were determined according to Field et al. [67].

Briefly, significant differences were detected by a robust 2-way independent analysis of variance (ANOVA) based on the median as M-estimator, with 2000 bootstrap samples. For this purpose, the pbad2way function from the WRS package [68] was used. The influence was counted as significant if the p-value was below 5 % ($p < 0.05$). Benjamini-Hochberg procedure was used to control the false discovery rate in data derived from the metagenomic datasets. Omega squared (ω^2) was calculated as an effect size to estimate the magnitude of observed influences of the analyzed factors. It can be interpreted as the percentage of variation in the dependent variable explained by the independent variable [69].

To detect global differences between samples, principal coordinate analysis (PCoA) ordinations of Bray-Curtis dissimilarity matrices were created using the pcoa function from the ape package [70]. Corrections for negative eigenvalues were performed by means of the Cailliez procedure. Bray-Curtis distances were calculated as an appropriate measure for community abundance data [71] using the vegdist function from the vegan package [72].

Spearman's rank correlation coefficient was used to identify whether the relative abundances of bacterial families and their functional genes were correlated. For this purpose, the function cor.test was used. The correlation was considered to be significant if $p < 0.05$. The average Rho was calculated based on absolute values.

Results

Initial soil substrate parameters

Bulk soils collected from both sites had similar low content of DOC and DON. DOC values were in the range of 4.57 ± 1.67 µg/g in samples collected from Chicken Creek, and 6.63 ± 0.46 µg/g in those from Lieberose, while DON was below detection limit in samples from both sites. Conversely, pH values differed between

294 soils from both locations. Soil from Chicken Creek was slightly alkaline (7.31 ± 0.30), whereas soil from
295 Lieberose was rather acidic (pH = 5.42 ± 0.39). Initial soil substrate parameters are presented in Table 2.

296

297 *BSC development*

298 BSCs developed in the microcosm experiment were in the initial stage of development. They consisted
299 mostly of bacterial and algal biofilms, which enmeshed soil particles and formed patches on the soil
300 surface. Mosses were also observed, but they did not form a dense surface cover yet. For BSCs developed
301 on the Chicken Creek soil, mosses and algae were already visible after the first four months of incubation
302 (T1). For the Lieberose soil, mostly biofilms around single soil particles were visible at T1, whereas mosses
303 and distinct BSC structures appeared after ten months of incubation (T2). Representative pictures are
304 presented in Supplementary material 1: Fig. S4.

305 ANOVA revealed a significant influence of incubation time on DOC ($p < 0.001$, $\omega^2 = 0.70$) and DON ($p <$
306 0.001, $\omega^2 = 0.81$). They accumulated over time and increased by one order of magnitude in BSCs at the
307 end of the experiment compared to the bulk soils at the beginning of the experiment. The water
308 repellency index at T2 was comparable between BSCs grown on soils taken from both locations. It
309 amounted to 1.12 ± 0.15 for BSCs originating from Chicken Creek, and 1.16 ± 0.25 for those from
310 Lieberose. BSC parameters are summarized in Table 2.

311 The exemplary CT images (Supplementary material 1: Fig. S5A-D) of BSCs from T2 grown on soil from
312 Chicken Creek showed a layer of smaller particles in the crust horizon compared to the underlying soil
313 substrate. Positive Euler numbers (Supplementary material 1: Fig. S5E) for both BSCs and the underlying
314 soil indicate more isolated pores than connections in the pore network. The connectivity of the pore space
315 was lower for the BSCs, especially when small pores (46 µm) were considered (Euler number of 8.6 mm^{-3}
316). In the underlying soil, the connectivity was the lowest for pore size class of 107 µm (Euler number of

317 4.6 mm^{-3}). The connectivity then increased towards larger pore sizes as indicated by decreasing Euler
318 numbers.

319

320 *Major characteristics of the shotgun sequencing libraries*

321 Shotgun sequencing of 18 libraries made from bulk soils from the beginning of the experiment (T0) and
322 BSCs from the four-months (T1) and ten-months (T2) samplings generated 18.3 Gbases of data in total.
323 This corresponded to 59,710,640 filtered reads. The number of filtered reads per sample varied between
324 2.1 and 5.3 million. Mean lengths of sequences after trimming ranged from 120 to 250 bp. Details of the
325 raw and filtered sequencing data are summarized in Supplementary material 2: Table S1.

326 The coverage of the microbial diversity by the metagenomic datasets, which was calculated using
327 Nonpareil, varied from 16.5 % to 67.3 % (Supplementary material 1: Fig. S1). As expected, metagenomes
328 from T0 (nonpareil diversity index of 19.24 ± 0.07) had higher coverage (41.9 ± 12.8 %) compared to
329 metagenomes obtained from T1 and T2 (nonpareil diversity index of 20.44 ± 0.31, coverage of 25.5 ± 6.0
330 %).

331

332 *Taxonomic analysis*

333 42.83 % of all metagenomic reads were assigned to *Bacteria*, which could be further differentiated into
334 366 families. Only these reads were further analyzed, as the main focus of this study was on EPS and LPS
335 producers of bacterial origin, and molecular data on other microbial polysaccharide producers in the
336 employed databases is poor. The principal coordinate analysis (PCoA) ordination plot (Fig. 1a) showed
337 that bacterial communities were highly similar at the family level in bulk soils, and underwent
338 differentiation during the development of BSCs. Dominant families were identified by selecting the five

most abundant families from each location at each time point, and sorting them according to their relative abundance of all metagenomes. Relative abundances of the dominant families are shown in Supplementary material 1: Fig. S6. As confirmed by ANOVA, the most characteristic families for T0 were *Burkholderiaceae*, *Comamonadaceae* and *Moraxellacea*e. *Flavobacteriaceae* were also highly abundant at T0, but showed additional differences between the two substrates, and had generally higher relative abundance in samples from Chicken Creek. Similarly, *Sphingomonadaceae* were typical for samples from Chicken Creek, but their relative abundance did not change significantly between the sampled time points. *Streptomycetaceae* had generally higher relative abundance in samples from Lieberose, and occurred mostly at T1 and T2. *Ktedonobacteraceae* and *Acidobacteriaceae* were typical at T1 and T2 for BSCs grown on soil substrate from the sand dune near Lieberose. *Bradyrhizobiaceae* were also characteristic for BSCs originating from Lieberose, but their abundance increased there only at T2. Cyanobacteria, including *Leptolyngbyaceae*, *Tolypothrichaceae* and *Nostocaceae*, were most abundant in BSCs grown on soil substrate from Chicken Creek at T1 and T2, while *Oscillatoriaceae* and *Microcoleaceae* dominated there at T1. Significance levels and ω^2 values are presented in Supplementary material 2: Table S2. Overall, the relative abundances of 13 families were influenced only by location, 125 – only by time, 63 – by both factors, and 130 – by interaction of both factors. The full list of impacted families can be taken from Supplementary material 2: Table S3.

The results of the taxonomic analysis of the whole metagenomic datasets based on the NCBI-NR database were supported by the 16S rRNA gene annotations with SILVA. Although only 0.0062 % of all metagenomic reads were assigned to the bacterial 16S rRNA gene, bacterial community composition did not differ when data from the analysis of the complete metagenomics datasets was compared to the phylogenetic analysis of subsampled 16S rRNA fragments (data not shown).

Functional annotation of metagenomic datasets

General function prediction in the metagenomic datasets was performed by means of the eggNOG database. In total, 73.08 % of bacterial reads were assigned to COG functional categories. The "function unknown" category was most abundant (~ 20 %), followed by "replication, recombination and repair" as well as "amino acid transport and metabolism" (each ~ 6 %). Relatively low abundant (< 0.5 %), but with a special importance to the initiation of BSC formation, were the "cell motility" and "extracellular structures" categories. ANOVA showed that these two categories were more abundant in bulk soils compared to BSCs. COG functional classification is presented in Supplementary material 1: Fig. S7, and significance levels and ω^2 values are listed in Supplementary material 2: Table S4.

Genes specific for the biosynthesis and excretion of alginate, colonic acid, levan and other EPSs as well as LPSs, which were identified using an approach combining HMM searches with blasts against sequences derived from the KEGG database, comprised 0.018 % of bacterial reads (Fig. 2). Key genes, with the overall relative abundance in all metagenomes in the range between 0.002 % and 0.005 %, were *wza*, *wcaB* and *wcaF* of the Wzy-dependent EPS synthesis pathway, and *lptF* and *lptG* of the LPS synthesis pathway. Moderately abundant (≥ 0.001 %) were *kpsE* of the ABC transporter-dependent EPS synthesis pathway, and *wzt* of the LPS synthesis pathway. Genes *wcaK*/*amsJ*, *algJ*, *sacB* and *lptC* were the least abundant (≤ 0.0003 %).

ANOVA revealed that the relative abundances of most investigated genes changed mainly between T0 and T1. However, the differences in the relative numbers of gene copies were also driven by the underlying soil substrate (Supplementary material 2: Table S5). In particular, the genes *wza* and *wcaF* increased at T1, and the increase was more pronounced in samples originating from Chicken Creek compared to those from Lieberose. Moreover, *wzt* increased in BSCs grown on soil substrate from Chicken Creek already at T1, while the increase was observed in BSCs grown on bulk soil taken from Lieberose only

385 at T2. Conversely, *kpsE* and *lptC* decreased at T1. Additionally, *kpsE* was relatively more abundant in
386 samples from Lieberose, whereas *lptG* was dominating in samples from Chicken Creek. Finally, *wcaK*/amsJ,
387 *algJ*, *sacB* and *lptF* were not significantly affected by either incubation time or soil substrate.

388

389 *Investigation of potential EPS/LPS producers*

390 The investigated genes were found in 210 different bacterial families, of which 11 families were found
391 harboring the genes in samples originating from both locations, taken at all three sampling time points
392 (Supplementary material 1: Fig. S8). The number of families harboring genes related to EPS and LPS
393 formation was higher at T1 and T2 compared to T0 (Fig. 3). At T0, the investigated genes were associated
394 with 33 families in bulk soil from Chicken Creek, and in 34 – in bulk soil from Lieberose. These numbers
395 increased at T1 to 150 families in samples originating from Chicken Creek, and 100 – in samples from
396 Lieberose. At T2, 146 families harbored the investigated genes in samples from Chicken Creek, and 87 –
397 in samples from Lieberose.

398 Taxonomy of bacteria potentially capable of synthesis and excretion of EPSs and LPSs is presented in Fig.
399 3 at the level of phylum or class (in case of *Proteobacteria*). At T0, the investigated genes were harbored
400 mainly by *Betaproteobacteria*, whereas at T1 and T2, the major potential producers of EPSs and LPSs were
401 members of *Cyanobacteria*, *Alphaproteobacteria* and *Chloroflexi*. Interestingly, differences were also
402 found in the diversity pattern of potential EPS and LPS producers in response to the different soil
403 substrates. In particular, *Cyanobacteria* were typical for BSCs grown on soil taken from Chicken Creek,
404 while *Chloroflexi* and *Acidobacteria* were characteristic for BSCs originating from Lieberose.

405 ANOVA identified a significant impact on the overall relative abundance of the investigated genes caused
406 by location alone in three families, by time alone – in 14 families, and by interaction of both factors – in
407 23 families. The full list of affected families can be taken from Supplementary material 2: Table S3. The

PCoA plot (Fig. 1b) indicated that the distribution pattern of the analyzed genes among bacterial families resembled that of the total bacterial community (Fig. 1a). In fact, Spearman's rank correlation analysis revealed a positive correlation between the total abundance of a given family and the amount of sequences related to EPS and LPS formation harbored by that family for 57 families of potential EPS and LPS producers (average Rho = 0.69, minimum Rho = 0.47, maximum Rho = 0.97). Three families showed a negative correlation (average Rho = 0.56, minimum Rho = 0.49, maximum Rho = 0.61), and 150 exhibited no correlation (average Rho = 0.23, minimum Rho = 0.00, maximum Rho = 0.46).

Of the 57 families that showed a positive correlation, 25 exceeded an abundance of 1 %, and encompassed altogether 43.26 % of all bacterial reads. Both the relative abundance as well as the potential for EPS and LPS synthesis and export of these families were strongly influenced by both incubation time and underlying soil substrate (Fig. 4). In fact, these factors selected the key producers of EPSs and LPSs already at the phylum level. *Betaproteobacteria* (especially *Burkholderiaceae*), as well as *Gammaproteobacteria* (*Moraxellaceae*) and *Bacteroidetes* (*Flavobacteriaceae*) were prevalent at T0, although most of their members were found also at T1 and T2. *Deltaproteobacteria* (*Myxococcaceae* and *Archangiaceae*) and *Planctomycetes* (*Gemmataceae* and *Planctomycetaceae*) occurred mainly at T1 and T2 in BSCs grown on soil taken from Chicken Creek. However, *Gemmataceae* were relatively abundant also at T2 in BSCs originating from Lieberose. *Cyanobacteria* were characteristic for Chicken Creek samples from T1 and T2, but some of their members (*Oscillatoriaceae* and *Leptolyngbyaceae*) could also be important for EPS and LPS production in Lieberose samples from T1 and T2. Typical for Lieberose samples from T1 and T2 were *Chloroflexi* (*Ktedonobacteraceae* and *Thermogemmatisporaceae*) and *Acidobacteria* (*Acidobacteriaceae*). *Alphaproteobacteria* were prevalent at T1 and T2 in general, but some of their members were more characteristic for one of the underlying substrates (e.g. *Sphingomonadaceae* for soil from Chicken Creek, and *Acetobacteraceae* for that from Lieberose).

Discussion

Bacterial communities of initial soils

In the present study, initial BSCs developed from indigenous communities of free-living microbes, which were highly similar in bulk soils from both sites. As carbon and nitrogen availability are one of the most important factors shaping bacterial community structure [73, 74], their low concentrations could be the primary influence selecting only the best-adapted bacteria in nutrient-poor habitats such as the Chicken Creek catchment and the Lieberose sand dune. In fact, the most abundant bacterial families in the bulk soils from our study were *Burkholderiaceae*, *Comamonadaceae*, *Moraxellaceae* and *Flavobacteriaceae*. These families exhibit oligotrophic traits, as their metabolic versatility and ability to degrade a wide range of compounds, such as various polymers, polycyclic aromatic compounds, phenols and halogenated aromatics, enables them to thrive even in environments with limited nutritional opportunities [75-78]. Consequently, these groups were isolated from habitats such as crude oil, desert soil, glacier ice or distilled water. Furthermore, many members of these families possess fimbriae and exhibit motility. This is in line with the higher amount of corresponding reads found in the bulk soils compared to the initial BSCs. These traits are especially important for free-living bacteria, as they assist in the first steps of cell attachment to a surface and establishment of biofilms [79]. In contrast, genes involved in the formation of EPSs and LPSs, which are particularly relevant in the later stages of biofilm development, were generally more abundant in the initial BSCs compared to the bulk soils.

Influence of initial BSC on soil stability and hydrological properties

EPSs and LPSs have protective functions, bind and mediate penetration of micronutrients into the cell, and function in cell-to-surface and cell-to-cell interactions, which are critical for biofilm development [16, 80, 81]. The prevalence of genes related to EPS and LPS synthesis and export in initial BSCs was therefore

455 expected. EPSs and LPSs also play an important role in improving soil stability, especially in initial BSCs
456 that harbor large amounts of bacteria, like in our study. Bacterial polysaccharides adhere around soil
457 particles, connecting them and cementing into larger aggregates [82]. Several studies demonstrated that
458 bacterial polysaccharides increased the amount of stable soil aggregates [83-85] and reduced rainfall-
459 induced erosion up to 98% [35]. Using the exemplary XCMT images of the ten-months-old samples from
460 Chicken Creek, we also confirmed the ability of initial BSCs to trap surface soil particles. Similar activity of
461 cyanobacterial crusts was captured on XCMT images for example by Raanan et al. [86].Moreover, the
462 increase of the potential for EPS and LPS formation in the initial BSCs compared to the bulk soils was
463 correlated in our study with an accumulation of dissolved organic carbon (data not shown). Altogether,
464 these point to an increased production of adhesive bacterial polysaccharides in our BSCs.

465 Additionally, we measured the influence of BSCs on soil hydrological properties, as the key role in altering
466 soil moisture dynamics seems to be played by polysaccharides [1]. On one hand, they tend to clog pores
467 through swelling, which may reduce soil infiltrability [32, 34, 49, 87]. On the other hand, they can increase
468 soil porosity, which is known to positively affect water penetration [88, 89]. Some researchers also
469 postulate that polysaccharides alter the hydrophobicity of BSC surfaces [90]. In our study, the water
470 repellency of BSCs incubated for ten months on both substrates was close to ideal wettability. Similar
471 water repellency was reported for very young BSCs also in other studies [49, 91]. In initial BSCs, the effect
472 on hydrological processes highly depends on the transient amount and chemical nature of polysaccharides
473 building the bacterial biofilms [92]. For example, water molecules as well as nutrients are bound mainly
474 by the hydrophilic polysaccharide fractions, while the hydrophobic fractions increase the stability of BSCs
475 and their ability to adhere to solid surfaces [93]. Furthermore, polysaccharides in bacterial biofilms are
476 subjected to constant modification and degradation processes, both enzymatic and abiotic [4]. Colica et
477 al. [94] underlined that polysaccharide content cannot be directly correlated with BSC age, as the transient
478 amount of polysaccharides in BSCs depends on the activity of both polysaccharide producers as well as

chemoheterotrophic organisms that use polymeric carbohydrates as a carbon source. Thus, the hydrological properties of BSCs are highly dynamic and may fluctuate during BSC development, as shown previously [49]. Comparing the structure of bacterial communities in BSCs with the composition and chemical properties of bacterial polysaccharides throughout the whole development of BSCs would surely shed more light on this issue. However, more research on the methods of extracting bacterial polysaccharides from BSC needs to be done before such measurements will be reliable and give additional information compared to the repellency index [4, 95].

Genes related to EPS and LPS formation

Although the total relative abundance of genes involved in the formation of EPSs and LPSs increased in the initial BSCs compared to the bulk soils, the individual genes showed different responses. Especially abundant and showing the strongest increase were genes from the Wzy-dependent EPS synthesis pathway and the LPS synthesis pathway. Most bacterial reads in our study belonged to phyla well-known for LPS production, such as *Proteobacteria* (40 %), *Cyanobacteria* (20 %) and *Bacteroidetes* (5 %) [96]. Moreover, recent evidence shows that LPS producers can be found even in phyla that are commonly considered as lacking LPSs [97, 98]. Therefore, the relatively high abundance of genes from the LPS synthesis pathways in our study was expected. Similarly, the relative abundance of genes from the Wzy-dependent pathway was expected, as it is the most widely distributed mechanism of EPS assembly and export [99, 100]. In particular, the *wza* gene encodes for an outer membrane protein Wza, which participates in the translocation across the outer membrane of a variety of EPSs in many different taxa [27]. In comparison, genes belonging to the other pathways of EPS assembly and export (ABC-dependent and synthase-dependent), as well as to the extracellular EPS synthesis, were less abundant in our metagenomes. However, these genes are found only in a limited number of bacteria [101-103].

In contrast to the other investigated genes, the relative abundances of the *kpsE* gene, which is part of the ABC-dependent EPS synthesis pathway, and the *lptC* gene of the LPS synthesis pathway, decreased in the BSCs compared to the bulk soils. The gene *kpsE* is associated with the synthesis of capsular polysaccharides, which enhance survival of bacterial cells in harsh environments [104]. This could explain the high relative abundance of *kpsE* in the low-nutrient bulk soils of the Chicken Creek catchment and the Lieberose sand dune. The LptC protein is part of the LptBFGC LPS export complex together with LptF and LptG. However, unlike LptF and LptG, LptC is not well-conserved among Gram-negative bacteria [105, 106], and may not even be essential for LPS formation [107].

The differences in the relative abundances of genes associated with EPS and LPS formation were observed mainly between the bulk soils and the BSCs. Conversely, very few differences in the relative abundances of the investigated genes were found between samples originating from Chicken Creek and Lieberose.

Differentiation of potential key producers of EPS and LPS during initial development of BSC on different soil substrates

Even though the soil substrate had little impact on the relative abundance of the investigated genes, it shaped the composition of bacterial communities in the developing BSCs. In fact, bacterial communities that were highly similar in the bulk soils underwent differentiation once BSCs started to develop. Furthermore, the taxonomic affiliation of the investigated genes reflected the overall composition of the bacterial communities in our study, and thus the differentiation of the overall bacterial communities was accompanied by the differentiation of the communities of potential producers of adhesive polysaccharides. This is in line with the theory about functional redundancy, which states that important functions are preserved by a community even if the community changes its composition [108]. Our results

524 indicate that the potential to form EPSs and LPSs is an important trait for initial BSCs, as it is maintained
525 despite the different development of bacterial communities on the two investigated substrates.

526 The importance of the potential to produce “soil glue” in the initial stage of BSC development is further
527 underlined by the fact that the highest numbers of sequences related to EPS and LPS biosynthesis were
528 harbored by the families dominating the initial BSCs. The potential key producers of adhesive
529 polysaccharides found in BSCs grown on soil from the Chicken Creek catchment and the Lieberose sand
530 dune were distinct already at the phylum level. In the Chicken Creek BSCs, the most abundant potential
531 producers of EPSs and LPSs were *Cyanobacteria*. They are well known for their capability to form external
532 polysaccharidic layers that enable them to survive in extreme environments [13]. In fact, the genetic
533 machinery of the LPS synthesis as well as the Wzy-dependent pathway of the EPS synthesis were both
534 found in *Cyanobacteria* before [27]. This explains the dominance of these particular polysaccharide
535 biosynthesis pathways in the metagenomes from the Chicken Creek BSCs. However, *Cyanobacteria* played
536 only a minor role in the community of potential EPS and LPS producers in the Lieberose BSCs, possibly
537 because they prefer alkaline environments [1]. In the BSCs grown on the soil from Lieberose,
538 *Cyanobacteria* were replaced by *Chloroflexi* and *Acidobacteria*, which favor acidic habitats [109-113].
539 While *Chloroflexi* lack the ability to synthesize LPSs, *Acidobacteria* are known LPS producers [96].
540 Furthermore, even though the information on the proficiency of both phyla in EPS formation is still limited,
541 sequences related to EPS synthesis were previously found in *Acidobacteria*, and a recent report suggests
542 that some members of this phylum produce large amount of EPSs [114]. *Acidobacteria* and *Chloroflexi* are
543 also members of communities that embed themselves in an EPS matrix, such as biofilms, microbial mats
544 and BSCs [115-117]. The low relative abundance of *Cyanobacteria* in BSCs grown on the soil from
545 Lieberose suggests that, besides *Chloroflexi*, the other major phototrophic organisms there could have
546 been algae, which are also well-known producers of EPSs. Algae dominate acidic soils, and are major
547 components of the natural BSCs found at Lieberose, except for the terminal successional stage that is

548 dominated by mosses and fungi [1, 44, 45, 49]. However, the identification of eukaryotes involved in
549 polysaccharide production is difficult using short-read shotgun sequencing, and would require a different
550 approach [118, 119]. In any case, our results show that potential producers of EPSs and LPSs dominate
551 bacterial communities of BSCs during the initial stage of BSC development. Consequently, the
552 differentiation of overall bacterial communities leads to the emergence of distinct potential key producers
553 of "soil glue".

554 The differentiation of bacterial communities in our study could have been on one hand triggered by soil
555 properties. For example, the two soil substrates used in our study differed in pH, which is one of the most
556 important edaphic parameters determining the composition of bacterial communities in soil [120], but
557 usually signifies that other edaphic parameters (e.g. micronutrient availability) also differ [121]. Therefore,
558 the experimental design of the current study prevents us from making any definite conclusions on the
559 influence of edaphic parameters on the community structure of potential "soil glue" producers. On the
560 other hand, the observed differentiation of bacterial communities could have resulted from various rare
561 species that were too low abundant to detect in the bulk soils, and started dominating during the initial
562 development of BSCs. To identify the main drivers shaping the community composition of potential
563 producers of EPSs and LPSs in initial BSCs, future experiments should involve multiple sterile soil
564 substrates with diverse edaphic parameters, inoculated with the same initial bacterial community.

565

566 **Conclusions**

567 Our study indicates that the potential to produce EPSs and LPSs is an important trait for bacterial
568 communities forming BSCs in the initial stage of BSC development, as (i) the relative abundance of genes
569 related to the biosynthesis of adhesive polysaccharides increases in the bacterial communities of initial
570 BSCs compared to the indigenous bacterial communities of bulk soils, (ii) the relative abundances of EPS

and LPS genes remain similar in initial BSCs with different composition of bacterial communities, and (iii) the highest numbers of sequences related to the "soil glue" production is found in families dominating initial BSCs. Furthermore, we demonstrate that the community composition of potential producers of EPSs and LPSs reflects the overall structure of bacterial communities in initial BSCs, and thus initial BSCs with different bacterial community composition harbor distinct potential key producers of adhesive polysaccharides. Whether the ability of BSCs to improve soil development in the long term is compromised by differences in the efficiency of polysaccharide formation, or the adhesive properties of EPSs and LPSs produced by different taxa, needs further investigation. Similarly, whether the differentiation of bacterial communities during the initial development of BSCs is primarily triggered by soil properties, or results from various rare species present in the initial bacterial community of bulk soil, remains to be determined.

Compliance with Ethical Standards

Conflict of Interest: The authors declare that they have no conflict of interest.

This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Belnap J, Lange OL (2003) Biological soil crusts: structure, function, and management. Springer, New York
2. Sancho LG, Maestre FT, Büdel B (2014) Biological soil crusts in a changing world: introduction to the special issue. Biodivers Conserv 23: 1611–1617.
3. Wu Y, Rao B, Wu P, Liu Y, Li G, Li D (2013) Development of artificially induced biological soil crusts in fields and their effects on top soil. Plant Soil 370: 115-124.
4. Rossi F, Mugnai G, De Philippis R (2018) Complex role of the polymeric matrix in biological soil crusts. Plant Soil 429: 19-34.

5. Totsche KU, Amelung W, Gerzabek MH, Guggenberger G, Klumpp E, Knief C, Lehndorff E, Mikutta R, Peth S, Prechtel A (2018) Microaggregates in soils. J Plant Nutr Soil Sc 181: 104-136.
6. Six J, Bossuyt H, Degryze S, Denef K (2004) A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. Soil Till Res 79: 7-31.
7. Abu-Lail NI, Camesano TA (2003) Role of lipopolysaccharides in the adhesion, retention, and transport of Escherichia coli JM109. Environ Sci Technol 37: 2173-2183.
8. Veste M, Littmann T, Breckle S-W, Yair A (2001) The role of biological soil crusts on desert sand dunes in the northwestern Negev, Israel. In: Breckle, S-W, Veste, M, Wucherer, W (eds.) Sustainable land use in deserts. Springer, Heidelberg, pp. 357-367
9. Selbmann L, Stingele F, Petruccioli M (2003) Exopolysaccharide production by filamentous fungi: the example of Botryosphaeria rhodina. Anton Leeuw Int J G 84: 135-145.
10. Martínez-Cánovas MJ, Quesada E, Martínez-Checa F, del Moral A, Bejar V (2004) Salipiger mucescens gen. nov., sp. nov., a moderately halophilic, exopolysaccharide-producing bacterium isolated from hypersaline soil, belonging to the α-Proteobacteria. Int J Syst Evol Micr 54: 1735-1740.
11. Suela Silva M, Naves Sales A, Teixeira Magalhães-Guedes K, Ribeiro Dias D, Schwan RF (2013) Brazilian Cerrado soil Actinobacteria ecology. Biomed Res Int 2013.
12. Wu N, Zhang Y, Pan H, Zhang J (2010) The role of nonphotosynthetic microbes in the recovery of biological soil crusts in the Gurbantunggut Desert, Northwestern China. Arid Land Res Manag 24: 42-56.
13. Pereira S, Zille A, Micheletti E, Moradas-Ferreira P, De Philippis R, Tamagnini P (2009) Complexity of cyanobacterial exopolysaccharides: composition, structures, inducing factors and putative genes involved in their biosynthesis and assembly. FEMS Microbiol Rev 33: 917-941.
14. Seviour R, Stasinopoulos S, Auer D, Gibbs P (1992) Production of pullulan and other exopolysaccharides by filamentous fungi. Crc Cr Rev Biotechn 12: 279-298.
15. Mahapatra S, Banerjee D (2013) Fungal exopolysaccharide: production, composition and applications. Microbiol Insights 6: 1-16.
16. Suresh Kumar A, Mody K, Jha B (2007) Bacterial exopolysaccharides–a perception. J Basic Microb 47: 103-117.
17. Zaady E, Kuhn U, Wilske B, Sandoval-Soto L, Kesselmeier J (2000) Patterns of CO2 exchange in biological soil crusts of successional age. Soil Biol Biochem 32: 959-966.
18. Fischer T, Gypser S, Subbotina M, Veste M (2014) Synergic hydraulic and nutritional feedback mechanisms control surface patchiness of biological soil crusts on tertiary sands at a post-mining site. J Hydrol Hydromech 62: 293-302.
19. Fischer T, Veste M (2018) Carbon cycling of biological soil crusts mirrors ecological maturity along a Central European inland dune catena. Catena 160: 68-75.
20. West NE (1990) Structure and function of microphytic soil crusts in wildland ecosystems of arid to semi-arid regions. Adv Ecol Res 20: 179-223.
21. Eldridge D, Greene R (1994) Microbiotic soil crusts-a review of their roles in soil and ecological processes in the rangelands of Australia. Soil Res 32: 389-415.
22. Kidron G, Barzilay E, Sachs E (2000) Microclimate control upon sand microbiotic crusts, western Negev Desert, Israel. Geomorphology 36: 1-18.
23. Yair A, Almog R, Veste M (2011) Differential hydrological response of biological topsoil crusts along a rainfall gradient in a sandy arid area: Northern Negev desert, Israel. Catena 87: 326-333.
24. Kidron GJ, Vonshak A (2012) The use of microbiotic crusts as biomarkers for ponding, subsurface flow and soil moisture content and duration. Geoderma 181: 56-64.
25. Belnap J (2006) The potential roles of biological soil crusts in dryland hydrologic cycles. Hydrol Process 20: 3159-3178.

26. Schmid J, Sieber V, Rehm B (2015) Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. Front Microbiol 6: 496.
27. Pereira SB, Mota R, Santos CL, De Philippis R, Tamagnini P (2013) Assembly and export of extracellular polymeric substances (EPS) in cyanobacteria: a phylogenomic approach. Adv Bot Res 65: 235-279.
28. Hunt F (1985) Patterns of LPS synthesis in gram negative bacteria. J Theor Biol 115: 213-219.
29. Wang X, Quinn PJ (2010) Lipopolysaccharide: Biosynthetic pathway and structure modification. Prog Lipid Res 49: 97-107.
30. Whitfield C, Trent MS (2014) Biosynthesis and export of bacterial lipopolysaccharides. Annu Rev Biochem 83: 99-128.
31. Costa OY, Raaijmakers JM, Kuramae EE (2018) Microbial extracellular polymeric substances: ecological function and impact on soil aggregation. Front Microbiol 9: 1-14.
32. Mazor G, Kidron GJ, Vonshak A, Abeliovich A (1996) The role of cyanobacterial exopolysaccharides in structuring desert microbial crusts. FEMS Microbiol Ecol 21: 121-130.
33. HuiXia P, ZhengMing C, XueMei Z, ShuYong M, XiaoLing Q, Fang W (2007) A study on an oligotrophic bacteria and its ecological characteristics in an arid desert area. Sci China Ser D 50: 128-134.
34. Colica G, Li H, Rossi F, Li D, Liu Y, De Philippis R (2014) Microbial secreted exopolysaccharides affect the hydrological behavior of induced biological soil crusts in desert sandy soils. Soil Biol Biochem 68: 62-70.
35. Kheirfam H, Sadeghi SH, Darki BZ, Homaee M (2017) Controlling rainfall-induced soil loss from small experimental plots through inoculation of bacteria and cyanobacteria. Catena 152: 40-46.
36. Mugnai G, Rossi F, Felde VJMNL, Colesie C, Büdel B, Peth S, Kaplan A, De Philippis R (2018) The potential of the cyanobacterium Leptolyngbya ohadii as inoculum for stabilizing bare sandy substrates. Soil Biol Biochem 127: 318-328.
37. Cania B, Vestergaard G, Krauss M, Fliessbach A, Schloter M, Schulz S (2019) A long-term field experiment demonstrates the influence of tillage on the bacterial potential to produce soil structure-stabilizing agents such as exopolysaccharides and lipopolysaccharides. Environ Microbiome 1: 1-14.
38. Gerwin W, Schaaf W, Biemelt D, Fischer A, Winter S, Hüttl RF (2009) The artificial catchment "Chicken Creek"(Lusatia, Germany)—A landscape laboratory for interdisciplinary studies of initial ecosystem development. Ecol Eng 35: 1786-1796.
39. Russell DJ, Hohberg K, Elmer M (2010) Primary colonisation of newly formed soils by actinedid mites. Soil Org 82: 237-251.
40. Zaplata MK, Winter S, Fischer A, Kollmann J, Ulrich W (2012) Species-driven phases and increasing structure in early-successional plant communities. Am Nat 181: E17-E27.
41. Lukešová A (2001) Soil algae in brown coal and lignite post-mining areas in central Europe (Czech Republic and Germany). Restor Ecol 9: 341-350.
42. Gypser S, Herppich WB, Fischer T, Lange P, Veste M (2016) Photosynthetic characteristics and their spatial variance on biological soil crusts covering initial soils of post-mining sites in Lower Lusatia, NE Germany. Flora 220: 103-116.
43. Dümig A, Veste M, Hagedorn F, Fischer T, Lange P, Spröte R, Kögel-Knabner I (2014) Organic matter from biological soil crusts induces the initial formation of sandy temperate soils. Catena 122: 196-208.
44. Fischer T, Veste M, Bens O, Hüttl RF (2012) Dew formation on the surface of biological soil crusts in central European sand ecosystems. Biogeosciences 9: 4621-4628.

45. Fischer T, Veste M, Eisele A, Bens O, Spyra W, Hüttl RF (2012) Small scale spatial heterogeneity of Normalized Difference Vegetation Indices (NDVIs) and hot spots of photosynthesis in biological soil crusts. Flora 207: 159-167.
46. Döhring T, Koefferlein M, Thiel S, Seidlitz HK (1996) Spectral shaping of artificial UV-B irradiation for vegetation stress research. J Plant Physiol 148: 115-119.
47. Thiel S, Döhring T, Köfferlein M, Kosak A, Martin P, Seidlitz HK (1996) A phytotron for plant stress research: how far can artificial lighting compare to natural sunlight? J Plant physiol 148: 456-463.
48. Brankatschk R, Töwe S, Kleineidam K, Schloter M, Zeyer J (2011) Abundances and potential activities of nitrogen cycling microbial communities along a chronosequence of a glacier forefield. ISME J 5: 1025-1037.
49. Fischer T, Veste M, Wiehe W, Lange P (2010) Water repellency and pore clogging at early successional stages of microbiotic crusts on inland dunes, Brandenburg, NE Germany. Catena 80: 47-52.
50. Hallett P, Young I (1999) Changes to water repellence of soil aggregates caused by substrate-induced microbial activity. Eur J Soil Sci 50: 35-40.
51. Urbanek E, Hallett P, Feeney D, Horn R (2007) Water repellency and distribution of hydrophilic and hydrophobic compounds in soil aggregates from different tillage systems. Geoderma 140: 147-155.
52. Köhne JM, Schlüter S, Vogel H-J (2011) Predicting solute transport in structured soil using pore network models. Vadose Zone J 10: 1082-1096.
53. Vogel H-J, Weller U, Schlüter S (2010) Quantification of soil structure based on Minkowski functions. Comput Geosci 36: 1236-1245.
54. Vestergaard G, Schulz S, Schöler A, Schloter M (2017) Making big data smart—how to use metagenomics to understand soil quality. Biol Fert Soils 53: 479-484. doi: 10.1007/s00374-017-1191-3
55. Schubert M, Lindgreen S, Orlando L (2016) AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res Notes 9: 88.
56. Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics 27: 863-864.
57. Schmieder R, Edwards R (2011) Fast identification and removal of sequence contamination from genomic and metagenomic datasets. PloS One 6: e17288.
58. Rodriguez-R LM, Konstantinidis KT (2014) Estimating coverage in metagenomic data sets and why it matters. ISME J 8: 2349-2351.
59. Menzel P, Ng KL, Krogh A (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun 7: 11257.
60. Kopylova E, Noé L, Touzet H (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28: 3211-3217.
61. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M (2015) eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res 44: D286-D293.
62. Haft DH, Selengut JD, Richter RA, Harkins D, Basu MK, Beck E (2013) TIGRFAMs and genome properties in 2013. Nucleic Acids Res 41: D387-D395.
63. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A (2016) The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44: D279-D285.
64. Rho M, Tang H, Ye Y (2010) FragGeneScan: predicting genes in short and error-prone reads. Nucleic Acids Res 38: e191-e191.

65. Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12: 59-60.
66. R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
67. Field A, Miles J, Field Z (2012) Discovering statistics using R. Sage publications, Thousand Oaks, California
68. Wilcox RR, Schönbrodt FD (2014) The WRS package for robust statistics in R. R package version 0.24.
69. Tunks T (1978) The use of omega squared in interpreting statistical significance. B Coun Res Music Ed: 28-34.
70. Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20: 289-290.
71. Legendre P, Legendre LF (2012) Numerical ecology. Elsevier, Amsterdam
72. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2018) vegan: community ecology package. R package version 2.5-1.
73. Sul WJ, Asuming-Brempong S, Wang Q, Tourlousse DM, Penton CR, Deng Y, Rodrigues JL, Adiku SG, Jones JW, Zhou J (2013) Tropical agricultural land management influences on soil microbial communities through its effect on soil organic carbon. Soil Biol Biochem 65: 33-38.
74. Cederlund H, Wessén E, Enwall K, Jones CM, Juhanson J, Pell M, Philippot L, Hallin S (2014) Soil carbon quality and nitrogen fertilization structure bacterial communities with predictable responses of major bacterial phyla. Appl Soil Ecol 84: 62-68.
75. Coenye T (2014) The family Burkholderiaceae. In: Rosenberg, E (ed.) The prokaryotes: Alphaproteobacteria and Betaproteobacteria. Springer, Heidelberg, pp. 759-776
76. Willems A (2014) The family Comamonadaceae. In: Rosenberg, E (ed.) The prokaryotes: Alphaproteobacteria and Betaproteobacteria. Springer, Heidelberg, pp. 777-851
77. Teixeira LM, Merquior VLC (2014) The family Moraxellaceae. In: Rosenberg, E (ed.) The prokaryotes: Gammaproteobacteria. Springer, Heidelberg, pp. 443-476
78. McBride MJ (2014) The family Flavobacteriaceae. In: Rosenberg, E (ed.) The prokaryotes: Other Major Lineages of Bacteria and the Archaea. Springer, Heidelberg, pp. 643-676
79. Vogeleer P, Tremblay YD, Mafu AA, Jacques M, Harel J (2014) Life on the outside: role of biofilms in environmental persistence of Shiga-toxin producing Escherichia coli. Front Microbiol 5: 317.
80. Lindhout T, Lau PCY, Brewer D, Lam JS (2009) Truncation in the core oligosaccharide of lipopolysaccharide affects flagella-mediated motility in Pseudomonas aeruginosa PAO1 via modulation of cell surface attachment. Microbiology+ 155: 3449-3460.
81. Kierek K, Watnick PI (2003) The Vibrio cholerae O139 O-antigen polysaccharide is essential for Ca2+-dependent biofilm development in sea water. P Natl Acad Sci USA 100: 14357-14362.
82. Huang Q, Wu H, Cai P, Fein JB, Chen W (2015) Atomic force microscopy measurements of bacterial adhesion and biofilm formation onto clay-sized particles. Sci Rep-UK 5: 16857.
83. de Caire GZ, De Cano MS, De Mule MZ, Palma R, Colombo K (1997) Exopolysaccharide of Nostoc muscorum (Cyanobacteria) in the aggregation of soil particles. J Appl Phycol 9: 249-253.
84. Carrasco L, Caravaca F, Azcón R, Roldán A (2009) Soil acidity determines the effectiveness of an organic amendment and a native bacterium for increasing soil stabilisation in semiarid mine tailings. Chemosphere 74: 239-244.
85. Rossi F, Li H, Liu Y, De Philippis R (2017) Cyanobacterial inoculation (cyanobacterisation): perspectives for the development of a standardized multifunctional technology for soil fertilization and desertification reversal. Earth-Sci Rev 171: 28-43.

86. Raanan H, Felde VJ, Peth S, Drahorad S, Ionescu D, Eshkol G, Treves H, Felix-Henningsen P, Berkowicz SM, Keren N (2016) Three-dimensional structure and cyanobacterial activity within a desert biological soil crust. Environ Microbiol 18: 372-383.
87. Issa OM, Défarge C, Trichet J, Valentin C, Rajot J-L (2009) Microbiotic soil crusts in the Sahel of Western Niger and their influence on soil porosity and water dynamics. Catena 77: 48-55.
88. Greene R (1992) Soil physical properties of three geomorphic zones in a semi-arid mulga woodland [Acacia aneura]. Aust J Soil Res 30: 55-69.
89. Eldridge DJ (2003) Biological soil crusts and water relations in Australian deserts. In: Belnap, J, Lange, OL (eds.) Biological soil crusts: structure, function, and management. Springer, Berlin pp. 327-337
90. Felde VJMNL, Rossi F, Colesie C, Uteau-Puschmann D, Horne R, Felix-Henningsen P, De Philippis R, Peth S (2016) Pore characteristics in biological soil crusts are independent of extracellular polymeric substances. Soil Biol Biochem 103: 294-299.
91. Pluis J (1994) Algal crust formation in the inland dune area, Laarder Wasmeer, the Netherlands. Vegetatio 113: 41-51.
92. Flemming H-C, Wingender J (2010) The biofilm matrix. Nat Rev Microbiol 8: 623–633.
93. Rossi F, Micheletti E, Bruno L, Adhikary SP, Albertano P, De Philippis R (2012) Characteristics and role of the exocellular polysaccharides produced by five cyanobacteria isolated from phototrophic biofilms growing on stone monuments. Biofouling 28: 215-224.
94. Colica G, Li H, Rossi F, Philippis RD, Liu Y (2015) Differentiation of the characteristics of excreted extracellular polysaccharides reveals the heterogeneous primary succession of induced biological soil crusts. J Appl Phycol 27: 24-32.
95. Redmile-Gordon M, Brookes P, Evershed R, Goulding K, Hirsch P (2014) Measuring the soil-microbial interface: Extraction of extracellular polymeric substances (EPS) from soil biofilms. Soil Biol Biochem 72: 163-171.
96. Lagier J-C, Million M, Hugon P, Armougom F, Raoult D (2012) Human gut microbiota: repertoire and variations. Front Cell Infect Mi 2: 136.
97. Antunes LCS, Poppleton D, Klingl A, Criscuolo A, Dupuy B, Brochier-Armanet C, Beloin C, Gribaldo S (2016) Phylogenomic analysis supports the ancestral presence of LPS-outer membranes in the Firmicutes. Elife 5: e14589.
98. Poppleton DI, Duchateau M, Hourdel V, Matondo M, Flechsler J, Klingl A, Beloin C, Gribaldo S (2017) Outer membrane proteome of veillonella parvula: a diderm firmicute of the human microbiome. Front Microbiol 8: 1215.
99. Whitfield C, Larue K (2008) Stop and go: regulation of chain length in the biosynthesis of bacterial polysaccharides. Nat Struct Mol Biol 15: 121-123.
100. Whitfield C (2010) Polymerases: glycan chain-length control. Nat Chem Biol 6: 403-404.
101. Rosenow C, Esumeh F, Roberts IS, Jann K (1995) Characterization and localization of the KpsE protein of Escherichia coli K5, which is involved in polysaccharide export. J Bacteriol 177: 1137-1143.
102. Bachtiar BM, Coloe PJ, Fry BN (2007) Knockout mutagenesis of the kpsE gene of Campylobacter jejuni 81116 and its involvement in bacterium–host interactions. FEMS Immunol Med Mic 49: 149-154.
103. Muhammadi, Ahmed N (2007) Genetics of bacterial alginate: alginate genes distribution, organization and biosynthesis in bacteria. Curr Genomics 8: 191-202.
104. Rendueles O, Garcia-Garcerà M, Néron B, Touchon M, Rocha EP (2017) Abundance and co-occurrence of extracellular capsules increase environmental breadth: Implications for the emergence of pathogens. PLoS Pathog 13: e1006525.

105. Putker F, Bos MP, Tommassen J (2015) Transport of lipopolysaccharide to the Gram-negative bacterial cell surface. FEMS Microbiol Rev 39: 985-1002.
106. Ruiz N, Gronenberg LS, Kahne D, Silhavy TJ (2008) Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of Escherichia coli. P Natl Acad Sci USA 105: 5537-5542.
107. Benedet M, Falchi FA, Puccio S, Di Benedetto C, Peano C, Polissi A, Dehò G (2016) The lack of the essential LptC protein in the trans-envelope lipopolysaccharide transport machine is circumvented by suppressor mutations in LptF, an inner membrane component of the Escherichia coli transporter. PloS One 11: e0161354.
108. Allison SD, Martiny JBH (2008) Resistance, resilience, and redundancy in microbial communities. P Natl Acad Sci USA 105: 11512-11519.
109. Jones RT, Robeson MS, Lauber CL, Hamady M, Knight R, Fierer N (2009) A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. ISME J 3: 442-453.
110. Lauber CL, Hamady M, Knight R, Fierer N (2009) Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl Environ Microbiol 75: 5111-5120.
111. Wilhelm RC, Niederberger TD, Greer C, Whyte LG (2011) Microbial diversity of active layer and permafrost in an acidic wetland from the Canadian High Arctic. Can J Microbiol 57: 303-315.
112. Santofimia E, González-Toril E, López-Pamo E, Gomariz M, Amils R, Aguilera Á (2013) Microbial diversity and its relationship to physicochemical characteristics of the water in two extreme acidic pit lakes from the Iberian Pyrite Belt (SW Spain). PLoS One 8: e66746.
113. Jones DS, Lapakko KA, Wenz ZJ, Olson MC, Roepke EW, Sadowsky MJ, Novak PJ, Bailey JV (2017) Novel microbial assemblages dominate weathered sulfide-bearing rock from copper-nickel deposits in the Duluth complex, Minnesota, USA. Appl Environ Microbiol 83: e00909-00917.
114. Kielak AM, Castellane TC, Campanharo JC, Colnago LA, Costa OY, Da Silva MLC, Van Veen JA, Lemos EG, Kuramae EE (2017) Characterization of novel Acidobacteria exopolysaccharides with potential industrial and ecological applications. Sci Rep-UK 7: 41193.
115. Rampadarath S, Bandhoa K, Puchooa D, Jeewon R, Bal S (2017) Early bacterial biofilm colonizers in the coastal waters of Mauritius. Electron J Biotechn 29: 13-21.
116. Prieto-Barajas CM, Valencia-Cantero E, Santoyo G (2017) Microbial mat ecosystems: Structure types, functional diversity, and biotechnological application. Electron J Biotechn 31: 48-56.
117. Mogul R, Vaishampayan P, Bashir M, McKay CP, Schubert K, Bornaccorsi R, Gomez E, Tharayil S, Payton G, Capra J (2017) Microbial community and biochemical dynamics of biological soil crusts across a gradient of surface coverage in the central Mojave desert. Front Microbiol 8: 1974.
118. De Vries M, Schöler A, Ertl J, Xu Z, Schloter M (2015) Metagenomic analyses reveal no differences in genes involved in cellulose degradation under different tillage treatments. FEMS Microbiol Ecol 91: fiv069.
119. Wooley JC, Godzik A, Friedberg I (2010) A primer on metagenomics. PLoS Comput Biol 6: 1-13.
120. Fierer N (2017) Embracing the unknown: disentangling the complexities of the soil microbiome. Nat Rev Microbiol 15: 579-590.
121. Lammel DR, Barth G, Ovaskainen O, Cruz LM, Zanatta JA, Ryo M, de Souza EM, Pedrosa FO (2018) Direct and indirect effects of a pH gradient bring insights into the mechanisms driving prokaryotic community structures. Microbiome 6: 106.

877 Table 1. Proteins related to exo- and lipopolysaccharide production with corresponding KO numbers,

878 HMM IDs and genes

Protein	KO number	HMM ID	Gene
polysaccharide export outer membrane protein Wza	K01991	PF02563	*wza*
colanic acid biosynthesis acetyltransferase WcaB	K03819	TIGR04016	*wcaB*
colanic acid biosynthesis acetyltransferase WcaF	K03818	TIGR04008	*wcaF*
colanic acid/amylovoran biosynthesis pyruvyl transferase WcaK/AmsJ	K16710	TIGR04006	*wcaK*/*amsJ*
capsular polysaccharide export system permease KpsE	K10107	TIGR01010	*kpsE*
alginate biosynthesis acetyltransferase AlgJ	K19295	PF16822	*algJ*
levansucrase SacB	K00692	PF02435	*sacB*
lipopolysaccharide transport system ATP-binding protein Wzt	K09691	PF14524	*wzt*
LptBFGC lipopolysaccharide export complex inner membrane protein LptC	K11719	TIGR04409, PF06835	*lptC*
LptBFGC lipopolysaccharide export complex permease LptF	K07091	TIGR04407	*lptF*
LptBFGC lipopolysaccharide export complex permease LptG	K11720	TIGR04408, PF03739	*lptG*

879

880 Table 2. DNA concentration, dissolved organic carbon (DOC), water repellency, dissolved organic nitrogen

881 (DON) and pH values. The mark "-" signifies that the parameter was not measured for the respective

882 samples, while "bdl" stands for "below detection limit"

Location	Time	DNA [ng/g]	DOC [µg/g]	DON [µg/g]	pH	Water repellency
Chicken Creek	T0	1.14 ± 0.29	4.57 ± 1.67	bdl	7.31 ± 0.30	-
	T1	28.89 ± 8.58	36.50 ± 4.31	1.32 ± 0.36	-	nd
	T2	30.95 ± 7.73	48.02 ± 18.06	1.14 ± 0.36	-	1.12 ± 0.15
Lieberose	T0	2.25 ± 1.04	6.63 ± 0.46	bdl	5.42 ± 0.39	-
	T1	5.79 ± 2.14	42.02 ± 6.97	0.83 ± 0.09	-	-
	T2	24.65 ± 4.63	81.03 ± 26.52	1.11 ± 0.15	-	1.16 ± 0.25

883

884

Figure titles and legends

Fig. 1 PCoA plots depicting differences on the family level in a – bacterial community composition, and b – taxonomic affiliation of genes related to EPS and LPS formation. Ellipses drawn around triplicates represent a 95 % confidence level

Fig. 2 Relative abundances of genes specific for the formation of EPSs and LPSs. Error bars show standard deviations

Fig. 3 Comparison of relative abundances of bacteria with and without the potential for EPS and LPS formation (labeled as „Present" and „Absent"). The distinction between the potential producers and non-producers was performed on the level of family. The families were then pooled according to their respective phyla or, in case of *Proteobacteria*, classes. Values above bars represent total numbers of displayed families

Fig. 4 Potential key families of EPS and LPS formation based on their relative abundance and the relative abundance of their genes related to EPS and LPS biosynthesis. Note the different color intensities between bacterial and gene abundances

Supplementary materials

Supplementary material 1: Supplementary figures S1-8

Supplementary material 2: Supplementary tables S1-5

Figure 1

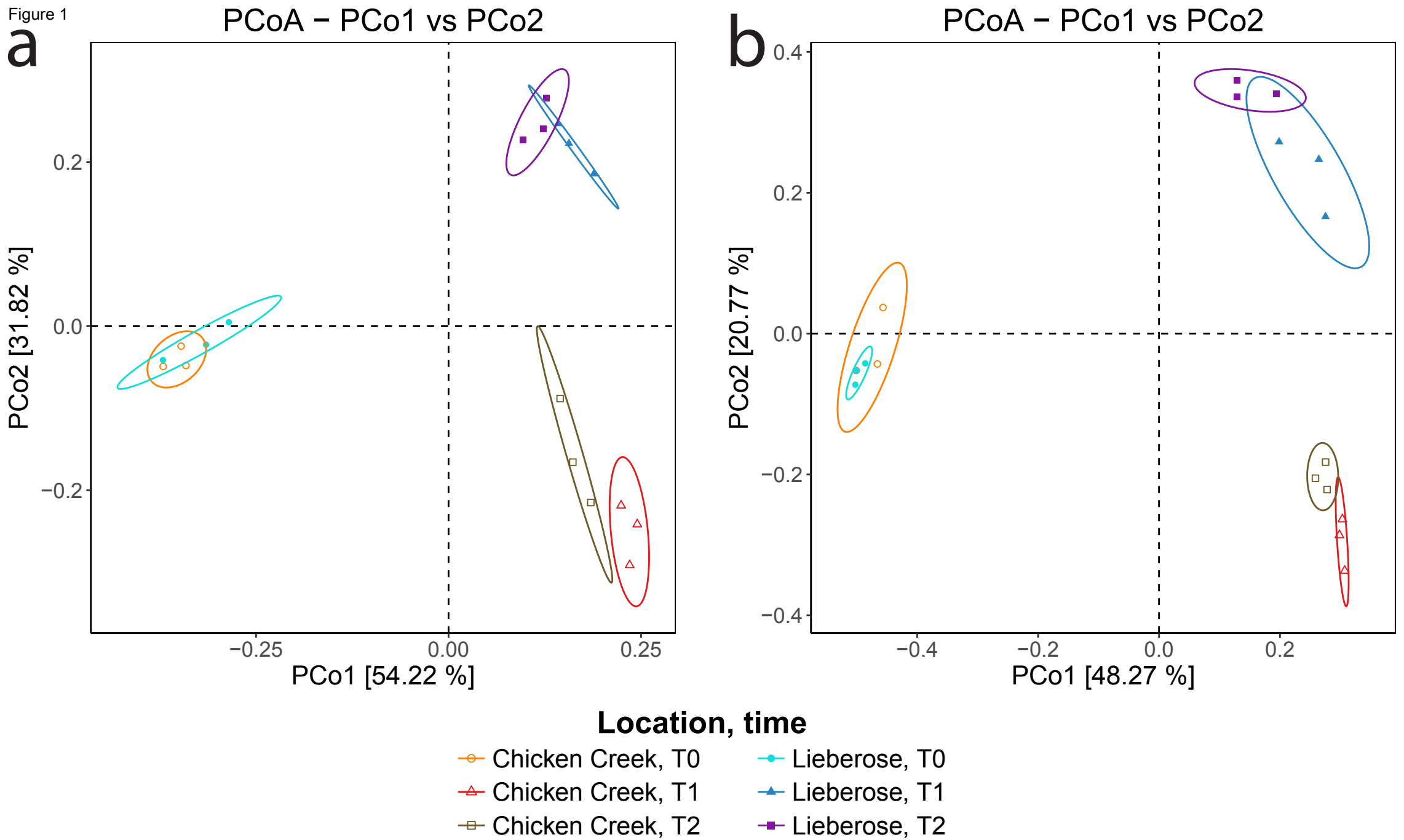

a
PCoA − PCo1 vs PCo2
PCo2 [31.82 %]
PCo1 [54.22 %]
0.2
0.0
−0.2
−0.25
0.00
0.25
b
PCoA − PCo1 vs PCo2
PCo2 [20.77 %]
PCo1 [48.27 %]
0.4
0.2
0.0
−0.2
−0.4
−0.4
−0.2
0.0
0.2
Location, time
Chicken Creek, T0
Chicken Creek, T1
Chicken Creek, T2
Lieberose, T0
Lieberose, T1
Lieberose, T2

Figure 2

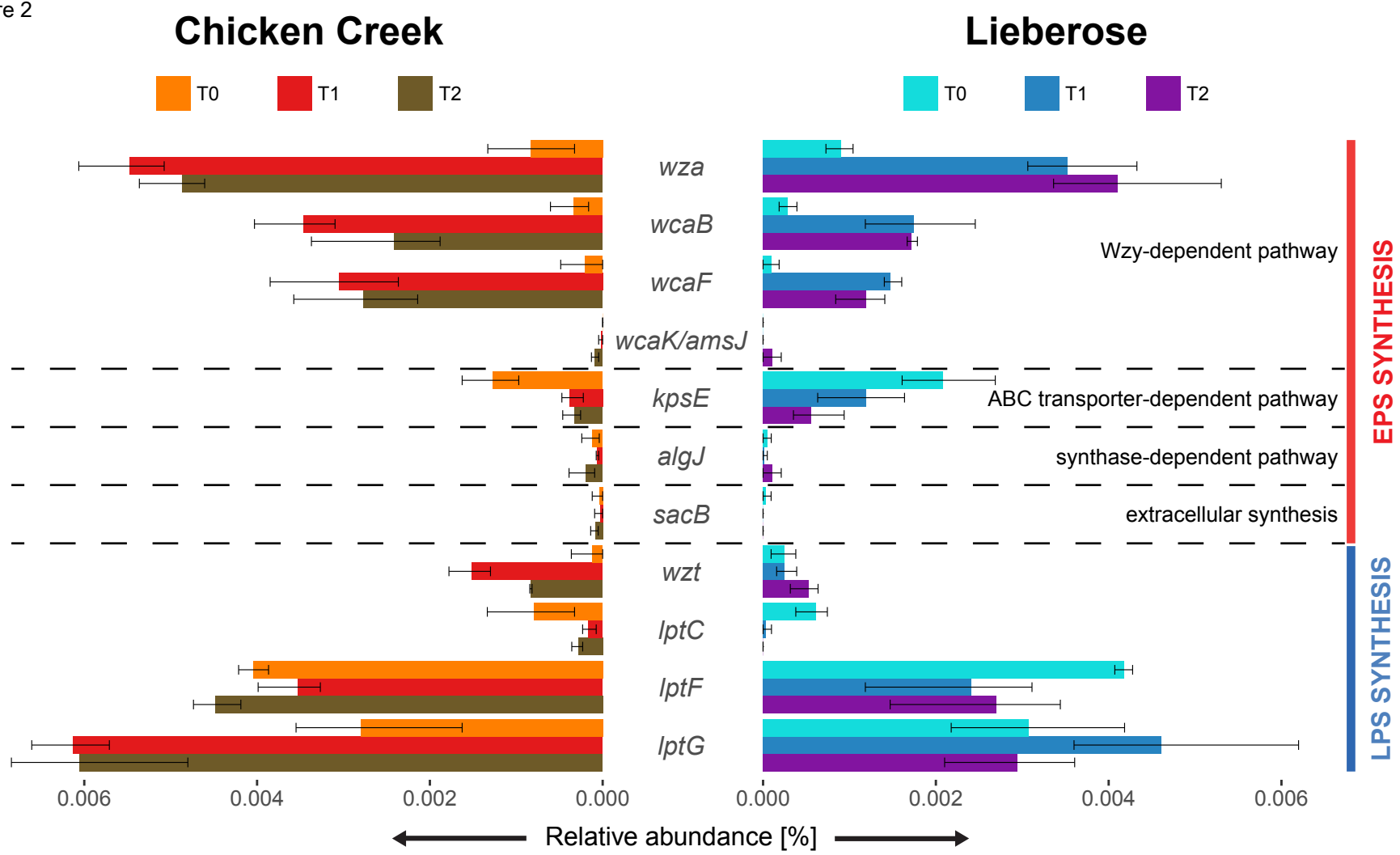
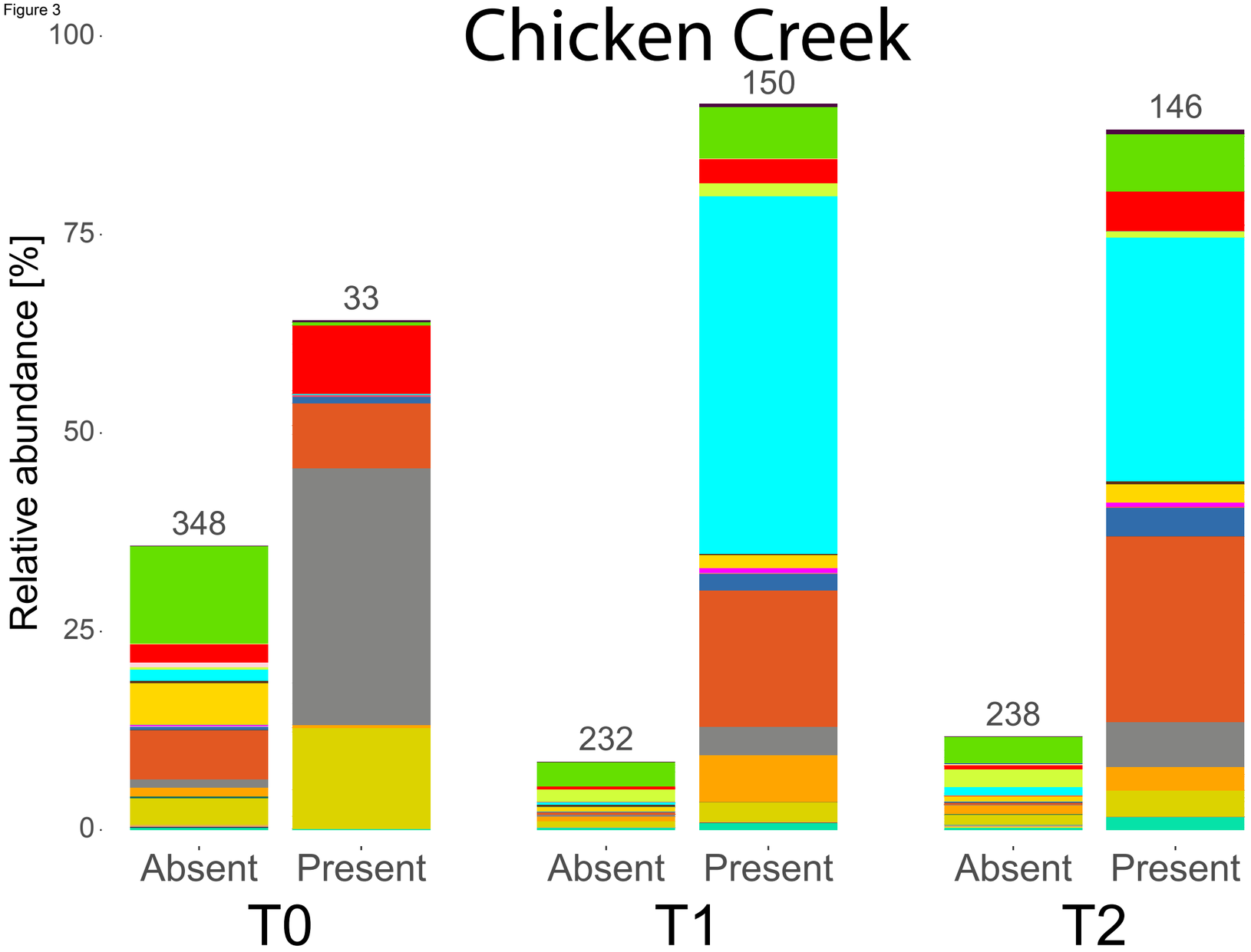
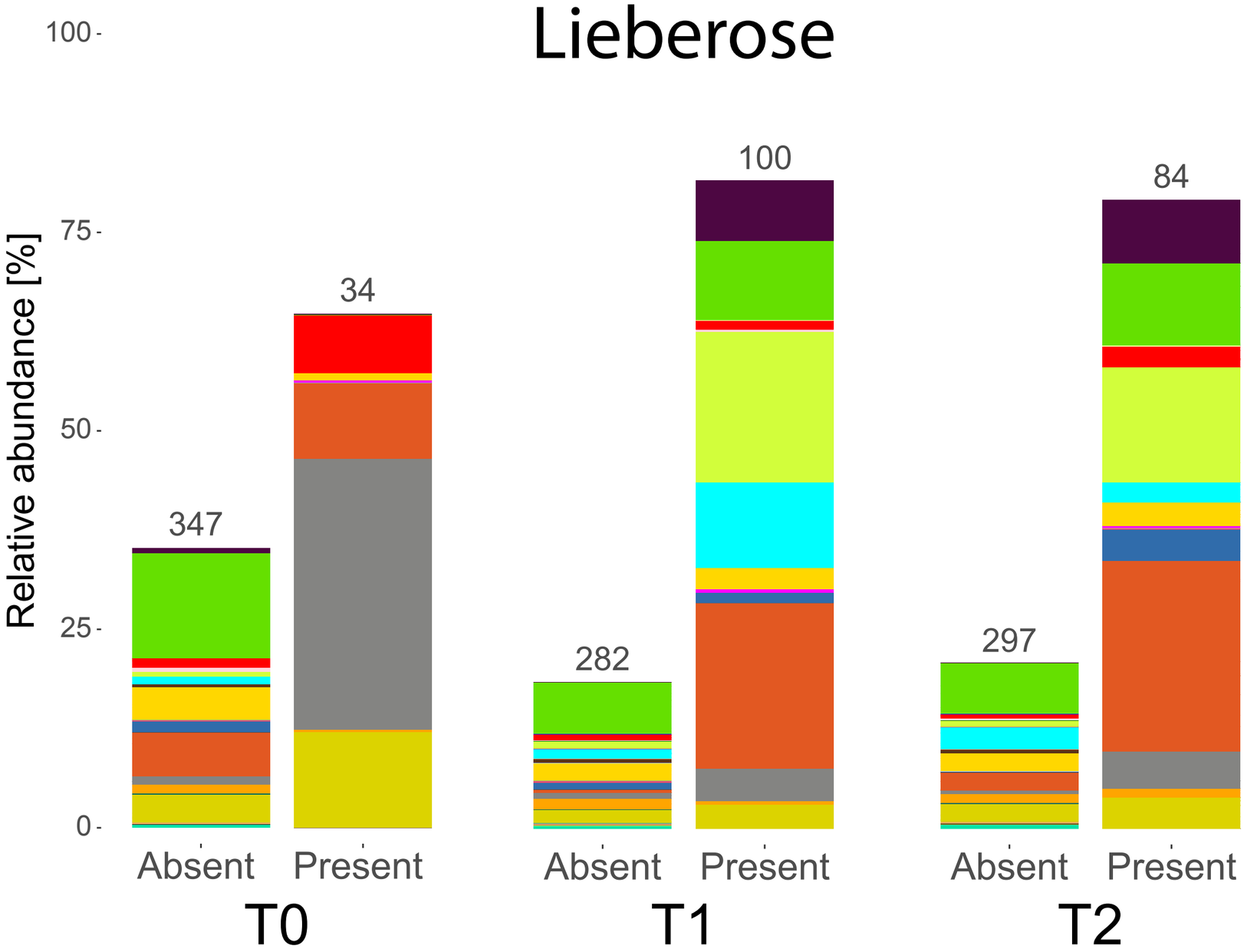

Chicken Creek
Lieberose
T0
T1
T2
T0
T1
T2
wza
wcaB
wcaF
wcaK/amsJ
kpsE
algJ
sacB
wzt
lptC
lptF
lptG
Wzy-dependent pathway
ABC transporter-dependent pathway
synthase-dependent pathway
extracellular synthesis
EPS SYNTHESIS
LPS SYNTHESIS
0.006
0.004
0.002
0.000
0.000
0.002
0.004
0.006
Relative abundance [%]

Figure 3

Chicken Creek



Lieberose



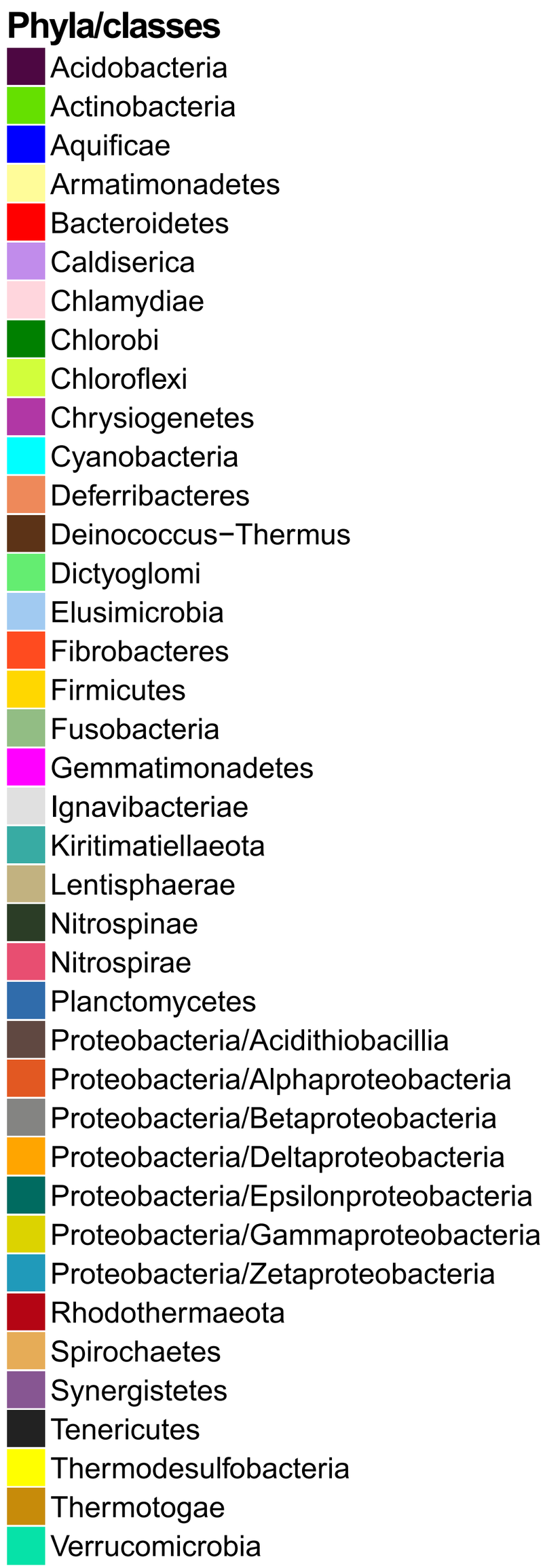
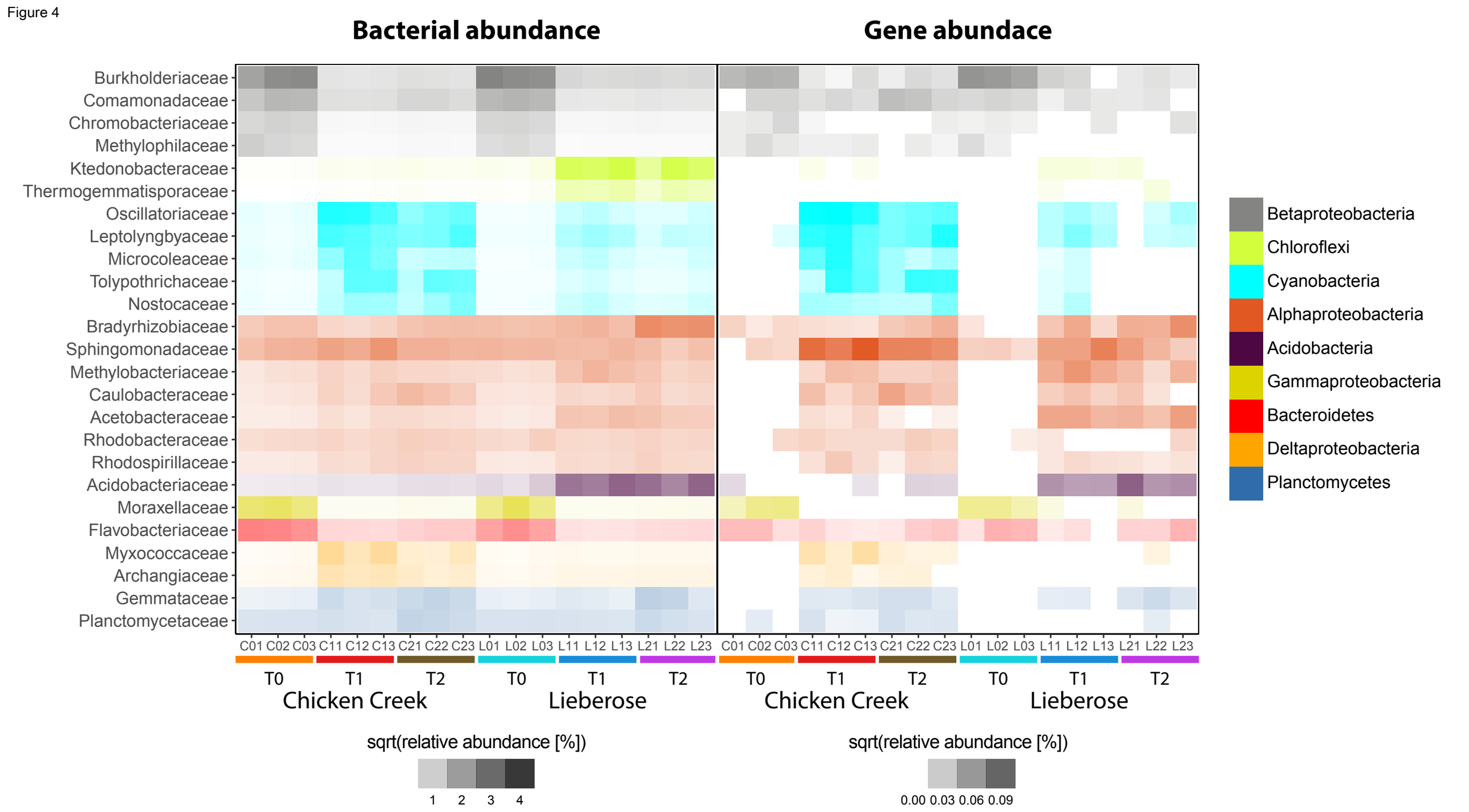


Figure 4


Bacterial abundance
Gene abundace
Burkholderiaceae
Comamonadaceae
Chromobacteriaceae
Methylophilaceae
Ktedonobacteraceae
Thermogemmatisporaceae
Oscillatoriaceae
Leptolyngbyaceae
Microcoleaceae
Tolypothrichaceae
Nostocaceae
Bradyrhizobiaceae
Sphingomonadaceae
Methylobacteriaceae
Caulobacteraceae
Acetobacteraceae
Rhodobacteraceae
Rhodospirillaceae
Acidobacteriaceae
Moraxellaceae
Flavobacteriaceae
Myxococcaceae
Archangiaceae
Gemmataceae
Planctomycetaceae
C01 C02 C03 C11 C12 C13 C21 C22 C23 L01 L02 L03 L11 L12 L13 L21 L22 L23
T0 T1 T2 T0 T1 T2
Chicken Creek
Lieberose
Betaproteobacteria
Chloroflexi
Cyanobacteria
Alphaproteobacteria
Acidobacteria
Gammaproteobacteria
Bacteroidetes
Deltaproteobacteria
Planctomycetes
sqrt(relative abundance [%])
1 2 3 4
0.00 0.03 0.06 0.09