Ecological functions and environmental fate of exopolymers of Acidobacteria

Costa, O.Y.A.

Ecological functions and environmental fate of exopolymers of *Acidobacteria*

Ohana Yonara de Assis Costa

Ecological functions and environmental fate of exopolymers of *Acidobacteria*
The research described in this thesis was performed at the Department of Microbial Ecology of the Netherlands Institute of Ecology - (NIOO-KNAW); O.Y.A. Costa was supported by an SWB grant from CNPq [202496/2015-5] (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brasil).

Cover and layout design by Ohana Yonara de Assis Costa.
Printed by GVO drukkers & vormgevers B.V. ||www.gvo.nl
ISBN: 978-94-6332-635-3

Ecological functions and environmental fate of exopolymers of *Acidobacteria*

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 9 juli 2020
klokke 16.15 uur

door

Ohana Yonara de Assis Costa

geboren te Manaus, Brazil

in 1990

Promotor :	**Prof. dr. J. M. Raaijmakers** Netherlands Institute of Ecology, Wageningen Leiden University
Copromotor:	**Prof. dr. E. E. Kuramae** Netherlands Institute of Ecology, Wageningen Utrecht University

Promotiecommissie:	**Prof. dr. G. P. van Wezel (voorzitter)** Leiden University
	Prof. dr. P. G. L. Klinkhamer (secretaris) Leiden University
	Prof. dr. ir. T. M. Bezemer Leiden University
	Dr. ir. M. M. Hefting Utrecht University
	Prof. dr. G. A. Kowalchuk Utrecht University
	Dr. K. Faust KU Leuven - University of Leuven

Contents

Chapter 1

General Introduction and Thesis Outline

Ohana Y. A. Costa & Eiko E. Kuramae

Adapted from: Kuramae EE and Costa OYA (2019). *Acidobacteria*. In: Schmidt, Thomas M. (ed.) **Encyclopedia of Microbiology**, 4th Edition. vol. 1, pp. 1-8. UK: Elsevier

1. Phylum *Acidobacteria*

Acidobacteria is a ubiquitous and abundant bacterial phylum in soil, but the factors underlying their ecological prevalence in the soil ecosystem remain unclear. This lack of fundamental knowledge is largely due to difficulties to isolate *Acidobacteria* and their slow growth *in vitro* (Kielak *et al.*, 2016). However, non-culturable approaches, mainly 16S rRNA-based sequence surveys, have revealed that *Acidobacteria* are metabolically diverse and widely distributed. Most *Acidobacteria* appear to be aerobes, but some can grow under reduced oxygen conditions (1%–2% O_2) (Eichorst *et al.*, 2018). The diversity and abundance of *Acidobacteria* have been reported in a variety of sites, such as diverse agricultural (Navarrete *et al.*, 2013) and contaminated soils (Wang *et al.*, 2016), sediments (Liao *et al.*, 2019), forest soils (Štursová *et al.*, 2012), peatland (Pankratov *et al.*, 2008), various water systems (Izumi *et al.*, 2012, López-López *et al.*, 2015), acid mine drainage (Wegner & Liesack, 2017) and surfaces of Paleolithic caves and catacombs (Schabereiter-Gurtner *et al.*, 2002, Zimmermann *et al.*, 2005). The few sequenced genomes of *Acidobacteria* indicate a broad substrate range of ABC transporters for nutrient uptake, suggesting an advantage of *Acidobacteria* in complex environments and adaptation to oligotrophic conditions, such as nutrient-limited soil conditions (Kielak *et al.*, 2016).

Acidobacteria form as much as 50% of the total soil bacterial community based on 16S rRNA gene phylogenetic sequence surveys (Pereira de Castro *et al.*, 2016) and compose on average 20% of the total microbial community in soils around the world (Janssen, 2006). Three subdivisions are particularly abundant in soils: class *Acidobacteriia* (former subdivision 1), class *Blastocatellia* (former subdivision 4) and subdivision 6. At present, members of class *Acidobacteriia* are the most readily culturable under laboratory conditions. Together with subdivision 3, *Acidobacteriia* are the most abundant groups in soils (Barns *et al.*, 1999, Janssen, 2006).

The existence of the phylum *Acidobacteria* was first recognized in 16S rRNA gene sequence-based studies, which revealed that *Acidobacteria* and *Proteobacteria* were the predominant phyla in diverse soil environments (Kielak *et al.*, 2016). It was predicted that *Acidobacteria* would be as diverse as the widely studied phylum *Proteobacteria* (Hugenholtz *et al.*, 1998). Only 4–5 subdivisions were initially described in 1997 (Kuske *et al.*, 1997, Ludwig *et al.*, 1997), a number that increased to 8 subdivisions in 1998 (Hugenholtz *et al.*, 1998) and 11 in 2005 (Zimmermann *et al.*, 2005). The diversity and phylogeny of *Acidobacteria* currently encompasses 26 known subdivisions (Figure 1) belonging to eleven described families: *Acidobacteriaceae*, *Bryobacteraceae* (within class *Acidobacteriia*), *Blastocatellaceae*, *Pyrinomonadaceae, Arenimicrobiaceae* (within class *Blastocatellia*), *Acanthopleuribacteraceae*, *Holophagaceae, Thermotomaculaceae* (within class *Holophagae*), *Vicinamibacteraceae* (within class *Vicinamibacteria*) and *Thermoanaerobaculales* (within class *Thermoanaerobaculia*).

Despite their high abundance in several environments, only 66 *Acidobacteria* species have been described so far. The first isolate belonging to this phylum was *Acidobacterium*

capsulatum, from which the name of the phylum was derived; the genus *Acidobacterium* was first proposed in 1991 for acidophilic, chemoorganotrophic bacteria isolated from an acidic mineral environment (Kishimoto *et al.*, 1991).

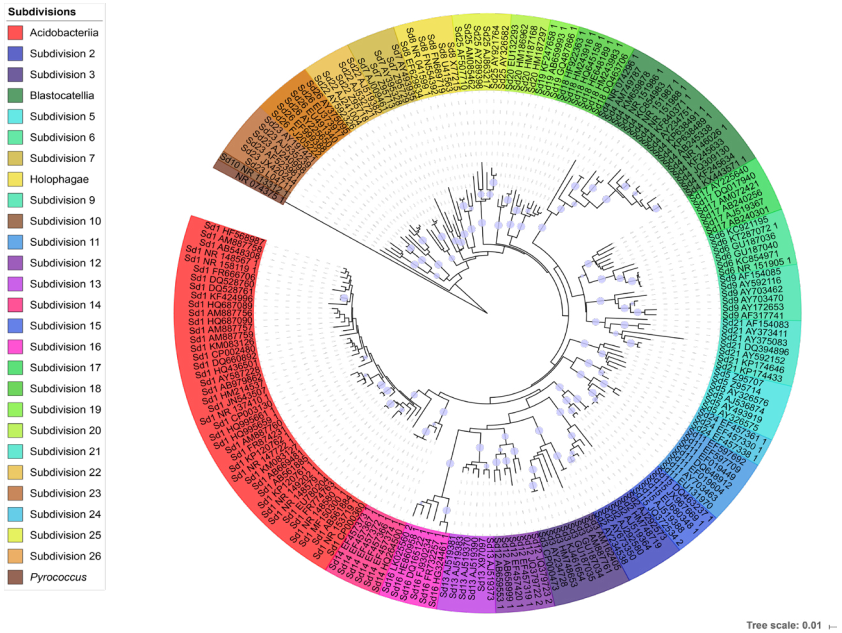

Figure 1: Dendogram of *Acidobacteria* subdivisions. The dendrogram was constructed using 16S rRNA gene sequences from 26 *Acidobacteria* subdivisions downloaded from the RDP 11 and NCBI databases. The sequences were aligned using the tool align.seqs in the software Mothur (Schloss et al., 2009) against Silva database version 132 (Quast et al., 2012). The software Mega 7 (Kumar et al., 2016) was used to build the dendrogram based on the neighbour-joining method with 1000 bootstraps. The circles represent bootstrap values above 0.75. Archaeon *Pyrococcus furiosus* was used as an outgroup.

Most *Acidobacteria* species belong to two classes: *Acidobacteriia* (46 species) and *Blastocatellia* (13 species), while four species belong to *Holophagae*, two to class *Vicinamibacteria*, and one species to class *Thermoanaerobaculia*. Two *Acidobacteria* isolates belong to *Candidatus* genera: '*Ca.* Koribacter', '*Ca.* Solibacter' (*Acidobacteriia*) and '*Ca.* Chloracidobacterium' (*Blastocatellia*). In 2018, three new *Candidatus* genera with features of dissimilatory sulfur metabolism were proposed based on metagenome-assembled genomes: '*Ca*. Sulfotelmatobacter', '*Ca*. Sulfotelmatomonas' and 'Ca. *Sulfopaludibacter*' (Class *Acidobacteriia*) (Hausmann *et al.*, 2018).

Recently, Dedysh and Yilmaz (2018) proposed a refinement to taxonomy of *Acidobacteria.* Based on 16S rRNA gene sequences, the authors distributed the 26 subdivisions to 15 class-level divisions, from which only 5 contain described members. Class *Acidobacteriia* contain subdivisions 1, 2, 3, 5, 11, 12, 13, 14, 15 and 24; class *Blastocatellia* contains subdivision 4; class *Vicinamibacteria* contains subdivisions 6, 9 and 17; class *Holophagae* consists of subdivisions

8 and 22, and class *Thermoanaerobaculia* contains subdivision 23.

The genus *Acidobacterium* (Kishimoto *et al.*, 1991) belongs to the family *Acidobacteriaceae* (subdivision 1), which also contains the genera *Edaphobacter* (Koch *et al.*, 2008), *Terriglobus* (Eichorst *et al.*, 2007), *Acidicapsa* (Kulichevskaya *et al.*, 2012), *Acidipila* (Okamura *et al.*, 2011), *Bryocella* (Dedysh *et al.*, 2012), *Granulicella* (Pankratov & Dedysh, 2010), *Occallatibacter* (Foesel *et al.*, 2016), *Telmatobacter* (Pankratov *et al.*, 2012), *Terracidiphilus* (García-Fraile *et al.*, 2016), *Silvibacterium* (Lladó *et al.*, 2016) and '*Candidatus* Koribacter' (Ward *et al.*, 2009). These bacteria are gram-negative chemoorganothrophs, with prevalent capsule formation and variable motility. They are aerobic or facultatively anaerobic and mostly mesophiles, although some are cold adapted. The members of this family use sugars as favorite source of carbon and energy and are able to degrade complex carbohydrates. Their genomic G+C content varies from 51.7 to 62.1% (Thrash & Coates, 2014, Foesel *et al.*, 2016).

The family *Bryobacteraceae* (subdivision 3) is formed by the genera *Bryobacter* (Kulichevskaya *et al.*, 2010), *Paludibaculum* (Kulichevskaya *et al.*, 2014) and '*Candidatus* Solibacter' (Ward *et al.*, 2009). These bacteria are chemoheterothrophic, gram negative, non-spore forming rods that are aerobes and facultative anaerobes and can use various sugars as growth substrates. In addition, members are mildly acidophilic, mesophilic and psychrotolerant. Their genomic G+C content varies from 55.5 to 61.9% (Dedysh *et al.*, 2017).

The family *Blastocatellaceae* (subdivision 4) contains the genera *Blastocatella* (Foesel *et al.*, 2013), *Aridibacter* (Huber *et al.*, 2014), *Tellurimicrobium* and *Stenotrophobacter* (Pascual *et al.*, 2015). The members of this family are gram-negative, non-spore forming, non-capsule forming bacteria. In addition, these aerobic bacteria are unable to reduce nitrate or ferment glucose and are slightly acidophilic to neutrophilic mesophiles, with a preference for complex proteinaceous growth substrates, although a few complex carbohydrates can be used. Their genomic G+C content ranges from 46.5% to 59.4%. (Pascual *et al.*, 2015).

The family *Pyrinomonadaceae* (subdivision 4) is composed of the genus *Pyrinomonas* (Crowe *et al.*, 2013). Members of this genus are gram-negative non-spore-forming, non-capsule-forming aerobic chemoheterothrophs that are unable to grow phototrophically, reduce nitrate or ferment glucose. They are thermophiles and mildly acidophiles. Furthermore, they prefer complex proteinaceous growth substrates and have a variable capability to hydrolyze polymers. The genomic G+C of the type strain *Pyrinomonas methylaliphatogenes* is 59.6 %(Crowe *et al.*, 2013).

The family *Arenimicrobiaceae* (subdivision 4) is formed by the genera *Brevitalea* and *Arenimicrobium* (Wüst *et al.*, 2016). These bacteria are gram-negative non-spore-forming and non-capsule forming rods. They are aerobic chemoorganoheterotrophs that prefer proteinaceous growth substrates and are not capable of nitrate reduction and glucose fermentation. Additionally, they are mesophiles with a wide pH tolerance range (Dedysh & Yilmaz, 2018). Their genomic G+C content varies from 54.7% to 66.9% (Wüst *et al.*, 2016).

Family *Acanthopleuribacteraceae* (subdivision 8) contains only the genus *Acanthopleuribacter*.

Cells belonging to this genus are gram-negative, motile, strictly aerobic rods that are able to use α-D-glucose, L-alanine, hydroxy-L-proline, L-serine, L-threonine, inosine, uridine and thymidine for growth. The genomic G+C content of the type species, *Acanthopleuribacter pedis* is 56.7% (Fukunaga *et al.*, 2008).

Family *Vicinamibacteraceae* (subdivision 6) contains the genera *Vicinamibacter* (Huber *et al.*, 2016) and *Luteitalea* (Vieira *et al.*, 2017). These bacteria are gram-negative, non-spore forming, aerobic chemoorganoheterotrophs that are capable of growth on organic/nucleic acids and simple sugars but prefer complex proteinaceous compounds. They are neutrophils that tolerate a wide range of pH and can be from psychrotolerant to mesophiles. Genomic G+C content varies from 64.7 to 65.9% (Huber & Overmann, 2018).

Family *Holophagaceae* (subdivision 8) is formed by genera *Holophaga* (Liesack *et al.*, 1994) and *Geothrix* (Coates *et al.*, 1999). Both are strict anaerobes chemoorganothrophs, non-spore-forming, gram negative, mesophile, neutrophilic, and non-motile. The genomic DNA G+C content of the type species *Holophaga foetida* is 62.5 % (Liesack *et al.*, 1994).

Family *Thermotomaculaceae* (subdivision 10) contains only the genus *Thermotomaculum,* a gram-negative, non-spore forming anaerobic heterotrophic thermophile that was isolated from a deep-sea hydrothermal vent. The genomic DNA G+C content of the type species *Thermotomaculum hydrothermale* is 51.6% (Izumi *et al.*, 2012).

The family *Thermoanaerobaculales* (subdivision 23) accommodates the genus *Thermoanaerobaculum,* a strictly anaerobic, thermophilic and chemo-organotrophic genus that was isolated from a freshwater hot spring. The genomic DNA G+C content of the type species *Thermoanaerobaculum aquaticum* is 62.7% (Losey *et al.*, 2013).

Last, *Chloracidobacterium* genus is not assigned to any currently described *Acidobacteria* family. These bacteria are thermophilic, anoxygenic, chlorophototrophic members of class *Blastocatellia* isolated from a hot spring (Tank & Bryant, 2015). The genus potentially represents a novel family and novel order, but studies addressing the taxonomy status of *Chloracidobacterium* are still ongoing (Dedysh & Yilmaz, 2018). In addition, genomes assigned to candidate phyla '*Candidatus* Aminicenantes' and '*Candidatus* Fischerbacteria' might belong to *Acidobacteria* (Dedysh & Yilmaz, 2018).

The number of isolates and described genera of *Acidobacteria* has gradually increased due insights into the metabolism of these bacteria provided by genomic and metagenomics studies, as well as through improvement of cultivation methods (Pascual *et al.*, 2015). Currently, in 2020, there are 16,286 *Acidobacteria* 16S rRNA gene sequences in RDP 11 Database (Cole *et al.*, 2014) and 51 complete genomes (42 *Acidobacteriia*, 3 *Blastocatellia*, 3 *Holophagae*, 1 *Thermoanaerobaculia*, 1 *Vicinamibacteria* and 1 unclassified *Acidobacteria*) in NCBI database (Table 1) (NCBI Resource Coordinators, 2016). Further efforts to unravel the metabolism of uncultured microorganisms through modern technologies, such as high-throughput sequencing and Nanoscale Secondary Ion Mass Spectrometry (NanoSIMS), may provide new insights into how to cultivate novel genera, thereby increasing the present knowledge on the

characteristics and potential functions of members of the phylum *Acidobacteria*.

Table 1: *Acidobacteria* complete genomes listed in NCBI*.

Genome	Accession	Size (Mb)
Acidobacteriia		
Acidipila dinghuensis str. DHOF10	NZ_SDMK00000000	5.1
Acidipila rosea str. DSM 103428	NZ_SMGK00000000	4.2
Acidipila sp. str. 4G-K13	NZ_QVQT00000000	5.0
Acidipila sp. str. EB88	NZ_QWEV00000000	4.5
Acidisarcina polymorpha str. SBC82	-	7.6
Acidobacteria bacterium str. KBS 146	NZ_JHVA00000000	5.0
Acidobacteriaceae bacterium str. KBS 83	-	12.5
Acidobacteriaceae bacterium str. KBS 89	-	12.0
Acidobacteriaceae bacterium str. KBS 96	-	13.4
Acidobacteriaceae bacterium str. TAA166	-	12.3
Acidobacteriaceae bacterium str. URHE0068	-	2.2
Acidobacteriia bacterium str. SbA2	NZ_OKRG00000000	2.7
Acidobacterium ailaaui str: PMMR2	NZ_JIAL00000000	3.7
Acidobacterium capsulatum str. ATCC 51196	-	8.3
Bryobacter aggregatus str. MPL3	NZ_JNIF00000000	5.7
Bryocella elongata str. DSM 22489	NZ_FNVA00000000	5.7
Candidatus Koribacter versatilis str. Ellin345	-	11.3
Candidatus Solibacter usitatus str. Ellin6076	-	19.9
Candidatus Sulfopaludibacter sp. str. SbA3	NZ_OKRF00000000	8.5
Candidatus Sulfopaludibacter sp. str. SbA4	NZ_OMOG00000000	10.0
Candidatus Sulfopaludibacter sp. str. SbA6	NZ_OKRH00000000	3.5
Candidatus Sulfotelmatobacter kueseliae str. SbA1	NZ_OMOD00000000	5.4
Candidatus Sulfotelmatobacter sp. str. SbA7	NZ_OKRE00000000	2.8
Candidatus Sulfotelmatomonas gaucii str. SbA5	NZ_OKRB00000000	5.3
Edaphobacter aggregans str. EB153	NZ_RSDW00000000	0.6
Edaphobacter aggregans str. DSM 19364	NZ_JQKI00000000	0.9
Edaphobacter dinghuensis str. EB95	NZ_RBIF00000000	4.5
Edaphobacter modestus str.DSM 18101	NZ_SHKW00000000	7.4
Granulicella mallensis str. MP5ACTX8	-	12.5
Granulicella pectinivorans str. DSM 21001	NZ_FOZL00000000	5.3
Granulicella rosea str. DSM 18704	NZ_FZOU00000000	5.3
Granulicella sibirica str.AF10	NZ_RDSM00000000	6.1
Granulicella sp. GAS466 str. GAS466	NZ_RJKT00000000	6.2
Granulicella tundricola str. MP5ACTX9	-	11.0
Occallatibacter savannae str. AB23	NZ_QFFY00000000	6.3
Silvibacterium bohemicum str. S15	NZ_LBHJ00000000	6.5
Terracidiphilus gabretensis str. S55	NZ_LAIJ00000000	5.3
Terriglobus albidus str. ORNL	NZ_CP042806	6.4
Terriglobus roseus str. GAS232	-	9.7
Terriglobus roseus str. DSM 18391	-	10.5
Terriglobus saanensis str. SP1PR4	-	10.2
Terriglobus sp. str. TAA 43	NZ_JUGR00000000	5.0
Blastocatellia		
Chloracidobacterium thermophilum str. OC1	NZ_LMXM00000000	3.6
Chloracidobacterium thermophilum B str. B	-	7.4
Pyrinomonas methylaliphatogenes str. K22	NZ_CBXV000000000	3.8
Holophagae		
Geothrix fermentans str. DSM 14018	-	2.0
Holophaga foetida str. DSM 6591	NZ_AGSB00000000	4.2
Holophagae bacterium str. FeB_10	NZ_PQAJ00000000	4.2
Thermoanaerobaculia		
Thermoanaerobaculum aquaticum str. MP-01	-	5.3
Vicinamibacteria		
Luteitalea pratensis str. DSM 100886	NZ_CP015136	7.5
Unclassified Acidobacteria		
Acidobacteria bacterium AB60	NZ_VANK00000000	6.7

*(https://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=57723, January 2020).

2. Carbohydrate metabolism

Carbon usage is one of the physiological requirements in *Acidobacteria* that has been widely studied. Genomic analyses demonstrated the presence of 131 glycoside hydrolase (GH) families across 24 *Acidobacteria* genomes, including important enzymes for plant cell wall breakdown (Eichorst *et al.*, 2018). Overall, *Acidobacteria* are able to use D-glucose, D-xylose, lactose, maltose, cellobiose, glucose and xylose as carbon sources and can degrade simple and polymeric carbohydrates. The ability to use glucose and xylose is evident, since those are the main carbon sources employed for acidobacterial isolation (Kielak *et al.*, 2016). However, most of subdivision 1 *Acidobacteria* are not able to use fucose or sorbose, sugars rarely observed in plant cell wall and soil (Li *et al.*, 2013, Kielak *et al.*, 2016). Genes related to the biosynthesis, transfer, breakdown and/or modification of carbohydrates typically represent 5%–9% of acidobacterial genomes. The genomes of the non-soil isolates *G. fermentans*, *H. foetida* and *C. thermophilum* B have the lowest percentages of genes related to carbohydrate-active enzymes, which indicates that soil *Acidobacteria* might have a higher proportion of their genomes involved in carbohydrate metabolism (Eichorst *et al.*, 2018). Moreover, members of subdivision 1 (*Acidobacteriia*) have a broader glycolytic capability than other subdivisions (Kielak *et al.*, 2016).

The genomes of *Acidobacteria* include genes encoding pathways for the degradation of various polysaccharides (starch, cellulose, hemicellulose, laminarin, xylan, xyloglucan, and gellan gum), but experimental evidence for hydrolytic capabilities not always support genomic predictions, which could be due to errors in gene annotation, variations in gene regulation or culture conditions (Kielak *et al.*, 2016, Belova *et al.*, 2018).

The GH families related to polymeric carbohydrate degradation with the highest percentages across sequenced acidobacterial genomes are GH109 and GH74 (Eichorst *et al.*, 2018). GH109 contais α-N-acetylgalactosaminidases that act on O-linked oligosaccharides, which are typically found in chitin, bacterial peptidoglycan and lipopolysaccharides (Liu *et al.*, 2007), while GH74 contains endoglucanases that act on β-1,4-linked glucans (Lombard *et al.*, 2014). Families involved in cellulose degradation, such as GH5 were observed in subdivisions 1, 3, 4 and 6, while families GH8, GH9, GH44 and GH12 were found in a few subdivision 1 and 3 genomes (*Terriglobus*. sp., '*Ca*. K. versatilis', *T*. *gabretensis*, and *G*. *mallensis* and 'Ca. S. usitatus'). Family GH3 β-glucosidases were detected in all sequenced genomes, while GH18 and GH19 putative chitinases were present in genomes belonging to subdivision 1, 3, 4, 6, and 8 (Eichorst *et al.*, 2018). Until recently, chitin usage had not been experimentally demonstrated for any member of *Acidobacteria* subdivision 1 (Kielak *et al.*, 2016). Nevertheless, Belova *et al.* (2018) isolated two strains of a novel genus and species, *Acidisarcina polymorpha*, bacteria with a wide repertoire of enzymes for the degradation of chitin, cellulose and xylan. The strains secreted chitinases linked to family GH18 and were capable of using chitin as carbon and nitrogen sources. In earlier studies, Ivanova (2016) identified *Acidobacteria* increased SSU rRNA transcript abundance in response to chitin availability in an acidic peatland

investigation.

3. Extracellular polymeric substances (EPS) production

EPS production has been reported for the *Acidobacteria* species *Granulicella paludicola*, *G. pectinivorans*, *G. aggregans*, *G. rosea* (Pankratov & Dedysh, 2010), *Acidicapsa borealis*, *A. ligni* (Kulichevskaya *et al.*, 2012) and *Terriglobus tenax* (Whang *et al.*, 2014). In addition, most acidobacterial genomes belonging to subdivision 1 (with the exception of '*Ca*. K. versatilis strain Ellin345) contain genes involved specifically in cellulose biosynthesis (Kielak *et al.*, 2016), which might be related to EPS production (Flemming *et al.*, 2007). EPS production is possibly contributing to acidobacterial cell protection and long-term survival in soil. The production of large amounts of EPS is related to abiotic stress, likely supporting dominance of *Acidobacteria* in acidic environments, resistance to heavy metals and pollutants like uranium, antimonium (Wang *et al.*, 2016), cadmium, lead, zinc, mercury (Guo *et al.*, 2017), petroleum compounds, linear alkylbenzene sulfonate (Sanchez-Peinado *et al.*, 2010) and p-nitrophenol (Paul *et al.*, 2006). To date, the only two acidobacterial EPSs that have been isolated and chemically characterized are produced by two strains of *Acidobacteria* subdivision 1, *Granulicella* sp. strain WH15 and strain 5B5. Both WH15EPS and 5B5EPS are able to emulsify oils and hydrocarbons, producing emulsions that are more thermostable over time than those of commercial EPS (Kielak *et al.*, 2017). In addition, EPS production allows these strains to colonize *Arabidopsis* roots, promoting plant growth (Kielak *et al.*, 2016). Interestingly, most of the characterized and industrially relevant EPS are composed of a maximum of four different monosaccharides (Rehm, 2010), while WH15EPS and 5B5EPS are heteropolysaccharides composed of 7 different monosaccharides, with xylose, mannose, glucose and galactose as main components (Kielak *et al.*, 2017). The composition of those EPS might be responsible for additional biological properties not present in EPS composed of more common sugar monomers (Roca *et al.*, 2015).

4. Genus *Granulicella*

The genus *Granulicella* currently contains 11 species: *G. paludicola* (T), *G. pectinivorans*, *G. aggregans*, *G. rosea,* isolated from *Sphagnum* peat bogs in Russia (Pankratov & Dedysh, 2010); *G. arctica*, *G. mallensis*, *G. tundricola*, *G. sapmiensis*, isolated from tundra soil in Finland (Mannisto *et al.*, 2012); *G. cerasi*, isolated from cherry bark in Japan (Yamada *et al.*, 2014); *G. acidiphila*, isolated from abandoned metal mines in Spain (Falagán *et al.*, 2017); and *G. sibirica*, isolated from organic tundra soil layer in Siberia (Oshkin *et al.*, 2019).

Granulicellas are gram-negative, non-spore forming, non-motile rods, occurring singly, in pair of short chains. These bacteria produce copious amount of EPS in culture media, and their colony colours vary from pale-pink to red, due to the production of carotenoid pigments. Strictly aerobic chemo-organotrophs, their preferred growth substrates are mono and polysaccharides. They grow in acidophilic and mesophilic conditions and are capable

to hydrolyze several polysaccharides, but cellulose and chitin breakdown has not yet been demonstrated (Pankratov & Dedysh, 2010).

The two strains used in my thesis, *Granulicella* sp. 5B5 and WH15, were isolated from decaying wood in the Netherlands (Valášková *et al.*, 2009). Comparisons among 16S rRNA gene sequences demonstrated that strain 5B5 is phylogenetically more related to *G. cerasi* and *G. paludicola*, while strain WH15 is closer to *G. tundricola* and *G. rosea* (Figure 2). The Average Nucleotide Identity (ANI) values (Table 2) showed that both strains do not belong to any of the species for which the genomes have been sequenced.

Table 2: Average Nucleotide Identity percentage values between *Granulicella* strains 5B5 and WH15 and other *Granulicella* strains' whole genomes available.

	G5B5	GHW15	GTUN	GPEC	GMAL	GSIB
G5B5	100	72.75	72.42	72.81	73.58	71.81
GHW15	72.75	100	74.12	73.32	73.03	73.19
GTUN	72.42	74.12	100	73.31	71.57	72.58
GPEC	72.81	73.32	73.31	100	71.57	74.88
GMAL	73.58	73.03	71.57	71.57	100	71.01
GSIB	71.81	73.19	72.58	74.88	71.01	100

GTUN –*G. tundricola*; GPEC-*G. pectinivorans*; GMAL-*G. mallensis*; GSIB-*G. sibirica*

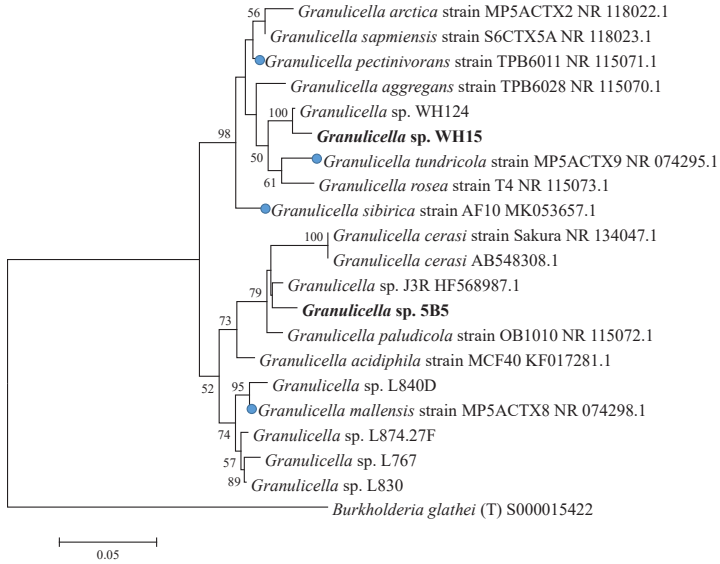

Figure 2: Dendogram showing the phylogenegtic relationships between *Granulicella* species based on comparisons between 16S rRNA gene sequences. The dendogram was constructed by using Maximum Likelihood (Tamura-Nei model) method (Tamura & Nei, 1993). Bootstrap values (expressed as percentages of 1,000 replications) are shown at branch points. *Burkholderia glathei* was used as outgroup. The strains used in this thesis are highlighted in bold letters. Blue circles indicate genomes compared to those of strains WH15 and 5B5 in ANI analysis

Thesis Outline

The high abundance and ubiquity of *Acidobacteria* in different environments, especially soils, raises intriguing questions about the physiological traits underlying their marked abundance. Genome sequences have provided relevant information, especially about *Acidobacteria* subdivision 1, the group for which various pure cultures are available. The increase in shotgun metagenomic studies and postgenomic analyses have enabled *de novo* assembly of acidobacterial genomes from environmental datasets and new insights into genome traits. Increased knowledge of the genomic features of different *Acidobacteria* subdivisions is critical for understanding their persistence in soil as well as their interactions with other soil microorganisms. Efforts to culture different acidobacterial genera and strains remain a top priority to decipher their genomic potential and to study their physiology and ecological functions.

The research presented in my thesis aimed at i) optimizing the growth of two strains of Granulicella, and ii) investigating the assimilation of the extracellular polymeric substances (EPS) of Granulicella strain WH15 by litter-topsoil microbial communities. To this end, I integrated different 'omic' approaches, including genomics, metagenomics, transcriptomics and proteomics, to expand the fundamental knowledge of their metabolism and interactions with other soil microbes, including the environmental fate of extracellular polymeric substances (EPS) produced by *Granulicella* (Figure 3).

EPS consists of highly hydrated polymers comprising polysaccharides, proteins and DNA (Wingender *et al.*, 1999) (Figure 4, Costa., unpublished). In **chapter 2**, we provide an overview

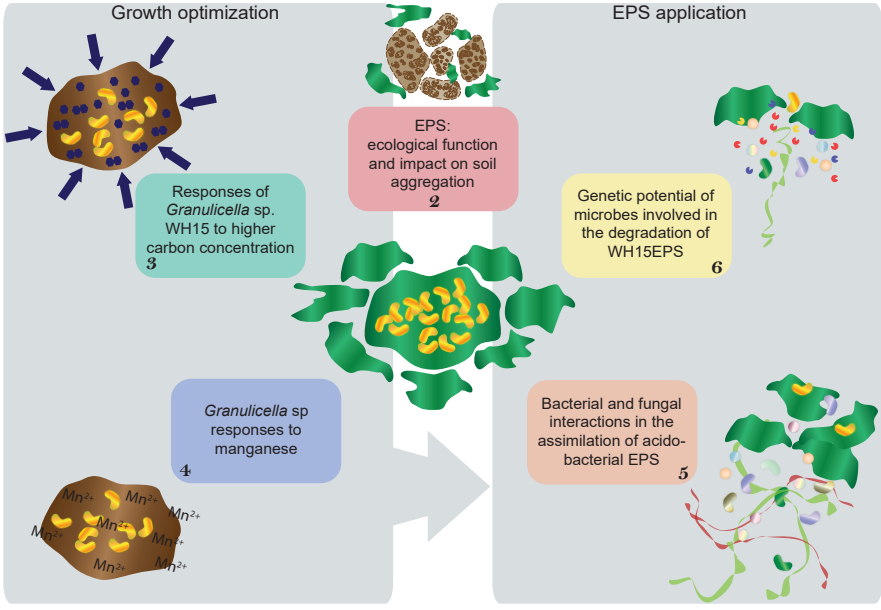

Figure 3: Schematic overview of the chapters presented in this thesis.

of the current knowledge on EPS biosynthesis, its chemical composition, factors influencing EPS production, the ecological functions of EPS and its application to improve soil particle aggregation. To study acidobacterial EPS, we first optimized growth conditions for higher bacterial biomass, necessary for EPS production and extraction. Simultaneously, we studied the impact of such growth optimizations on the metabolism of the *Granulicella* strains.

In general, strains of *Acidobacteria* exhibit slow growth under laboratory conditions, requiring low nutrient concentrations (Kielak *et al.*, 2016). However, growth at higher carbon concentrations was demonstrated for some isolates (de Castro *et al.*, 2013), including strains WH15 and 5B5 for which culture medium PSYL5 was developed (Campanharo *et al.*, 2016). In **chapter 3,** we evaluated transcriptional and proteomic responses of *Granulicella* strain WH15 grown at different concentrations of cellobiose. Our results demonstrated that higher cellobiose concentrations resulted in the higher expression of excretory functions and the reallocation of resources to maintenance of basic cell metabolism instead of production of new cell material.

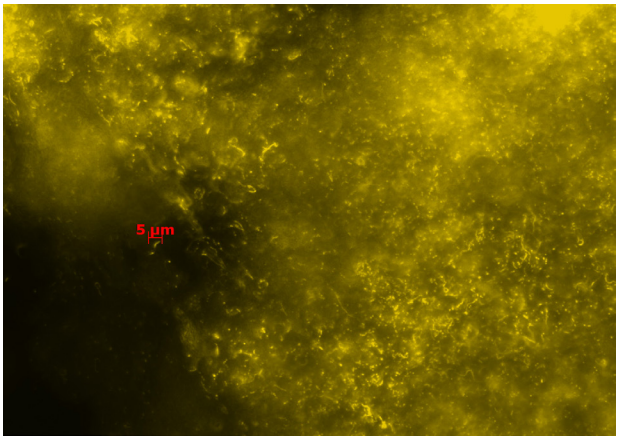

Figure 4: Extracellular DNA present in the EPS of liquid cultures of *Granulicella* sp.WH15 detected by Sytox™ Orange fluorescent nucleic acid stain (Invitrogen ™) (Unpublished data).

It is widely known that trace elements are important for microbial metabolism (Puri *et al.*, 2010). In **chapter 4,** we evaluated the impact of different trace elements on the growth of *Granulicella* strains WH15 and 5B5. The addition of trace element solution SL10 improved significantly the growth of both strains. When we further evaluated the effect of each of the trace elements separately, the results showed that primarily manganese (Mn) had a positive effect on the growth of both strains. To understand the effect of Mn on the metabolism of the two *Granulicella* strains, we adopted proteomics and genomics. Our results showed that the strains had different proteomic profiles and several uncharacterized metal ion transporters that could be involved in metal ion homeostasis. We postulate that these transporters could contribute to survival under high manganese concentrations present in the wood decomposition environment from where the strains were originally isolated.

Optimization of carbon concentration and manganese in culture medium allowed our strains, especially WH15, to grow faster in laboratory conditions, producing extractable amounts of EPS. The EPS of strain WH15 (WH15EPS) is mainly composed of polysaccharides (Kielak *et al.*, 2017) with a unique sugar composition that can be used as a nutrient source for other microorganisms. In **chapter 5**, we labeled WH15EPS with ^{13}C and investigated its effect on the assembly and co-occurrence of the active bacterial and fungal communities in topsoil by the stable isotope probing (SIP) approach. Our results demonstrated that WH15EPS was mainly assimilated by *Planctomycetes*, *Verrucomicrobia, Ascomycota* and *Basidiomycota* and co-inertia analysis suggested overall relationships between these kingdoms. Furthermore, comparisons among co-occurrence networks from labeled and unlabeled treatments demonstrated that hidden potential interactions can be unraveled by more specific and targeted metabolism studies. For instance, we observed the incorporation of WH15EPS by *Singulisphaera* and its connections to other *Planctomycetes* and *Acidobacteria*, which were not reported before.

The metabolization of WH15EPS and other biopolymers requires the production of a wide range of enzymes, such as glycoside hydrolases (GHs). GHs have applications in several industrial sectors, including biofilm removal, food processing and biofuel production. In **chapter 6,** we applied WH15EPS as an enrichment factor to target microorganisms and functions involved in EPS degradation through culture-independent and culture-dependent techniques. For this, we used topsoil samples obtained from the environment where *Granulicella* sp. WH15 was originally isolated. Our results showed a large diversity of glycoside hydrolase families with biotechnological potential and a high number of unclassified microorganisms that could be targeted for further studies.

In **chapter 7,** I integrate the overall findings of my thesis and discuss the most important observations concerning the impact of carbon sources and trace elements on the physiology of *Granulicella* and, more general, the ecological functions and environmental fate of EPS of *Acidobacteria*.

Chapter 2

Microbial extracellular polymeric substances: ecological function and impact on soil aggregation

Ohana Y. A. Costa, Jos M. Raaijmakers, and Eiko E. Kuramae

Adapted from: Costa OYA, Raaijmakers JM, Kuramae EE (2018). Microbial Extracellular Polymeric Substances: Ecological Function and Impact on Soil Aggregation. **Frontiers in Microbiology**. 9:1636.

Abstract

A wide range of microorganisms produce extracellular polymeric substances (EPS), highly hydrated polymers that are mainly composed of polysaccharides, proteins and DNA. EPS are fundamental for microbial life and provide an ideal substrate for chemical reactions, nutrient entrapment and protection against environmental stresses such as salinity and drought. Microbial EPS can enhance the aggregation of soil particles and benefit plants by maintaining the moisture of the environment and trapping nutrients. In addition, EPS have unique characteristics, such as biocompatibility, gelling and thickening capabilities, with industrial applications. However, despite decades of research on the industrial potential of EPS, only a few polymers are widely used in different areas, especially in agriculture. This review provides an overview of current knowledge on the ecological functions of microbial extracellular polymeric substances (EPS) and its application in agricultural soils to improve soil particle aggregation, an important factor for soil structure, health and fertility.

Keywords: EPS production, microorganisms, biosynthesis, ecological functions, soil aggregation.

1. Introduction

EPS are polymers biosynthesized by several strains of microorganisms. Composed mainly of polysaccharides, proteins and DNA, the production of these slimes is triggered primarily by environmental stresses. Since their biosynthesis is energetically expensive, they should generate some kind of advantage to the producer microorganism (Flemming & Wingender, 2010). Therefore, EPS production and functions have been studied for decades.

The polysaccharides are the most studied components of EPS. The investigation of EPS from numerous strains of microorganisms has demonstrated that the polysaccharides in these biopolymers vary immensely in composition and structure. They can be composed of one or many structural units, and the arrangement of these units is also exclusive for each different kind of EPS (Roca *et al.*, 2015). Aside from the carbohydrates, recently the interest in the structural proteins, enzymes and e-DNA has also been increasing. The analysis of e-DNA present in the EPS of a variety of strains has shown that the DNA is not innocuous, but can be a source of genetic exchange, signaling, attachment and moreover a very important structural component (Flemming & Wingender, 2010).

Besides the diversity of structures, EPS vary in their functions. A significant number of functions has been attributed to EPS, most of them related to protection. The matrix produced by EPS around microbial cells has the capability of shielding them against antimicrobial compounds and heavy metals; EPS matrix can also retain water, protecting microbes and the environment where it's contained against drought. In addition, other functions, such as adhesion, communication with other microbes and plants, antioxidant, aggregation, carbon storage, entrapment of nutrients have also been reported (Wingender *et al.*, 1999, Vardharajula & Ali, 2015, Wang *et al.*, 2015).

One of the roles of the EPS matrix that has been explored for decades is the capacity to aggregate soil particles, a function that is important for soil structure, health and fertility. Since EPS have a slimy texture and ionic charges, it can act like a glue, getting attached to clay and ions, holding solid particles together (Chenu, 1995). On the other hand, as stated before, EPS structures are variable, therefore their application efficiency in soils will vary accordingly. These polymers that are studied and produced in laboratorial conditions can be applied to soils for improvement of soil structure, fertility and quality. In this review we collate and synthesize the available information on EPS composition, biosynthesis, factors affecting EPS production, as well the ecological functions of microbial EPS and its application on soil particle aggregation.

2. EPS constituents

EPS are composed mainly of polysaccharides, proteins and DNA. However, the proportion of each component varies depending on the microbial strain and the method used to extract the EPS. Physical extraction methods, such as centrifugation, avoid the destruction of cells,

whereas aggressive methods, such as NaOH extraction, lyse cells, releasing their content into the EPS. Therefore, it is possible to find in the literature quantities of protein varying from 7.9 to 54.6 g per gram of EPS derived from the same sample using different extraction methods (Liu & Fang, 2002).

In general, polysaccharides are the main constituents of EPS and represent approximately 40% to 95% of the polymer (Flemming *et al.*, 2007). The polysaccharides can be classified as homopolysaccharides composed of a single type of monosaccharide or heteropolysaccharides composed of two or more types of monosaccharides (Sutherland, 2004). Common monosaccharides in EPS are D-glucose, D-galactose, D-mannose, L-fucose, L-rhamnose, D-arabinose, D-ribose and L-altrose. Less frequent are L-colitose, N-acetyl-L-fucosamine, N-acetyl-L-talosamine, L-iduronic acid, D-riburonic acid and 2-deoxy-D-arabino-hexuronic acid (Sutherland, 2004, Mishra & Jha, 2013). Polymers enriched with rare sugars are of potential interest because their unusual composition and structure may confer additional attributes, such as anti-inflammatory and antioxidant properties.

A significant number of proteins with different functions have been observed in EPS, including several trapped extracellular enzymes. The products of these enzymes consequently remain close to the cell, facilitating their uptake by the bacteria (Wingender *et al.*, 1999, Flemming & Wingender, 2010). Some EPS-modifying enzymes are capable of degrading the polymer during starvation. However, this process is slow, since no single enzyme is capable of degrading all of the polysaccharides present in the EPS matrix. In general, highly specific enzymes are required for this task (Sutherland, 2004, Flemming & Wingender, 2010). In addition, structural proteins are involved in the formation and stabilization of the polysaccharide chain and are responsible for the connection between the cell surface and the extracellular EPS (Flemming & Wingender, 2010).

Extracellular DNA (e-DNA) of different origins is an important EPS component in biofilms. Although the function and origin of e-DNA have not been completely elucidated, studies have shown that it is responsible for the structure of certain EPS and plays a role in adhesion to surfaces and signaling. E-DNA is likely an important structural component of *Staphylococcus aureus* biofilms but is not essential in biofilms produced by *Streptococcus epidermidis* (Flemming & Wingender, 2010). This conclusion is based on the fact that treatment with DNase I inhibits biofilm formation and detachment of preformed biofilms by *S. aureus* but not *S. epidermidis* (Izano *et al.*, 2007). In *Pseudomonas aeruginosa*, e-DNA is essential for biofilm formation, as DNase I inhibits this process (Whitchurch *et al.*, 2002). Moreover, *Bacillus cereus* mutants produce a weaker biofilm when lacking a purine biosynthesis gene involved in e-DNA production (Vilain *et al.*, 2009).

3. EPS biosynthesis

EPS production has been reported for bacteria and cyanobacteria as well as microalgae

(Parikh & Madamwar, 2006, Boonchai *et al.*, 2014), yeasts (Pavlova & Grigorova, 1999), basidiomycetes (Hwang *et al.*, 2004, Elisashvili *et al.*, 2009) and protists (Jain *et al.*, 2005). EPS are formed by the polymerization of repeating units of similar or identical monomers and are classified as loosely bound or tightly bound depending on their association with the cell (More *et al.*, 2014).

Initially, EPS was used as an abbreviation for "extracellular polysaccharides", "exopolymers" or "exopolysaccharides"; however, studies have shown that the matrix is much more complex and includes structural proteins, enzymes, nucleic acids, lipids and other compounds such as humic acids (Wingender *et al.*, 1999, Wingender *et al.*, 1999, Flemming & Wingender, 2010). The mucoid substances present in EPS are not only produced by the microorganism but are also derived from cellular lysis, hydrolysis of macromolecules and absorption from the environment. Each component contributes to the physicochemical characteristics of the matrix (Nielsen & Jahn, 1999). EPS physically involve microbial cells and mediate contact and exchange processes within microbial communities as well as with the environment. The EPS matrix provides a hydrated and buffered environment that facilitates chemical reactions (Wingender *et al.*, 1999).

Studies of genes involved in EPS production have focused on a few polymers, and the biosynthesis pathways of the polysaccharides composing EPS have been widely described. Four main biosynthesis pathways are known: (1) the Wzx-Wzy-dependent pathway; (2) the synthase-dependent pathway, (3) the ABC transporter-dependent pathway, and (4) extracellular synthesis by sucrase enzymes (Schmid *et al.*, 2015). In the Wzx-Wzy-dependent secretion pathway, the individual repeating units are assembled on an undecaprenyl-phosphate carrier located in the cytoplasmic portion of the inner membrane and then transported to the periplasm by a Wzx flippase (Whitney & Howell, 2013, Schmid *et al.*, 2015). Once in the periplasm, the putative polymerase Wzy assembles the polymer units, which are transported across the outer membrane by a complex formed by a polysaccharide copolymerase and an outer membrane polysaccharide exporter (Figure 1). The size of the chain is regulated by the polysaccharide copolymerase Wzz (Cuthbertson *et al.*, 2009). Xanthan produced by *Xanthomonas campestris* (Vorhölter *et al.*, 2008) and gellan produced by *Sphingomonas paucimobilis* (Wang *et al.*, 2006) are examples of EPS synthesized via this pathway.

In the ABC transporter-dependent pathway, the polymer is fully synthesized in the cytoplasm and then transported across the inner membrane by a dedicated ABC transporter. As in the Wzx-Wzy-dependent pathway, transport to the outside of the cell is accomplished by the polysaccharide copolymerase and polysaccharide exporter complex (Cuthbertson *et al.*, 2009, Whitney & Howell, 2013, Schmid *et al.*, 2015) (Figure 1). This pathway is involved in capsular polysaccharide production (Whitney & Howell, 2013).

For the synthase-dependent pathway, it has been proposed that biosynthesis and transport to the periplasmic space are accomplished by the same protein, a polymerizing

glycosyltransferase (synthase) (Hubbard *et al.*, 2012, Schmid *et al.*, 2015). Transport across the outer membrane is accomplished by a tetratricopeptide repeat (TRP) domain-containing protein and a β-barrel porin (Keiski *et al.*, 2010, Whitney & Howell, 2013) (Figure 1).

In general, this pathway is used for the production of EPS consisting of only one type of sugar precursor (Schmid *et al.*, 2015), such as the bacterial cellulose from *Komagataeibacter medellinensis* (Matsutani *et al.*, 2015) and curdlan from *Agrobacterium sp.* (Stasinopoulos *et al.*, 1999). Extracellularly synthesized EPS are assembled by glucansucrases, enzymes that are secreted and anchored to the cell wall. These enzymes catalyze the transfer of glucose from sucrose to the growing polysaccharide chain (Rehm, 2010, Schmid *et al.*, 2015). In some strains of microorganisms, such as the dextran producer *Leuconostoc mesenteroides*, the expression of glucansucrases is induced by sucrose (Kim & Robyt, 1994), whereas in some levan/inulin-producing *Lactobacillus strains*, the genes *levS* (levansucrase) and *inu* (inulosucrase) are expressed constitutively (Tieking *et al.*, 2004, Schwab *et al.*, 2007). The mechanism of sucrose induction, however, remains unknown. A summary of the genes, structure and producing microorganisms of industrially relevant EPS is provided in Table 1. For a review on the biosynthesis pathways of industrially important EPS, refer to Schmid et al. (2015).

EPS production and regulation have been studied for several decades because these polymers have biotechnological applications and are widely used in the pharmaceutical and food industries. For some polymers, such as xanthan, alginate and curdlan, the genes

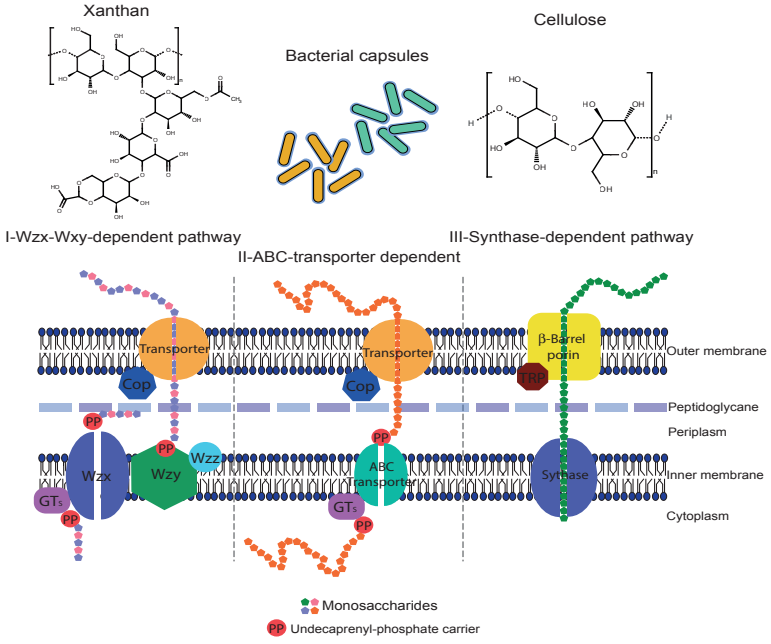

Figure 1: Simplified schematic representation of the three main intracellular biosynthetic pathways for microbial EPS production and example structures for each pathway. COP: copolymerase; GTs: glycosyl hydrolases; TRP: tetratricopeptide repeat protein. Figure modified according to Schmid *et al.* (2015).

responsible for production and structure have already been described and deeply explored. Current studies are focused on applying engineering strategies to improve the yields and characteristics of these polymers by targeting precursors and genes involved in their production or regulation. However, knowledge of the structures and functions of other types of polymers, especially EPS secreted by lesser-known strains, is relatively deficient.

4. Factors influencing EPS production

EPS production by microorganisms can be triggered by environmental and physiological conditions, such as carbon source, nitrogen starvation and ionic strength (Janczarek, 2011, Carzaniga *et al.*, 2012). Under stress conditions, EPS are synthesized to establish a physical barrier around the cell (Kehr & Dittmann, 2015). Microbial EPS production has mainly been observed in pure culture and varies according to environmental conditions (Chenu, 1995). In laboratory culture media, physiological conditions can be controlled to achieve high EPS yield and modify polymer characteristics, including relative molecular mass, polymer pattern, number of residues and degree of branching (Dumitriu, 2005). The composition of the polymer, however, is strain specific, and only some bacteria, such as *Enterobacter* strain A47 (Torres *et al.*, 2012), can be induced to change the polymer pattern (Roca *et al.*, 2015). Moreover, there are no standard conditions that promote high EPS production, since carbon and nitrogen sources, mineral requirements, optimal pH and temperature differ for each microorganism (Kumar *et al.*, 2007).

4.1. Carbon source

Carbon source is one of the main factors influencing EPS yield; therefore, many studies of EPS production have assessed the influence of a variety of carbon sources on EPS yield and the biomass of microorganisms. These carbon sources include glucose, fructose, lactose, maltose, mannitol, sorbitol, starch and sugar concentrates (Neosorb™, Cerelose™) (Kumar *et al.*, 2007). Glucose and fructose typically deliver the highest amount of EPS. Two strains of *Lactobacillus delbrueckii* subsp. *bulgaricus* (Petry *et al.*, 2000), *Chryseobacterium indologenes* MUT.2 (Khani *et al.*, 2016), *Lactobacillus delbrueckii* subsp. *bulgaricus* strains B3 and G12 and *Streptococcus thermophilus* W22 (Yuksekdag & Aslim, 2008) produce EPS more efficiently with glucose as the carbon source in the culture medium. *Trametes versicolor*, however, produces EPS more efficiently when fructose is the carbon source (Bolla *et al.*, 2010). Other strains have different requirements, such as *Rhizobium leguminosarum* biovar *trifolii* TA-1, which has a preference for mannitol, and *Halomonas alkaliantarctica* strain CRSS, which requires acetate (Poli *et al.*, 2004) for high EPS production.

Table 1: Industrially relevant polymers: genes, structures and biosynthesis pathways.

Polymer	Genes	Structure (example of producing microorganism)	Biosynthesis pathway	Reference
Alginate	*alg8, alg44, algD, algG, algX, algK, algE*	β-(1-4)-linked non-repeating heteropolymer (*Pseudomonas aeruginosa, Azotobacter vinelandii*)	Synthase-dependent	(Remminghorst & Rehm, 2006, Rehm, 2010)
Bacterial cellulose	*bcsA, bcsB, bcsC, bcsZ*	β-(1-4)-Glucan (*Acetobacter xylinum)*	Synthase-dependent	(Rehm, 2010)
Curdlan	*crdS, crdA, crdR, cdrC* and *crdB*	β-(1-3)-Glucan (*Agrobacterium* strains*)*	Synthase-dependent	(Rehm, 2010, Schmid *et al.*, 2015, Yu *et al.*, 2015)
Dextran*	*dsr* variations (*dsrF*, *dsrD*, *dsrX*, *dsrN*, *dsrR*)	α-(1-6)-glucose backbone with α-(1→2), α-(1→3), α-(1→4) branched linkages (*Leuconostoc mesenteroides*)	Extracellular (Dextransucrase)	(Kim & Robyt, 1994, Schmid *et al.*, 2015)
Gellan	pgmG, ugpG	β-(1,3)-linked repeating polymer composed by tetrasaccharide units (*Sphingomonas paucimobilis*)	Wzx/Wzy-dependent	(Sá-Correia *et al.*, 2002, Schmid *et al.*, 2015)
Inulin*	*inuJ*, *inuG*, *inuB*, *islA*, *ftf*	b(2-1) fructan (*Lactobacillus* and *Leuconostoc* strains)	Extracellular, (Inulosucrase)	(Anwar *et al.*, 2010, Schmid *et al.*, 2015)
Levan*	*levG*, *lsdA*, *levS*, *sacB*, *sacY*	β-(2,6) D-fructose backbone with β-(2,1) D-fructose and glucose branches (*Lactobacillus*, *Bacillus* and *Aerobacter* strains)	(Extracellular, Levansucrase)	(Srikanth *et al.*, 2015)
Succinoglycan	*exoC, exoB, exoN, exoY, exoF, exoA, exoL, exoM, exoO, exoU, exoW, exoP, exoT, exoQ, exoZ, exoH, exoV, exoK*	Octasaccharide repeating units with acetyl, pyruvyl, and succinyl substitutions (*Rhizobium*, *Agrobacterium*, *Alcaligenes* and *Pseudomonas* strains)	Wzx/Wzy-dependent	(Reuber & Walker, 1993, Schmid *et al.*, 2015)
Xanthan	*gumB, gumM, gumH, gumK, gumI, gumJ, gumE, gumC, gumB, gumD, gumL, gumF, gumG*	β-(1-4)-linked glucose backbone and a side chain composed of two mannose units and one glucuronic acid (*Xanthomonas campestris)*	Wzx/Wzy-dependent	(Becker *et al.*, 1998, Rehm, 2010)

*The genes are not part of a cluster; each encodes a different glucansucrase in a different strain.

4.2. Nitrogen source

The main nitrogen sources employed in EPS production are ammonium sulfate, peptone, sodium nitrate, urea and yeast extract (Kumar *et al.*, 2007). However, the highest growth rates and EPS yields are reached when complex nitrogen sources are involved, probably due to the presence of growth factors (Farrés *et al.*, 1997), for which requirements vary among microorganisms. In addition, carbon found in the nitrogen source increases the carbon/nitrogen (C/N) ratio and thereby enhances EPS production (Kumar *et al.*, 2007).

Similar to carbon sources, several studies have compared different compounds to identify the best nitrogen supply for a variety of microbial strains reflecting the metabolic diversity among EPS producers. Among the nitrogen sources applied to optimize EPS production, the compounds that generally induce the highest yields are yeast extract and different types of peptones. However, inorganic nitrogen supplies can also induce high polymer production; ammonium sulfate is the best source for a high EPS yield from *Gluconacetobacter hansenii* LGM1524 (Valepyn *et al.*, 2012), whereas ammonium nitrate and sodium nitrate are optimal for *Bacillus megaterium* (Gandhi *et al.*, 1997) and some rhizobial strains (Kumar and Ram (2014).

4.3. Carbon/nitrogen ratio

The C/N ratio is as important as carbon type or nitrogen source and greatly affects microbial metabolism and, consequently, EPS production. Many studies have reported maximization of EPS production under nitrogen limitation and carbon excess. However, like other culture nutrient variables, there is no fixed ideal C/N ratio for all microorganisms (More *et al.*, 2014). EPS production by *Rhizobium tropici* reaches its maximum yield at a C/N of 20 (Staudt *et al.*, 2011), whereas *Rhodoblastus acidophilus* (formerly known as *Rhodopseudomonas acidophila*) requires a C/N ratio of 7.7 at low concentrations of carbon ($C_4H_4Na_2O_4$) and nitrogen sources ($(NH_4)_2SO_4$) (Sheng *et al.*, 2006). By contrast, for some strains of lactic acid bacteria (LAB), nitrogen limitation does not increase EPS yield. The production of EPS by *Streptococcus thermophilus* is dependent on high carbon and nitrogen concentrations (De Vuyst *et al.*, 1998). In addition, the effect of the C/N ratio may depend on the culture medium used for growth. Gonzalez Garcia et al. (2015) evaluated the effects of variable C/N ratios on EPS production by *Saccharophagus degradans* in basal culture medium (BM) and nutrient-limited medium (NL). In BM, variation of the C/N ratio did not affect EPS production, whereas an enhancement of EPS production was observed in NL, indicating a possible effect of the combination of the C/N ratio and nutrient limitation (N, P, K, Ca, Mg, and Fe).

4.4. Other nutrients and trace elements

In addition to carbon and nitrogen, nutrients such as Mn, Zn, Co, Mo, vitamins, P and O_2 are required for EPS synthesis and influence the conversion of precursors into polysaccharide (More *et al.*, 2014, González-García *et al.*, 2015). However, metal ion requirements differ among

microbial strains. Mg^{2+} appears to enhance EPS production by *Lactobacillus rhamnosus* C83 (Gamar-Nourani *et al.*, 1998) and *Stemphylium sp*. (Banerjee *et al.*, 2009), whereas the addition of phosphate in the medium decreases the EPS yield of *Klebsiella* I-174 (Farrés *et al.*, 1997). The presence of Na^{+} increases the EPS yields of *Rhodopseudomonas acidophila* (Sheng *et al.*, 2006) and *L. rhamnosus* (Gamar-Nourani *et al.*, 1998), suggesting a defensive response of the bacteria to salt stress. In addition to metal ions, compounds such as histidine, tyrosine, phenylalanine and xanthine are important for EPS production by *Stemphylium sp*. (Banerjee *et al.*, 2009).

4.5. Temperature and pH

Temperature and pH often influence EPS production (Kumar *et al.*, 2007, More *et al.*, 2014), but the optimal values of both parameters vary among microorganisms. Incubation at a temperature lower than the optimal temperature for bacterial growth typically enhances EPS biosynthesis because when cells grow slowly, the synthesis of the cell wall is slower, and more sugar precursors are available for EPS production (Sutherland, 2001). Most strains that produce EPS grow at a temperature range of 25-30 °C (More *et al.*, 2014). However, EPS production has been reported for psychrophilic microorganisms such as *Pseudoalteromonas* strain CAM025, which has an increased polymer yield between -2 °C and 10 °C (Nichols *et al.*, 2005), and *Colwellia psychrerythraea* strain 34H, which exhibits the highest yield at -8 °C (Marx *et al.*, 2009). EPS secretion has also been observed between 60 °C and 65 °C for the thermophilic bacteria *Bacillus thermodenitrificans* DSM 465 (Nicolaus *et al.*, 2000) and *Bacillus thermoantarcticus* (Manca *et al.*, 1996).

Many microbial strains grow and produce EPS in a neutral pH range, and EPS synthesis generally requires a stable pH for maximum production (Kumar *et al.*, 2007). Thus, extreme pH variation can decrease the polymer yield. Xanthan, curdlan and gellan, polymers used in industry, are produced at a pH range of 7.0–7.5 (Kalogiannis *et al.*, 2003, Nampoothiri *et al.*, 2003, Shih *et al.*, 2009). However, since the optimal pH varies among microorganisms, EPS formation has been observed at a wide range of pH. *Rhizobium tropici* (Staudt *et al.*, 2011) and *Rhizobium ciceri* (Küçük & Kivanç, 2009) produce EPS at neutral pH, with a drastic decrease in yield under acidic conditions. Synthesis of EPS by *Enterobacter* strain A47 decreases significantly with increasing pH of the medium (Torres *et al.*, 2012). By contrast, *Halomonas alkaliantarctica* strain CRSS can produce EPS at pH 8.0 and 9.0 (Poli *et al.*, 2004), and the optimum pH for highest EPS yield is 5.0 for *Antrodia camphorata* (Shu & Lung, 2004). These data illustrate the diversity of conditions for EPS production and the efforts of researchers to increase the yields of EPS secreted by different microorganisms. However, due to the high diversity of available EPS, many strains and polymers remain to be evaluated.

5. Ecological functions

EPS biosynthesis is an energy-demanding process. Therefore, its production requires selective advantages in the environment of the producing microorganism. In laboratory cultures, the production of EPS does not impact cell viability or growth and thus appears not to be essential for survival. However, in natural environments, most microorganisms live in aggregates, such as flocs and biofilms, for which EPS are structurally and functionally essential (Wingender *et al.*, 1999). Most of the functions attributed to EPS are related to protection of the producing microorganism. Diverse variations in abiotic conditions such as drought, temperature, pH and salinity can trigger the production of EPS as a response to environmental stresses (Wingender *et al.*, 1999, Kumar *et al.*, 2007, Vardharajula & Ali, 2015). The functions of EPS are summarized in Figure 2.

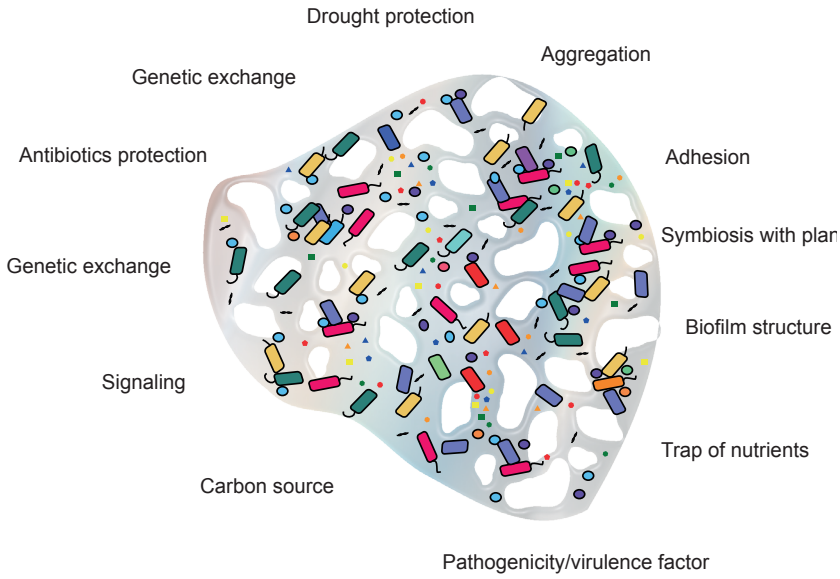

Figure 2: Conceptual framework of the functions of microbial extracellular polymeric substances (EPS) in soil.

5.1. Functions of EPS in interactions with the environment

5.1.1. Adhesion/ Cohesion/ Genetic material transfer

EPS are responsible for the cohesion of microorganisms and adhesion of biofilms to surfaces, influencing spatial organization, allowing interactions among microorganisms, and acting as adhesives between cells (Wolfaardt *et al.*, 1999). These functions are important for the establishment and biological activities of biofilms and flocs. The polymers mechanically stabilize the microbial aggregates via several types of interactions between the macromolecules, including dispersion forces, electrostatic interactions and hydrogen

bonds. The resultant formation of a gel-like tridimensional structure around the cells allows the microorganisms to be retained near each other to establish stable consortia (Flemming *et al.*, 2000). For example, EPS of *Sphingomonas paucimobilis* have surface-active properties that promote and enhance attachment via the formation of polymeric bridges (Azeredo & Oliveira, 2000). The quantity of EPS can also influence cell adhesion, as demonstrated by Tsuneda *et al.* (2003). For the 27 bacterial strains evaluated, small quantities of EPS inhibited cell adhesion by electrostatic forces, whereas large amounts enhanced adhesion via interactions between functional groups in the EPS, such as uronic acids and acetyl groups. The nature of the interactions between the functional groups in EPS, however, is unknown. In addition, the matrix formed by EPS can facilitate chemical communication and even influence predator-prey interactions (Flemming *et al.*, 2007). Joubert *et al.* (2006) observed that ciliated protists preferred feeding on planktonic cells and the EPS matrix rather than on attached and biofilm-derived cells. In addition, the presence of protists appeared to enhance yeast metabolic activity in the biofilm.

Together with different protein adhesins, EPS are believed to be involved in the initial steps of microbial adhesion to surfaces. For instance, the polysaccharide produced by *Caulobacter crescentus*, called holdfast, is crucial for the initial surface attachment, together with other cellular structures (Entcheva-Dimitrov & Spormann, 2004, Wan *et al.*, 2013). However, the characteristics of each polymer are defined by their composition, as adhesiveness depends heavily on chain conformation, internal substituents and internal/external interactions (Berne *et al.*, 2015). Therefore, the extent to which the type of polymer contributes to the adhesive properties of bacterial cells remains to be determined.

In addition to polysaccharides, extracellular DNA (exDNA) seems also to be responsible for the adhesive properties of some EPS. Although the functions of exDNA have not been completely elucidated, studies have shown that it is responsible for the cohesion and structure of certain EPS and plays a role in adhesion to surfaces and signaling (Okshevsky & Meyer, 2013). Released by autolysis or active secretion by microorganisms, exDNA is likely an important structural component of *Staphylococcus aureus*, *Pseudomonas aeruginosa* and *Ralstonia solanacearum* biofilms (Whitchurch *et al.*, 2002, Minh Tran *et al.*, 2016); however, it is not essential in biofilms produced by *Streptococcus epidermidis* (Flemming & Wingender, 2010). This conclusion is based on the fact that treatment with DNase I inhibits biofilm formation and detachment of preformed biofilms by *S. aureus* but not *S. epidermidis* (Izano *et al.*, 2007). Enhancement of genetic material transfer between microorganisms is another property of extracellular polymers. ExDNA of different origins is an important EPS component in biofilms, where microorganisms are surrounded by an EPS matrix. Although studies in this area are scarce, the rates of natural transformation and conjugation of bacteria appear to be higher within biofilms. Bae et al. (2014) demonstrated that *Campylobacter jejuni* transfers antibiotic resistance genes by natural transformation more frequently in biofilms than in planktonic cells. Other studies have shown that biofilm age and DNA concentration influence the frequency

of transformation events, whereas a high density of planktonic cells inhibits transformation in biofilms (Hendrickx *et al.*, 2003). Moreover, the number of events observed can depend greatly on the technique used to detect conjugative gene transfer in biofilms. For instance, Hausner and Wertz (1999) detected 1000-fold higher conjugation rates using confocal laser scanning microscopy than by classic plating techniques. It has been suggested that exDNA fractions can be used in environmental studies as an alternative method for microbial activity measurement. However, exDNA fraction separation and evaluation in complex samples, such as soils, has yet to be improved (Nagler *et al.*, 2018). Estimates of microbial community composition can be influenced by the presence of exDNA (Carini *et al.*, 2016).

5.1.2. Symbiosis

EPS play an important role in the establishment of symbiosis between nitrogen-fixing rhizobia and plants. Rhizobial surface polysaccharides are fundamental for nodule formation by some legumes, although the underlying mechanisms are not yet fully resolved. For example, to invade alfalfa nodules and establish successful symbiosis, *Sinorhizobium meliloti* Rm1021 must produce succinoglycan (Cheng & Walker, 1998). Mutants that do not synthesize succinoglycan produce modified polymers or overproduce EPS, reduce the capacity of *S. meliloti* Rm 1021 to infect and establish symbiosis. Although capable of producing nodules, *Rhizobium leguminosarum* biovar *viciae* glucomannan (*gmsA*) mutants are strongly outcompeted by wild-type bacteria in mixed inoculations of *Pisum sativum* (Williams *et al.*, 2008). The interaction between the EPS of *Mesorhizobium loti* strain R7A and *Lotus japonicus* was recently shown to be mediated by a receptor expressed by the plant. *Lotus japonicus* produces a receptor (EPR3) that binds to and permits infection by only bacteria that produce EPS with a specific structure; mutants with truncated EPS are less successful in infection (Kawaharada *et al.*, 2015). The expression of this receptor demonstrates that the plant is capable of recognizing the structure of EPS produced by rhizobia.

5.1.3. EPS as pathogenicity/virulence factors

For some bacteria, polymers function as pathogenicity and virulence factors. For example, the high virulence of *Erwinia amylovora* is a result of the production of amylovoran and levan. Both polymers contribute to the pathogenesis of the bacteria, and the absence of either amylovoran or levan dramatically decreases plant colonization (Koczan *et al.*, 2009). In addition, EPS can serve as a mechanical barrier between bacteria and plant defense compounds by decreasing the diffusion rates of these compounds. For example, the polymers of *Pseudomonas syringae* pv. *phaseolicola* and *Sinorhizobium meliloti* protect the bacteria against reactive oxygen species (ROS) produced by the plant host during infection, thereby decreasing oxidative stress (Király *et al.*, 1997, Lehman & Long, 2013). *Sinorhizobium meliloti* mutants overproducing EPS protect polymer-deficient mutants against H_2O_2 (Lehman & Long, 2013). Alginate, the EPS produced by *Pseudomonas aeruginosa*, a human

opportunistic pathogen, protects the bacteria against the inflammatory process of the host, avoiding free radicals, antibodies and phagocytosis and thereby aggravating the prognosis of patients infected by *P. aeruginosa* (Ryder *et al.*, 2007). Althought is known that EPS may act as an antioxidant, less in known the chemical mechanism of protection against ROS.

5.2. EPS and nutrition

5.2.1. Carbon reserves

EPS produced by microorganisms might act as carbon reserves, but few studies have investigated the role of EPS in nutrition or cross-feeding between organisms. Since EPS are generally complex molecules, their complete degradation would require a wide range of different enzymes (Flemming & Wingender, 2010). *Rhizobium* NZP 2037 can use its own poly-b-hydroxybutyrate (PHB) and EPS as sole sources of carbon for survival in carbon-restriction situations (Patel & Gerson, 1974). However, EPS is a higher potential carbon source than PHB. Stable isotope probing (SIP) is a powerful strategy for detecting microorganisms that can degrade polymers. Wang et al. (2015) labelled the EPS of *Beijerinckia indica* and observed that the polymer was assimilated by bacteria with low identities to known species, particularly members of the phylum *Planctomycetes*. In addition, the authors isolated bacteria that used the EPS as a sole carbon source, demonstrating the potential utility of these polymers for isolating new microbial species.

5.2.2. Nutrient trap

In addition to supplying carbon, EPS can accumulate other nutrients and molecules. The retention of extracellular enzymes in the EPS matrix promotes the formation of an extracellular digestion system that captures compounds from the water phase and permits their use as nutrient and energy sources (Flemming & Wingender, 2010). Many studies have investigated the adsorption of metal ions by EPS for heavy-metal remediation and recovery of polluted environments. The EPS of *Paenibacillus jamilae* adsorbs multiple heavy metals (Pb, Cd, Co, Ni, Zn and Cu) with strongler interaction with Pb, a maximum binding capacity of 303.03 mg/g, tenfold higher than the binding capacities for other metals (Morillo Pérez *et al.*, 2008). The polymers produced by *Anabaena variabilis* and *Nostoc muscorum* possess similar affinities for Cu, Cd, Co, Zn and Ni, with the highest affinity for Cu and the lowest for Ni. Both bacterial EPS are promising for the removal of toxic heavy metals from polluted water (El-Naggar *et al.*, 2008). The EPS of *Pseudomonas sp*. CU-1 has a high Cu-binding capacity and thus, protect bacterial cells against this metal ion (Lau *et al.*, 2005).

5.3. EPS in protection against abiotic and biotic stresses

5.3.1. Drought protection

EPS production can confer advantages to microorganisms in environments under drought

stress. A high water-holding capacity was observed for an EPS produced by a *Pseudomonas* strain isolated from soil; this EPS can hold several times its weight in water. When added to a sandy soil, the EPS altered its moisture by allowing the amended soil to hold more water than unamended soil (Roberson & Firestone, 1992). According to the authors, the EPS protected the bacteria against desiccation by acting like a protective sponge, thereby giving the bacteria time to make metabolic adjustments. This polymer exhibits significant structural modifications during desiccation and may be an important protection factor traps a reservoir of water and nutrients for bacterial survival (Roberson *et al.*, 1993). Cyanobacteria isolated from arid regions, such as *Nostoc calcicola* (2014) and *Phormidium* 94a (2004), are also capable of producing EPS, which may represent a strategy for water/nutrient retention and survival.

5.3.2. Salt tolerance

Some studies have revealed that microbial polymers are involved in tolerance to salt stress, not only for the producer microorganisms but also for the associated plants. The production of polymer by NaCl-tolerant isolates can decrease Na uptake by plants by trapping and decreasing the amount of ions available (Upadhyay *et al.*, 2011). Therefore, the polymer prevents nutrient imbalance and osmotic stress, which can promote survival of the microorganisms and benefit of the plant. *Sinorhizobium meliloti* strain EFBI cells severely reduce EPS production when inoculated in culture medium with low salt concentration. Since this strain was isolated from the nodules of a plant growing in a salt marsh with a salinity level of 0.3 M, a lower amount of salt can be considered a stressful condition. However, the relevance of this EPS for survival and symbiosis was not further studied (Lloret *et al.*, 1998).

5.3.3. Protection against low/high temperatures

The production of EPS at low temperatures is an important factor in the cryoprotection of sea-ice organisms as well as a natural adaptation to low temperatures and high salinities. High concentrations of EPS have been observed in samples collected from Arctic sea ice; the EPS shields diatoms against the severe environmental conditions during the winter season (Krembs *et al.*, 2002). In addition, EPS alter the microstructure and desalination of growing ice, consequently improving microbial habitability and survivability (Krembs *et al.*, 2011).
EPS can be a protection factor for thermophilic bacteria by shielding microorganisms from very high temperatures. The polymers produced by *Bacillus sp*. strain B3-72 and *Geobacillus tepidamans* V264 are not easily dissolved at high temperatures (Nicolaus *et al.*, 2000, Kambourova *et al.*, 2009). A few studies (Manca *et al.*, 1996, Nicolaus *et al.*, 2000, Nicolaus *et al.*, 2004) have evaluated EPS production by thermophilic bacteria and archaea for potential applications of these polymers in industry and the recovery of polluted environments. However, the structure and the ecological function of these slimes remain to be established.

5.3.4. Protection against antimicrobials

The matrix that surrounds microorganisms in biofilms plays an important role in decreased susceptibility to antimicrobials. In general, biofilm matrices possess a negative charge and therefore bind positively charged compounds, protecting the innermost cells from contact. In addition, electrostatic repulsion can reduce the diffusion rates of negatively charged antimicrobials through the biofilm (Everett & Rumbaugh, 2015). Many studies have tested the inhibitory potential of bacterial EPS against antimicrobial compounds, particularly for clinically important bacterial strains. A few studies have demonstrated that the slime produced by *Staphylococcus* sp. is an effective antagonist to vancomycin, perfloxacin and teicoplanin, acting as a barrier to the compounds or even interfering with their action in the cell membrane (Farber *et al.*, 1990, Souli & Giamarellou, 1998). The EPS produced by *Acinetobacter baumannii* is also protective against tobramycin exposure and is effective regardless of the bacterial species exposed. By contrast, the polymer from *S. aureus* has no protective effect against tobramycin (Davenport *et al.*, 2014). EPS can also protect microorganisms against disinfection agents. Alginate produced by *Pseudomonas aeruginosa* enhances bacterial survival in chlorinated water, and removal of the slime eliminates bacterial chlorine resistance (Grobe *et al.*, 2001).

The few EPS isolated thus far have a wide range of functions, but a huge diversity of polymers produced by microorganisms with different functions awaits exploration and discovery. The different functions already discovered are consequences of the diverse EPS structures, and are connected to the benefits they can have when applied to soils. The production of EPS is not only and advantage to the microbes, but to the soil environment in general. The adhesiveness is important for gluing soil particles together; high water holding capacity protects microorganisms and plants against drought, as well as permits the diffusions of nutrients in the environment. EPS production also influences and is influenced by interactions between plants and microorganisms, increasing the availability of nutrients as a whole, promoting plant and microbial growth. In the next section, we summarize how the currently known EPS are applied to agricultural soils and their benefits for soil aggregation.

6. Application of EPS on soil aggregation

6.1. Soil aggregates and microbial communities

Aggregates are the basic units of soil structure and are composed of pores and solid material produced by rearrangement of particles, flocculation and cementation. These units define the physical and mechanical properties of soil, such as water retention, water movement, aeration and temperature, which in turn affect physical, chemical and biological processes (Alami *et al.*, 2000, Tang *et al.*, 2011). Aggregates are important for the improvement of soil fertility, porosity, erodibility and agronomic productivity by influencing plant germination and root growth (Dinel *et al.*, 1992, Bronick & Lal, 2005). Aggregate formation involves

numerous factors: vegetation, soil fauna, microorganisms, cations and interactions between clay particles and organic matter (Kumar *et al.*, 2013). The stability of aggregates depends on their internal cohesion, pore volume, connectivity, tortuosity and pore-wall hydrophobicity (Chenu & Cosentino, 2011). A good soil structure, dependent on aggregation, is fundamental for sustaining agricultural productivity and environmental quality, sustainable use of soil and agriculture (Amézketa, 1999).

The hierarchical model for classifying soil aggregates suggests that larger aggregates are composed of smaller units, which are formed from even smaller aggregates (Tisdall & Oades, 1982) (Figure 3). In persistent microaggregates (2-20 µm diameter), clay particles are united by inorganic amorphous binding agents such as aluminosilicates, oxides, humic substances and soil polysaccharides associated with metal ions. These persistent microaggregates are bound together into larger microaggregates (20-250 µm diameter) by plant roots, root hairs, and fungal hyphae. Microaggregates are glued to each other by transient binding agents such as polysaccharides and polyuronides to form macroaggregates (>250 µm diameter). Aggregation is influenced by the soil microbial community, mineral and organic compounds, plant community composition and past soil handling (Tisdall & Oades, 1982). For many decades, the microbial communities inside different classes of aggregates have been investigated using several techniques and experimental designs (Blaud et al., 2012, Zhang et al., 2018). Many studies determined the microbial community inside the different aggregate

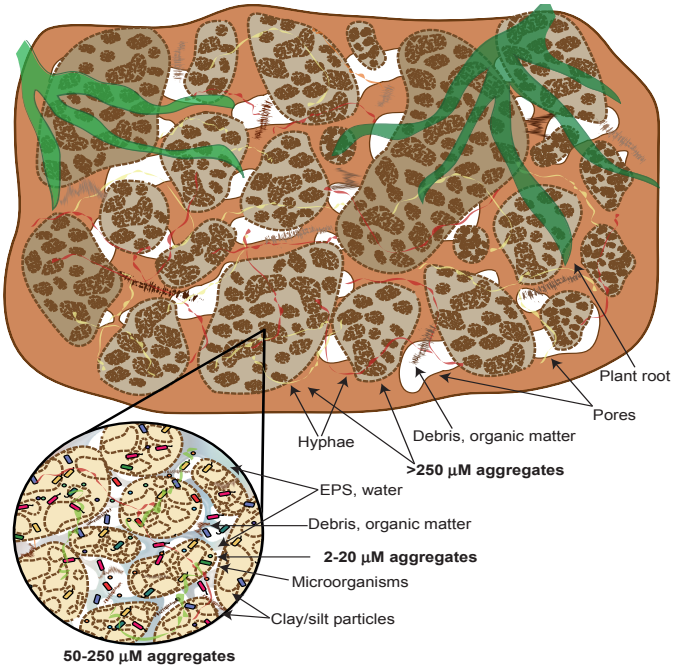

Figure 3: The hierarchical model of soil aggregate classification. Larger aggregates are composed of smaller units, which are formed from even smaller aggregates.

sizes in different agriculture management systems (Sessitsch *et al.*, 2001, Mummey & Stahl, 2004, Kravchenko *et al*., 2014), however, no studies evaluated the microbial community responsible for the aggregation. Studies on microbial effect on soil aggregations were limited to microbial isolated strains, albeit the role of microorganisms and their polysaccharides in soil aggregation have been studied for decades. Cesar Ton-That *et al*. (2007) used microaggregates (250 to 50 µm) from two agricultural ecosystems (40 years tillage and 9 years no tillage) to isolate bacteria and test their aggregation potential, as well as profiled both systems using Fatty Acid Methyl Ester (FAME). They observed that *Stenotrophomonas*, *Sphingobacterium*, *Bacillus* and *Pseudomonas* species could stabilize and increase aggregate strength in artificial aggregates, and that these species were mere frequent in partially undisturbed soils. In other study, Cesar Ton-That *et al*. (2014) investigated if soil aggregation and the culturable aggregating bacteria present in soils were influenced by different irrigation, tillage and cropping systems. In the irrigated no tillage and conservation areas, higher proportion of soil aggregating bacteria were isolated (81, compared to ~35). They were able to isolate 50 aggregating bacteria (from 1296 isolates), which were dominated again by *Pseudomonas* sp. and *Bacillus* sp. Interestingly, *Bacillus* and *Pseudomonas* are genus widely known to produce biofilms and EPS, which are involved in the stabilization of soil structure.

6.2. Inoculation of EPS producers in soils

The role of microorganisms and their polysaccharides in soil aggregation have been studied for decades. Microorganisms are fundamental for soil aggregation. Fungi and bacteria contribute to stabilization of soil structure by producing extracellular polymers and degrading aromatic humic materials that generate clay-metal-organic matter complexes (Umer & Rajab, 2012). Fungi also contribute by anchoring particles through hyphae, albeit with less persistence. However, the influence of microorganisms on soil structure stabilization varies and depends on the microbial species, available substrates and soil management (Beare *et al.*, 1994, Umer & Rajab, 2012). The aggregating potentials of numerous bacterial and fungal strains have been tested, demonstrating that the effect of microbial pure cultures on soil aggregation is dependent on the microbial species. Therefore, different microbial slimes and EPS have been explored as aggregation-capable components in different types of soils, for the recovery of soil quality and fertility.

Among the bacterial EPS producers that are the most investigated for soil aggregation potential are strains of *Pseudomonas*, *Bacillus* and *Paenibacillus* genera easily grown in laboratorial conditions, producing high amounts of EPS. However, other genera strains of *Pseudomonas*, *Streptomyces* and *Penicillium* had shown a significant positive effect on soil loss and erodibility, after rainfall simulation (Gasperi-Mago & Troeh, 1979). *Pseudomonas putida* strain GAP-P45 inoculation in soil increased aggregate stability in more than 50% in soils subjected to temperature, salt and drought stresses. Under stress conditions, the strain produced more EPS, protecting the bacteria against water stress and contributing to soil

structure (Vardharajula & Ali, 2015). An unidentified bacterium isolated from biological soil crusts from the Gurbantünggüt Desert stabilized sand surface, producing aggregation and slowing the soil water evaporation after only 8 days of inoculation. In addition, the EPS of the bacterium produced the conglutination of sand particles, as observed by scanning electron microscopy (HuiXia *et al.*, 2007). Another isolate from Gurbantünggüt Desert, *Paenibacillus* KLBB0001 – a strong EPS producer – was inoculated in the desertic soil to improve the recovery of biological soil crusts (BSC). After one year of field experiments, the strain stimulated the heterotrophic community in the soil and increased the numbers of bacteria, available nitrogen and phosphorus. Microscope images of the inoculation area revealed a glue-like polymer connecting sand grains, confirming the presence of EPS (Wu *et al.*, 2014). The studies showed the potential of the strains for the recovery of soil structure, especially under nutrient- and water-limited conditions.

Due to its high EPS production, *Bacillus amyloliquefaciens* strain HYD-B17, *B. licheniformis* strain HYTAPB18 and *B. subtilis* RMPB44 inoculation in soil improved aggregate stability in both the absence of stress and under drought stress conditions. For these strains, it was also observed a better aggregation effect with a larger bacterial population size, as well as an important role of larger incubation periods for EPS production and soil aggregation. All the strains produced more EPS under drought conditions, and strain HYD-B17 was the most efficient for aggregation among the strains studied. The differences in the performances of the strains could be explained by the different compositions of their EPSs. The performance of the strains demonstrate that they are also interesting for inoculation in situations of abiotic stresses (Vardharajula & Ali, 2014). Strains of *Pseudomonas* and *Bacillus* were also important for the stabilization of sand on the beach and at the edge of a dune in the study of Foster (Forster, 1979).

Microbacterium arborescens-AGSB is another example of an EPS-producer strain that can be used for the recovery of soils; its inoculation produces strong binding in sandy soil. In addition, the bacteria produced better aggregation in a sandy soil than in agricultural and mine reject soils, showing that the effect of microbial inoculation varies according to the soil type (Godinho & Bhosle, 2009).

In addition to other bacterial genera, the inoculation of soil with cyanobacteria has long been proved to be beneficial to soil structure and parameters. These bacteria were recognized as important in the stabilization of soil surfaces, primarily because of EPS production. In arid environments, cyanobacteria are major components of biological soil crusts (BSCs). BSCs are microbial assemblages developed on the top soil of drylands (Malam Issa *et al.*, 1999). They are integral components of arid and semi-arid ecosystems, which biological activities are important for soil fertility and reduction of erosion, influencing soil temperature, C and N content, hydrological dynamics and plant germination (Chamizo *et al.*, 2012, Rossi *et al.*, 2017, Velasco Ayuso *et al.*, 2017). Their main components are species of bacteria, microalgae, fungi, lichen and mosses, but their specific composition is variable (Wu *et al.*, 2014, Mugnai *et*

al., 2017). The use of cyanobacteria for recovery of drylands and BSCs will not be discussed in this review since the focus is in agriculture soils.

Cyanobacterization improves soil structure, fertility, bioavalability of nutrients, benefits that are extended also to the subcrust. Recently, they have been investigated for improvement of quality of arable lands and treatment of degraded and desertified environments (Rossi & De Philippis, 2015). Characteristics such as stress tolerance drought resistance and oligotrophy make them optimal candidates, and their EPS improves soil stability and moisture content at the topsoil, stimulating soil biological activity (Guo *et al.*, 2007). Cyanobacteria exert a mechanical effect on soil particles, as they produce a gluing mesh, binding soil particles with their EPS. They promote the formation of hard entangled superficial structures that improve the stability of semi-arid soil surfaces, protecting them from erosion. In addition, they play a significant role in water storage, because of the hygroscopic properties of the EPS (Mugnai *et al.*, 2017). For instance, the inoculation of *Nostoc muscorum* improved the aggregate stability of a poorly structured silt loam soil in a greenhouse experiment. In this study, the authors investigated the effect of the inoculation of *N. muscorum* on the microbial population, soil nutrient status and fertility. The addition of the microorganism increased soil aggregation by an average of 18%, as well as increased soil total carbon by ~60% and total N by more than 100%; it also increased microbial population numbers and the emergence of lettuce seedlings in more than 52% (Rogers & Burns, 1994). Another strain of the genus *Nostoc* caused a positive impact in the physical characteristics of poorly aggregated soils from Guquka (Eastern Cape, South Africa). A dense superficial network of cyanobacterial EPS filaments covered soil surface after 4 and 6 weeks of incubation. The improvement appeared a short while after incubation, and increased with time and cyanobacteria growth (Malam Issa *et al.*, 2006). Other strains of cyanobacteria, such as *Oscillatoria*, *Lyngbya* and *Schizothrix delicatissima* AMPL0116 also showed positive effects in soil structure, by improving soil hydrological responses to rainfall, soil particle connections, soil permeability and water absorption (Mugnai *et al.*, 2017, Sadeghi *et al.*, 2017).

Inoculation of pure cultures of filamentous fungi is known to increase soil aggregation, however with different effectiveness than that of bacteria. Fungi not only can produce EPS that bind soil particles together but also produce hyphae that can enmesh aggregates (Baldock, 2002). The presence of *Strachybotrys atra* increased the aggregation of fumigated Peorian loess soil. However, the fungus was only able to produce this effect in a situation of reduced microbial community, demonstrating its establish as the dominant microorganism (McCalla *et al.*, 1958). In the study of Aspiras (1971), *Alternaria tenuis*, *Stachybotrys atra*, *Aspergillus niger*, *Mucor hiemalis* and the streptomycetes *Streptomyces purpurascens* and *S. coelicolor* promoted the stabilization of artificial soil particles from 3 different soils. The aggregation was a result of binding agents closely associated to the hyphae. Swaby et al (1949) tested the aggregation capacity of pure cultures of 101 bacteria, 5 yeasts and 50 filamentous fungi, finding that fungi had the best results. Among the best fungi there were species of *Absidia*,

Mucor, *Rhizopus*, *Chaetomium*, *Fusarium* and *Aspergillus*. For bacteria, *Achromobacter*, *Bacilllus* and unclassified *Actinomycetes* had the best aggregation potentials. A saprophytic lignin-decomposer evaluated by Caesar-TonThat & Cochran (2000) was able to aggregate and stabilize sandy soil, producing 90% of water stable aggregates. The fungus excreted insoluble extracellular compounds that acted as binding agents, forming a fibrillary network observed in soil micrographs. *Azotobacter chroococcum*, *Lipomyces starkey* and strains of *Pseudomonas* sp. and *Mucor hiemalis* were also able to promote soil stabilization (Lynch, 1981).

In addition of pure cultures, a combination of microorganisms can be an interesting option for soil inoculation. However, few studies investigate the addition of microbial consortia for improvement of soil aggregation. Nonetheless, when complex mixed cultures of microorganisms are inoculated (Swaby, 1949) in particles, aggregation is maximized as a result of interactions between different strains. Different species have different EPS properties; furthermore, EPS can have a complementary effect when associated with other EPS and other aggregating factors, such as EPS-coated fungal hyphae, resulting in greater adherence of soil particles compared to only physical involvement by the hyphae (Aspiras *et al.*, 1971). Moreover, the combination of organic fertilizers with microbial inoculants can strengthen microbial aggregation effects by enhancing EPS production, consequently improving soil structure, function, and quality (Rashid *et al.*, 2016).

6.3. Plant inoculation with EPS producers

Plant inoculation with Plant Growth Promoter Rhizobacteria (PGPR) and Arbuscular Mycorrhizal Fungi (AMF) is a very important agricultural practice. Microorganisms stablish interactions with plants, promoting plant growth, which stimulates the microbial community with the production of exudates. The organic carbon released by plant roots stimulates the growth of the microbial communities in the rhizosphere, which in turn, produce mucilaginous EPS, promoting soil aggregation and increasing Root Adhering Soil (RAS). RAS aggregation is important because it forms the immediate environment where plants take up water and nutrients for their development. The inoculation of plants with beneficial microbes can, in addition, increase the availability of nutrients, such as N, P, K and iron (Rashid *et al.*, 2016).

Among the best and most investigated bacterial candidates for plant inoculation are strains of *Bacillus*, *Pseudomonas*, *Rhizobium* and *Pantoea*, all known EPS producers and plant growth promoters. These strains can be inoculated directly in soil, or in seedlings, where they will also be beneficial for crop yield (Cipriano et al., 2016). The production of EPS in the rhizosphere of plants protects the environment against drying and fluctuations in the water potential, increasing nutrient uptake by plants and promoting plant growth. It protects seedlings from drought stress and stimulates root exudates. The improvement in aggregation and soil structure improves the growth of seedlings, because it promotes an efficient uptake of nutrients and water (Alami *et al.*, 2000, Bezzate *et al.*, 2000, Vardharajula *et al.*, 2009). Several studies have evaluate the effect of PGPR and AMF, however no focus on soil

aggregation, since the most are focused were on the plant growth.

Rhizobium strain KYGT207, which was isolated from an arid Algerian soil, is a wheat (*Triticum durum* L.) growth promoter and EPS-producing bacterium with significant soil structure-improving capacity. Inoculation of the strain on wheat increased the root-adhering soil dry mass/root dry mass ratio by 137% and enhanced the percentage of water-stable aggregates due to reduction of soil water stress by the EPS (Kaci *et al.*, 2005). Equally significant are the effects of *Pantoea agglomerans* NAS206 and its polymer on the rhizosphere of wheat and on soil aggregation. The strain can colonize the wheat rhizosphere, causing significant aggregation and stabilization of root-adhering soil. It also increased aggregate mean diameter weight, formation of water-stable aggregates (diameter > 0.2 mm) and RAS macroporosity. Thus, *Pantoea agglomerans* NAS206 is an interesting candidate for inoculation, since it can play an important role in regulating water content in the rhizosphere of wheat and improving soil aggregation (Amellal *et al.*, 1998, Amellal *et al.*, 1999).

The levan produced by *Paenibacillus polymyxa* CF43 also has notable effects on the aggregation of soil adhering to wheat roots (Bezzate *et al.*, 2000). The authors tested the role of levan in aggregation of soil adhering to wheat roots by producing a mutant strain. In comparison with the mutant, the wild type EPS producing strain increased the mass of RAS, demonstrating the influence of the EPS in aggregation and suggesting that the production of levan is the main mechanism involved in the improvement of the RAS structuration.

The role of EPS in soil aggregation has also been evaluated under the application of different environmental stresses. Inoculation of chickpea plants (*Cicer arietinum* var. CM-98) with the EPS-producer strains *Halomonas variabilis* HT1 and *Planococcus rifietoensis* RT4 protected the plants from salinity, promoted plant growth and improved soil aggregation in more than 75% under elevated salt stress. These results demonstrated that both bacteria can be applied to enhance plant growth and soil fertility under salinity (Qurashi & Sabri, 2012). In another study, the EPS-producer *Rhizobium* YAS-34 positively affected soil aggregation and water and nitrogen uptake by sunflower plants under normal and water stress conditions. It increased RAS in up to 100%. The strain acted as a plant growth promoter, increasing shoot and root biomass and also increased soil macropore volume. These effects were attributed to EPS production, which increased soil water hold capacity (WHC) and reduced water loss (Alami *et al.*, 2000). The strains of *Bacillus* and *Aeromonas* evaluated by Ashraf et al (2004) increased the aggregation around roots of wheat in a moderate saline soil, restricting Na^+ uptake by plants and promoting plant growth.

The effects of plant inoculation of several fungi have also been extensively evaluated, also with more focus on plant growth promotion than in rhizosphere soil aggregation. The mechanisms involved in the aggregate stabilization by fungi are entanglement of the soil particles by hyphae as well as the production of EPS. AMF also produce glomalin, a glycoprotein that acts as a glue (Kohler *et al.*, 2006). Forster and Nicholson (1981) examined the effects of the interactions among grass (*Agropyron junceiforme*) and microorganisms

(*Penicillium* sp and *Glomus fasciculatus*) in the aggregation of sand from an embryo dune. Experiments showed that the addition of selected microorganisms increased both plant growth and soil aggregation. Even though roots alone affected sand aggregation, the best results were due to the association of microorganism inoculation and plants.

The mycorrhizal inoculation of *Olea europaea* and *Rhamnus lycioides* with *Glomus intraradices* showed beneficial effects for rhizosphere aggregation. Together with the addition of composted residue, inoculation increased rhizosphere aggregation in comparison with non-rhizosphere soil by 1.8 fold (Caravaca *et al.*, 2002). The effects of the inoculation of *Glomus intraradices*, and *Pseudomonas mendocina* were evaluated by Kohler (2006) in lettuce. The inoculation of both strains increased the percentage of water soluble carbohydrates and stable aggregates. *P. mendocina* also had a positive effect on soil enzymatic activities, such as dehydrogenase and phosphatase. The combination of *P. mendocina* with inorganic fertilization increased stable aggregates in 84% compared to the control.

Inoculated microorganisms can have a significant effect on soil properties and quality by interacting with natural microorganisms in the rhizosphere, in addition to the improvement of plant productivity. Good soil structure and aggregate formation are important for controlling germination and root growth. Microbial inoculants have been studied for decades, but there is still a need for the enhancement of microbial growth conditions, for the production of high quality inoculants, with higher biomass and EPS production. Therefore, strains will be able to have an optimal performance in field conditions, with efficient colonization and dominance over the native microbial community.

6.4. Addition of pure EPS to soil

Several studies link microbial products to soil aggregate stability. It has been long know that polysaccharides, are involved in the maintenance of soil structure, even though they are not the primary aggregating agents. Other molecules, such as humic acids are also responsible for soil structure. The treatment of natural and synthetic soil aggregates with various chemical substances, such as periodate and tetraborate frequently does not result in a consistent pattern, demonstrating that polysaccharides are important, but more than one single substance are the main factors sustaining soil aggregates (Mehta *et al.*, 1960, Sparling & Cheshire, 1985). Angers and Mehuys (1989) observed that the correlation between aggregate mean weight and carbohydrate content suggested that at least part of the water-stable aggregation was related to carbohydrates in soils . Treatment of the soil with sodium periodate prior to wet sieving confirmed partial involvement of carbohydrates in the stabilization of aggregates by crops.

The resistance of the biopolymers to degradation may be related to its importance for the soil structure. The greater the resistance, the longer is its persistence in soils. The association of polymers with metal ions and colloids, such as clay may also influence the degradation rates of polymers, because of their influence in enzymatic activity. Since the addition of polymers

to soil started to be investigated, it has been demonstrated that the binding power of plant and microbial polysaccharides is variable. However, characteristics of the soil such as pH also influence the action of polysaccharides, because the charges of molecules are essential for binding particles (Martin, 1971). Some characteristics of polysaccharides that influence their binding activity are linear structure and length and flexible nature, that allow the formation of Van der Waals forces; large number of OH for hydrogen bonding and presence of acyl groups, allowing ionic binding to clays (Martin, 1971).

The effects of many different EPS produced by fungi and bacteria were already tested as soil aggregating agents. The direct application of polymers in soil can be an alternative to the inoculation of microorganism. The aggregating potential of the EPS of *Bacillus subtilis*, *Leuconostoc dextranicum* and *L. mesenteroides* were evaluated by Geoghegan and Brian (1948). The different EPS had a significant in soil aggregation tested by wet sieving, and even small amounts of levan (0.125 to 0.05%) were able to stabilize aggregates. The EPS of *Chromobacterium violaceum* had also an interesting effect in soil, being more resistant to degradation than a variety of plant polysaccharides. It exhibited the best binding performance among all polysaccharides tested, improving the hydraulic conductivity of a soil with neutral pH (Martin & Richards, 1963).

Some EPS molecules have a very high WHC. A xanthan tested by Chenu and Roberson (1996) demonstrated a WHC of 15 times its weight. The dextran tested in the same study had a lower WHC, due to differences in structure. For both EPS, diffusion of glucose was tested, and it was observed that diffusion rates were slower than in water. A high WHC of EPS can protects microorganisms, soil and plants against drought stress, promoting hydrating conditions and bridging among soil particles and clay. In addition, the nutrients are still able to diffuse until the microorganisms during low water potentials, maintain physiological functions even during dry periods. The EPS of a *Pseudomonas* strain isolated from soil can also hold several times its weight in water. When added to a sandy soil, the EPS altered its moisture by allowing the amended soil to hold more water than unamended soil. The addition of a small amount of EPS increased the amount of water held by the sand (Roberson & Firestone, 1992).

There are evidences that xanthan stabilizes soil against disruptive effect of wetting and drying cycles (Czarnes *et al.*, 2000). In comparison with control soil and dextran, soils amended with xanthan were less sensitive to this kind of stress. Differences in structure of both polysaccharides could explain their different behaviors. Rosenzweig et al. (2012) also tested the WHC of two sandy soils amended with xanthan, and observed that the addition of >1% xanthan increased dramatically the water holding capacity of the soil, as well as soil porosity.

Many of the studies that evaluate the application of microbial biopolymers in soil are in the engineering area. There are several studies that evaluate the application of microbial biopolymers and plant polymers, such as guar gum and cellulose for stabilization and soil binding for constructions. Such studies in the engineering area also confirm the usefulness of biopolymers application in soil, but with different purposes.

The strength of biopolymers can be observed by their application in the fields of construction and geotechnical engineering, as soil binders (Chang *et al.*, 2017). The commercial polymer from *Aureobasidium pullulans* was efficient in the treatment and stabilization of a residual Korean soil, increasing the compressive strength of soil more than 200% (Chang & Cho, 2012). It was considered an economically competitive and environmentally friendly alternative for soil binding. In another study, a very small amount of microbial EPS (such as xanthan and gellan gum 0.5%) mixed with soil resulted in a higher compression strength in comparison to the addition of a large amount of cement. Xanthan forms connection bridges between particles, enhancing particle alignment, improving strength. The effect is a result of the matrix strength and electrostatic bonds between xanthan and fine soil particles. These polymers can be naturally decomposed, not requiring construction demolition. They are promising for construction as building materials (Chang *et al.*, 2015). The application of xanthan gum can also be used to treatment of collapsible soil, reducing collapsible potential (Ayeldeen *et al.*, 2017).

In addition to the direct application of EPS to soil, there are evidences that EPS production in soil can be modulated by N management. Roberson et al. (1995) evaluated the effect of the N addition in EPS production and soil aggregation, by indirect measurements, carbohydrate content and monosaccharide composition. While intermediate and high amount of N fertilization gave similar crop yield, the soil properties had different results. Intermediate N fertilization induced better aggregation, saturated hydraulic conductivity and the monosaccharide composition was more related to microbial polysaccharide. Therefore, the addition of nutrients could also induce EPS production directly in soil, consequently improving soil aggregation.

Many studies have demonstrated that EPS production can increase soil aggregation, improve soil quality and contribute to soil fertility. Moreover, in addition to improving soil structure, the presence of EPS in soil and in plant roots can improve nutrient uptake and water availability for both plants and microorganisms, thus benefiting not only the producer but also the environment as a whole. Several works show that both bacteria and fungi are important for soil aggregation, Their EPS are capable of binding soil particles, and their interactions, as well as their interaction with plants, and the addition of organic fertilizers altogether are enhancers of soil structure and stability. Microorganisms have an enormous potential, which can be enhanced by the improvement of the knowledge of the structure of EPS, as well as the development of microbial consortia and large-scale EPS production.

7. Conclusion and perspectives

Microorganisms have developed several approaches to survive environmental conditions, especially in soils. EPS production is an important strategy for providing a moist environment, entrapping nutrients, facilitating chemical reactions, and protecting cells

against environmental conditions, antibiotics and attack by predators. Microbial extracellular polymers are highly diverse compounds with multiple functions that depend on their composition and structure.

EPS have long been of interest due to their biodegradability, biocompatibility, and thickening, gelling and emulsion capacities. The polymers and their production can be manipulated to achieve high yields, but such manipulations are dependent on the characterization and physiological study of EPS-producing microorganisms. Improving polymer production requires an understanding of the underlying mechanisms and regulatory pathways. In contrast to the intensive work focused on improving EPS yield and altering the characteristics of well-known polymers, novel EPS and polymers produced by less-studied microbial strains are still underexplored. The investigation of the genetic mechanisms involved in the biosynthesis of any type of molecule involves complex and time-consuming techniques, and thus the development of knowledge in this area may proceed slowly. Many microorganisms produce EPS, and because each polymer is different, many opportunities remain for investigation and discovery.

EPS are complex substances and our understanding of their composition, structures, functions and genetic regulation, although very broad, is far from complete. There is a need for a fundamental understanding of the genes and mechanisms involved in the biosynthesis and regulation of EPS. Furthermore, the discovery and characterization of new polymers could lead to interesting other applications, especially for the environment. EPS can be employed in wastewater treatment, recovery of polluted environments, and, potentially, in the recovery of soil aggregation and improvement of soil fertility. Advances in modern techniques and approaches, such as high-throughput sequencing, confocal laser scanning microscopy, nuclear magnetic resonance and scanning electronic microscopy, stable isotope probing in association with classic microbiology techniques will enhance efforts to discover and characterize new EPS and their functions in the soil ecosystem. The understanding of structure and properties of EPS is fundamental for understanding their interactions with soil. The combination of classic microbiology techniques with modern high-throughput methods and integration of different fields are fundamental for increasing knowledge on EPS composition, structure and function and applications.

Acknowledgments

O.Y.A. Costa was supported by an SWB grant from CNPq [202496/2015-5] (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

Chapter 3

Transcriptional and proteomic responses of Granulicella sp. WH15 to increasing concentrations of cellobiose

Ohana Y.A. Costa[#], Marcelo M. Zerillo[#], Daniela Zühlke, Anna M. Kielak, Agata Pijl, Katharina Riedel, Eiko E. Kuramae

[#] equal contribution

Modified version published as: Costa OYA, Zerillo M, Zühlke D, Kielak AM, Pijl A, Riedel K, Kuramae EE (2020). Responses of *Acidobacteria Granulicella* sp. WH15 to High Carbon Revealed by Integrated Omics Analyses. **Microorganisms**. 8(2).

Abstract

The phylum *Acidobacteria* is widely distributed in soils, but few representatives have been cultured. In general, *Acidobacteria* are oligotrophs and exhibit slow growth under laboratory conditions. We sequenced the genome of *Granulicella* sp. WH15, a strain obtained from decaying wood, and determined the transcriptome and proteome when grown in poor medium with a low or high concentration of cellobiose. We detected the presence of 217 carbohydrate-associated enzymes in the genome of strain WH15. Integrated analysis of the transcriptomic and proteomic profiles suggested that high concentration of cellobiose triggered the expression of stress-related proteins. As part of this response, transcripts related to cell wall stress, such as sigma factor σW and toxin-antitoxin (TA) systems, were upregulated, as were several proteins involved in detoxification and repair, including the multidrug resistance protein MdtA and outer membrane protein OprM. KEGG metabolic pathway analysis indicated repression of carbon metabolism upon high cellobiose concentration, especially the pentose phosphate pathway, and repression of protein synthesis, carbohydrate metabolism and cell division, suggesting arrest of cell activity and growth. In summary, the stress response of *Granulicella* sp. WH15 induced by the presence of a high cellobiose concentration in the medium resulted in the enhanced expression of functions associated with secretion of metabolic byproducts and with reallocation of resources to cell maintenance instead of growth.

Keywords: Genome, Transcriptome, Proteome, Oligotroph, stress signal, transporters, sigma factor σW

1. Introduction

Acidobacteria is one of the most abundant bacterial phyla in soil, yet little is known about its physiology, ecological function, and impact on the soil environment (Kielak *et al.*, 2016). The first species of *Acidobacteria* was described in the 1990s (Kishimoto *et al.*, 1991), and the ubiquity of the phylum was only recognized after the introduction of 16S rRNA gene sequencing and metagenomics (Kielak *et al.*, 2016). This phylum constitutes 20–50% of the soil bacterial community (Kuramae *et al.*, 2012, Navarrete *et al.*, 2013, Pan *et al.*, 2014, Kielak *et al.*, 2016, Kielak *et al.*, 2016), but the few species that have been isolated exhibit slow growth under standard laboratory conditions, resulting in a relatively small number of cultured representatives. Genome analyses have revealed only one or two copies of the 16S rRNA gene in species sequenced to date (Wingender *et al.*, 1999, Pankratov & Dedysh, 2010, Whang *et al.*, 2014), which may indicate slow growth of these bacteria under natural conditions as well. Ribosomal RNA operon copy number has been potentially linked to low environmental resource availability and slow growth rate in bacteria (Valdivia-Anistro *et al.*, 2016). Consequently, the factor(s) responsible for the prevalence and successful adaptation of *Acidobacteria* to soil conditions remain unclear.

A strong negative correlation between *Acidobacteria* abundance based on 16S rRNA amplicon next generation sequences and soil organic carbon content has been observed in diverse microbiome studies, suggesting that the phylum is composed of oligotrophic bacteria (Fierer *et al.*, 2007, Foesel *et al.*, 2014). Oligotrophs are mainly characterized by their capacity to grow under low nutrient availability and their higher substrate utilization efficiency than copiotrophs, producing a higher biomass yield per unit of substrate utilized. In general, they are able to thrive in poor nutrient environments and exhibit slow growth under laboratory conditions (Ho *et al.*, 2017). Although most acidobacterial isolates have been obtained from low-nutrient culture media, some isolates are capable of growing at higher sugar concentrations (de Castro *et al.*, 2013, Campanharo *et al.*, 2016, Kielak *et al.*, 2016).

Many soil *Acidobacteria* are able to degrade a wide range of carbon sources, mainly mono- and disaccharides, such as glucose, xylose, mannose, galactose and cellobiose. Predictions of genes associated with the degradation of polysaccharides in acidobacterial genomes have not always been confirmed by experimental data (Kielak *et al.*, 2016). This knowledge gap highlights the need for studies with cultured strains. We recently established a culture medium and incubation conditions permitting larger amounts of acidobacterial biomass (cells and/or exopolysaccharides, EPS) to be harvested after 4 days of incubation (Campanharo *et al.*, 2016); by contrast, on other reported media, most cultivated acidobacterial species form visible colonies only after weeks of incubation (Ward *et al.*, 2009). As no study has addressed the response of acidobacterial strains under different cellobiose concentrations, the aim of this study was to sequence the genome of an acidobacterial strain, *Granulicella* sp. WH15, and determine the transcriptome and proteome responses under conditions of low (0.025%) and high (3%) cellobiose concentrations.

2. Material and Methods

2.1. Genome

The *Granulicella* sp. WH15 strain (Valášková *et al.*, 2009) obtained from the collection of the Netherlands Institute of Ecology (NIOO-KNAW) was grown on 1/10 TSB agar medium (Valášková *et al.*, 2009) at pH 5.0 for 3 days at 30 ˚C. The bacterial cells were harvested and the genomic DNA was extracted using MasterPure™ DNA Purification Kit (Epicentre, Madison, WI) according to manufacturer's instructions. A total of 10 mg of DNA was sent to the Genomics Resource Center (Baltimore, USA) for a single long insert library (15kb-20kb), that was constructed and sequenced in one SMRTcell using the PacBio RS II (Pacific Biosciences, Inc.) sequencing platform. *De novo* assembly was performed with the help of SMRT Analysis software v2.2.0 (Pacific Biosciences) featuring HGAP 2 (Chin *et al.*, 2013), and subsequent correction with Pilon 1.16 (Walker *et al.*, 2014) to reveal a circular replicon: a 4,675,153-bp chromosome (G+C content 60,7%; 58.4× coverage). Automatic gene prediction and annotation was performed by using Prokka (Seemann, 2014), RAST genome annotation server (http://rast.nmpdr.org/) (Aziz *et al.*, 2008). Genes were mapped to COG and KEGG IDs using the COG database (2014 release) (Galperin *et al.*, 2015) and KEGG database (release 2013) (Kanehisa, 2000), using in house scripts. The CAZyme contents of WH15 genome were determined by identifying genes containing CAZyme domains using the dbCAN2 meta server (cys.bios.niu.edu/dbCAN2) (Zhang *et al.*, 2018), according to the CAZy (Carbohydrate-Active Enzyme) database classification (Lombard *et al.*, 2014). Only CAZyme domains predicted by at least two of the three algorithms (DIAMOND, HMMER and Hotpep) employed by dbCAN2 were kept. Circular genome map was drawn used CGView software (Stothard & Wishart, 2004). The *Granulicella* sp. WH15 strain genome data is deposited at NCBI with accession number CP042596.

2.2. Growth experiments

CAZyme annotation was compared to *in vitro* carbohydrate utilization assays. Modified 1/10 TSB (1.7 g of casein hydrated (acid) (Oxoid™), 0.3 g of tryptone (vegetable) (Fluka 95039), 0.5 g of sodium chloride, 0.25 g of dipotassium phosphate (Sigma-Aldrich) in 1 litter of distilled water) pH 5.0 medium was supplemented with sole carbon sources: pectin (apple – Sigma- Aldrich), glycogen (Sigma- Aldrich), glucosamine (Sigma- Aldrich) and cellulose (Sigma- Aldrich), at 1% (w/v), and D-glucose, D-galactose, D-mannose, D-xylose, L- arabinose, L-rhamnose, D-galacturonic acid, cellobiose, D-lactose, and sucrose at 25 mM. Plates were inoculated with 100 µL of cell suspension at OD_{600nm}= 1 obtained from modified TSB (3% of cellobiose), homogenized and incubated at 30 °C for 72 h, when growth was evaluated. After 72 h of growth, colonies were reinoculated in a new plate with the same carbohydrate concentration, and the process was repeated three times, in order to confirm growth.

2.3. Transcriptome

Bacteria were grown in two sugar concentrations, based on solid (1.5% agarose) modified 1/10 TSB medium pH 5.0 (as described above) supplemented with 0.025% and 3% of cellobiose, for low and high cellobiose conditions, respectively. Plates were inoculated with 100 µL of cells at OD_{600nm}= 1 nm/mL, homogenized and incubated at 30 °C for 72 h. RNA extraction was performed using Aurum Total RNA kit (Bio-Rad) in a final volume of 50 µL of elution buffer. Total of 6 mg of RNA was rRNA depleted with the Ribo-Zero Bacteria Kit from Illumina. The sequencing library of each sample was prepared with the Kapa Biosystems Stranded RNA-Seq kit, and sequenced according to the Illumina TruSeq v3 protocol on the HiSeq2000 with a single read 50 bp and 7bp index at Erasmus Center for Biomics (Rotterdam). Quality of the raw fastq files was checked and filtered with FastQC. Read alignment was performed on the genome of WH15 using Bowtie2 (Langmead & Salzberg, 2012). Differential expression profiling was performed between low and high cellobiose treatments. Gene expression levels were quantified using software package RNASeq by Expectation Maximization (RSEM) (Li & Dewey, 2011). The matrix of fragment counts from each sample was used for the differential expression analysis via edgeR package. All samples were normalized by trimmed mean of M-values (TMM). (Robinson *et al.*, 2009). Differentially expressed genes between the two treatments, using low cellobiose treatment as a reference, were identified at significance level of 0.05 with a false discovery rate (FDR) correction method and the $\log_2$ fold change (logFC) equaled to 1. 1% differentially expressed genes were selected to generate a heatmap. COG (Galperin *et al.*, 2015) and KEGG (Kanehisa, 2000) analysis were performed for differentially expressed genes. Transcriptome data is deposited at NCBI GEO database with accession number GSM4017160-65.

2.4. Analysis of the cytosolic proteome by mass spectrometry and data analysis

Bacteria were grown in the same low cellobiose (0.025% cellobiose) and high cellobiose (3% of cellobiose) conditions for transcriptome analyses. Bacterial biomass from each low and high cellobiose treatment were collected from plates and resuspended in 1 ml TE buffer. Bacterial cells were harvested by centrifugation 10,015 x g at 4 °C for 10 min. Pellets were washed twice with 1 mL of TE buffer and finally resuspended in 1 ml TE buffer. 500µL of cell suspension were transferred into 2 mL screw cap tubes filled with 500 µL glass beads (0.1 mm in diameter; Sarstedt, Germany) and mechanically disrupted using Fastprep (MP Biomedicals) for 3 x 30 sec at 6.5 m/s; with on ice incubation for 5 min between cycles. To remove cell debris and glass beads samples were centrifuged for 10 min at 4 °C at 21,885 x g, followed by a second centrifugation (30 min at 4 °C at 21,885 x g) to remove insoluble and aggregated proteins. The protein extracts were kept at -20 °C. Protein concentration was determined using RotiNanoquant (Carl Roth, Germany). Proteins were separated by SDS-PAGE. Protein lanes were cut into ten equidistant pieces and in-gel digested using trypsin as described earlier (Grube *et al.*, 2014). Tryptic peptides were separated on an EASY-nLC II

coupled to an LTQ Orbitrap Velos using a non-linear binary 76 min gradient from 5 – 75 % buffer B (0.1 % acetic acid in acetonitrile) at a flow rate of 300 nL/min and infused into an LTQ Orbitrap Velos (Thermo Fisher Scientific, USA) mass spectrometer. Survey scans were recorded in the Orbitrap at a resolution of 60,000 in the m/z range of 300 – 1,700. The 20 most-intense peaks were selected for CID fragmentation in the LTQ. Dynamic exclusion of precursor ions was set to 20 seconds; single-charged ions and ions with unknown charge were excluded from fragmentation; internal calibration was applied (lock mass 445.120025). For protein identification resulting MS/MS spectra were searched against a database containing protein sequences of *Granulicella* sp. strain WH15 and common laboratory contaminants (9,236 entries) using Sorcerer-Sequest v.27, rev. 11 (Thermo Scientific) and Scaffold v4.8.4 (Proteome Software, USA) as described earlier (Stopnisek *et al.*, 2016). Relative quantification of proteins is based on normalized spectrum abundance factors (NSAF;(Zhang *et al.*, 2010)). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol *et al.*, 2019) partner repository with the dataset identifier PXD015715. Statistical analysis was done using MeV (Saeed *et al.*, 2003); t-test was applied for proteins that were identified in at least two replicates of the respective condition. Hierarchical clustering and t-test of z-transformed normalized data were performed with the following parameters: unequal group variances were assumed (Welch approximation), P-values based on all permutations with P=0.01, significance determined by adjusted Bonferroni correction. Only significantly changed proteins showing at least two-fold changes between conditions were considered for further analysis. Furthermore, so-called on/off proteins, that were only identified in one condition were analysed. Functional classification of *Granulicella* sp. strain WH15 proteins was carried out using Prophane software (www.prophane.de; (Schneider *et al.*, 2011)), COG (Galperin *et al.*, 2015) and KEGG databases (Kanehisa, 2000). Voronoi treemaps were generated with Paver software (Decodon GmbH, Germany).

3. Results

3.1. Genome annotation and CAZymes

The assembled genome of *Granulicella* sp. WH15 is 4,675,153 bp, with 60.7% GC content, 3,849 proteins and only one rRNA operon. Functional annotation using COG (Cluster of Ortholog Groups) and RAST analyses resulted in the classification of 2,620 genes into 23 COG functional groups and the annotation of 1,456 genes to RAST subsystems. The properties and distribution of genes into COGs/RAST functional categories are listed in Table 1 and Figure 1, respectively. A circular genome map of WH15 is depicted in Figure 2.

Table 1. Genomic statistics from *Granulicella* sp. WH15 genome.

Genome	***Granulicella* sp. WH15**
Size (bp)	4,673,153
G+C content (%)	60.7
Number of coding sequences	3,939
Number of features in Subsystems	1,456
Number of RNA genes	51
Number of contigs	1

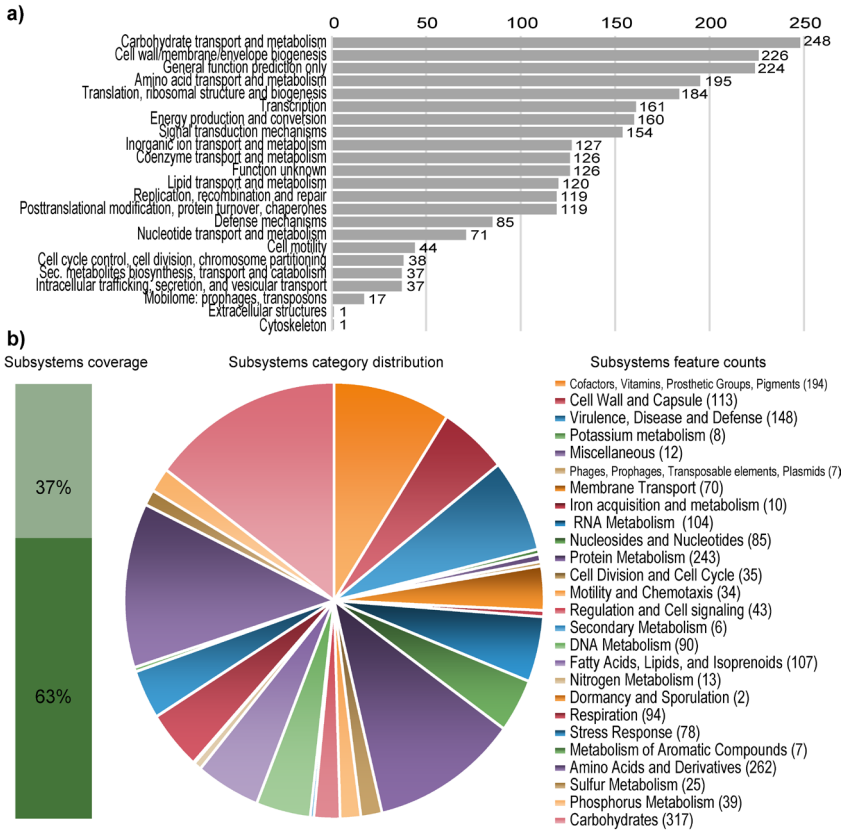

Figure 1: Statistics of the COG and RAST subsystem annotations of *Granulicella* sp. WH15. a) COG category distribution showing the number of genes annotated in each category. b) Subsystem category distribution. The light green bar represents the percentage of proteins that could be annotated by the RAST Server, and the dark green bar represents the proteins that were not annotated. The pie chart represents the percentage of proteins annotated to each subsystem category.

RAST analysis showed that only 37% (1456/3939) of the annotated genes were assigned to subsystems. Among the subsystem categories present in the genome, carbohydrates and dormancy and sporulation had the highest and lowest feature counts, respectively (Figure 1). A comparison with five publicly available genomes of *Granulicella* strains showed that the

genome of strain WH15 is larger only than that of G. *pectinivorans* (Table 2).
We also performed automatic annotation followed by manual curation of carbohydrate-active enzymes (CAZymes) by using a dbCAN2 search, which revealed 217 CAZyme genes in the strain WH15 genome (E-value <10^{-05}) (Table 3). This value is similar to that of other soil *Acidobacteria* Gp1 species determined using the same parameters (Eichorst *et al.*, 2018).
Our annotation revealed a poor gene apparatus of strain WH15 for pectin degradation, in contrast to most sequenced soil isolates of Acidobacteria. Based on the presence of polygalacturonase genes (GH28), strain WH15 might be capable of breaching the complex heteropolysaccharides of pectins but lacks key genes for breaking the pectin backbone (PL1 and PL4), genes targeting D-xylose substitutions and side chains (GH53, GH54, GH93 and GH127) or genes for further pectin saccharification (GH88 and GH105).

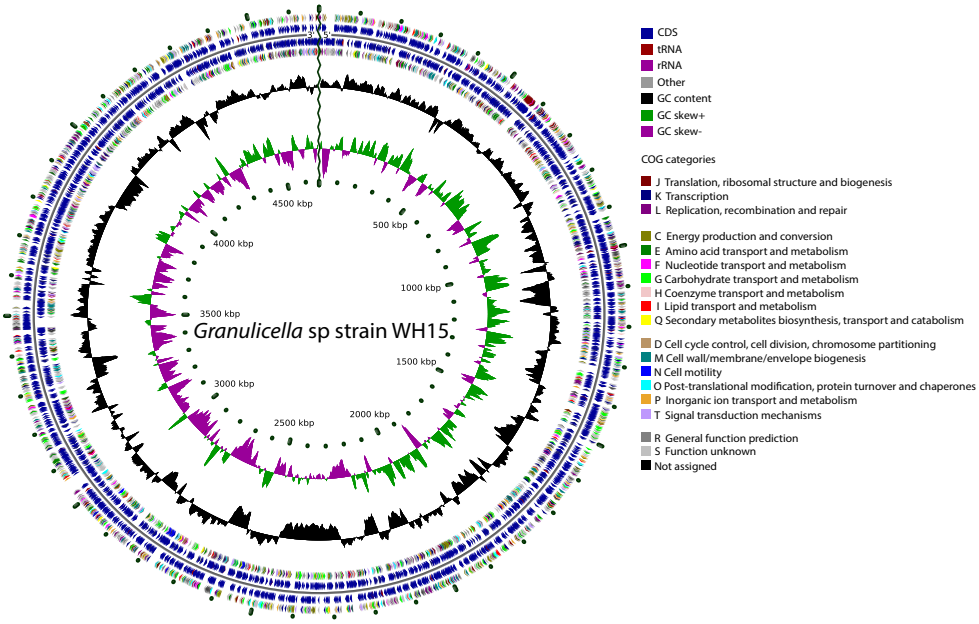

Figure 2: Graphical circular genome map of *Granulicella* sp. WH15. The rings indicate coding sequences, COG categories, GC content and GC skew.

Table 2: Comparison of coding sequences, RNA and subsystems among *Granulicella* species with complete genomes sequences available at NCBI.

Genome	Size (bp)	G+C content	Number of coding sequences	Number of features	Number of subsystems	Number of RNAs
Granulicella spp. WH15	4,673,153	60.7	3871	1496	374	51
Granulicella pectinivorans[1]	4,439,413	61.2	3681	1066	302	41
Granulicella mallensis MP5ACTX8 [2]	6,237,577	57.9	5008	1662	394	50
Granulicella tundricola MP5ACTX9[3]	5,503,984	60	4730	1526	361	49
Granulicella rosea DSM 18704[1]	5,293,785	62.9	4515	864	250	55

[1]:(Pankratov & Dedysh, 2010); [2]:(Rawat *et al.*, 2013); [3]:(Rawat *et al.*, 2013).

Table 3: CAZYme families observed in the genome of *Granulicella* sp. strain WH15.

CAZYme family	Counts
Auxiliary activity (AA)	13
Carbohydrate binding module (CBM)	22
Carbohydrate esterase (CE)	41
Cohesin	1
Glycoside hydrolase (GH)	86
Glycosyl transferase (GT)	52
Polyssacharide lyase (PL)	2
Total	**217**

Accordingly, we did not observe growth of strain WH15 on media supplemented with pectin or galacturonic acid, but growth was obtained on media containing sucrose, glucose, cellobiose, xylose, arabinose, mannose, rhamnose, galactose, or lactose (Table S1). These results are consistent with the presence of α-L-rhamnosidase and β-galactosidase genes in the WH15 genome. In addition, WH15 genome contained *acsC*, *acsAB* and *bcsC*, genes involved in cellulose biosynthesis.

Analysis with the ANTISMASH 4.2.0 database revealed the presence of 33 biosynthetic gene clusters, including 11 of defined type (Table 4) and 22 classified as putative (not shown). Among the putative clusters, 2 have similarities with known clusters involved in the production of O-antigen and thuggacin (Table 4). In addition, other identified clusters showed potential for the production of terpenes, bacteriocins, polyketide synthases (PKS), fatty acids, lasso peptides and saccharides.

Table 4: Biosynthetic gene clusters in the genome of *Granulicella* sp. WH15 reveled by analysis with ANTISMASH.

Cluster	Type	Most similar known cluster
Cluster2	putative	O-antigen BGC (19% of genes show similarity)
Cluster 3	Cf saccharide	
Cluster4	t3pks-cf fatty acid	
cluster 7	Cf saccharide	
cluster 11	lassopeptide	
cluster 13	bacteriocin	
cluster 16	Cf saccharide	
cluster 17	Cf fatty acid-terpene	Malleobactin BGC 11% of genes show similarity)
cluster 19	t3pks	
cluster 22	terpene	
cluster 26	bacteriocin	
cluster 27	Cf saccharide	
Cluster 28	putative	Thuggacin BGC (15% of genes show similarity)

3.2. Transcriptome analysis

The total numbers of reads in the low and high cellobiose conditions are listed in Table S2. In total, 106 (53 upregulated and 53 downregulated) genes were significantly differentially expressed (p-value <0.05) in both treatments, of which only 44 could be annotated, reflecting current gaps in knowledge of the genomes of Acidobacteria in general. Gene expression analysis of Granulicella sp. WH15 grown in low and high cellobiose conditions showed that 28 genes were upregulated at log2 >1.0-fold in the high cellobiose treatment, of which 17 were induced at log2 >1.5-fold (Table S3). In addition, 30 genes were downregulated at log2 >1.0-fold in the high cellobiose treatment, of which 12 were repressed at log2 >1.5-fold (Figure 3) (Table S3). The comparison between the treatments demonstrated that the high cellobiose condition mainly induced the expression of genes related to stress, as well as several unknown hypothetical proteins. Among the annotated genes, a transfer RNA (tRNA-Asn-GTT) was the most upregulated (2.7-fold). In addition, the expression of the genes *gfo4* (glucose-fructose oxidase), tRNA-Asp (GTC), *higA* (HigA antitoxin) and *lytR* (sensory transduction protein LytR) was upregulated (>1.5-fold) in the high cellobiose condition, and the expression of genes coding for the stress response sigma factor SigW (σW), peroxiredoxin TsaA and toxin HigB-1 was upregulated >1.0-fold. By contrast, the high cellobiose condition suppressed the expression of genes related to cell cycle/division and energy metabolism. Downregulation of expression (<1.5-fold) was observed for the genes *trpF* (N-(5'-phosphoribosyl) anthranilate isomerase) and *ssrA* (transfer-messenger RNA), which are involved in amino acid/protein synthesis; two copies of *xerC* (tyrosine recombinase XerC) and the 16S ribosomal RNA gene, which are involved in cell division; *tpiA* (triosephosphate isomerase) and *glpF* (glycerol uptake facilitator protein), which are involved in carbohydrate transport/metabolism; and *hspA* (spore protein SP21), a heat shock chaperone. The expression of genes encoding the housekeeping sigma factor RpoD, 6-phosphogluconate dehydrogenase (*gndA*), RNase P and Lon protease 2 (*lon2*) was also significantly downregulated (>1.0-fold). COG category

assignments of the differentially expressed transcripts are presented in Figure 4a. All of the significantly up- and downregulated genes (>1.0 fold) and their annotations are described in Table S3. Among the upregulated transcripts, two genes could be assigned to KEGG pathways: *tsaA* (K11188 - metabolic pathways, biosynthesis of secondary metabolites and phenylpropanoid biosynthesis) and *lytR* (K07705 - two-component system) (Figure 4a). Of the downregulated transcripts, 5 genes were linked to metabolic pathways in the KEGG database: *trpF* (5 pathways), *tpiA* (10 pathways), *dnaK* (1 pathway), *trpC* (5 pathways) and *gndA* (8 pathways) (Figure 4b). These genes are mostly related to the production of antibiotics and secondary metabolites, as well as amino acid and carbon metabolism. The metabolic pathways for each gene are described in Table 5.

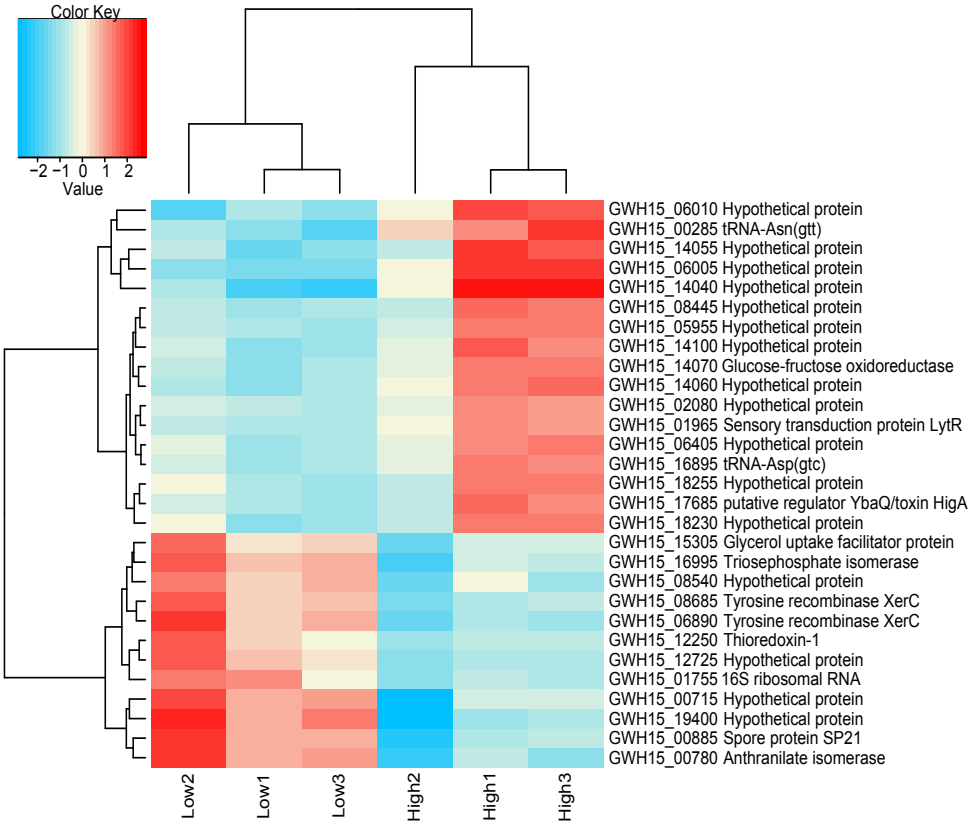

Figure 3: Heatmap of transcriptome data from *Granulicella* sp WH15 showing the top differentially expressed genes (log_2 >1.5-fold) and hierarchical clustering analysis in high cellobiose versus low cellobiose concentrations. High to low expression is shown by a gradient color from red to blue, respectively. Low – 0.025% cellobiose concentration. High – 3% high cellobiose concentration. Differentially expressed genes between the two treatments, using low cellobiose treatment as a reference, were identified at significance level of 0.05 with a false discovery rate (FDR) correction method. Each treatment had three replicates.

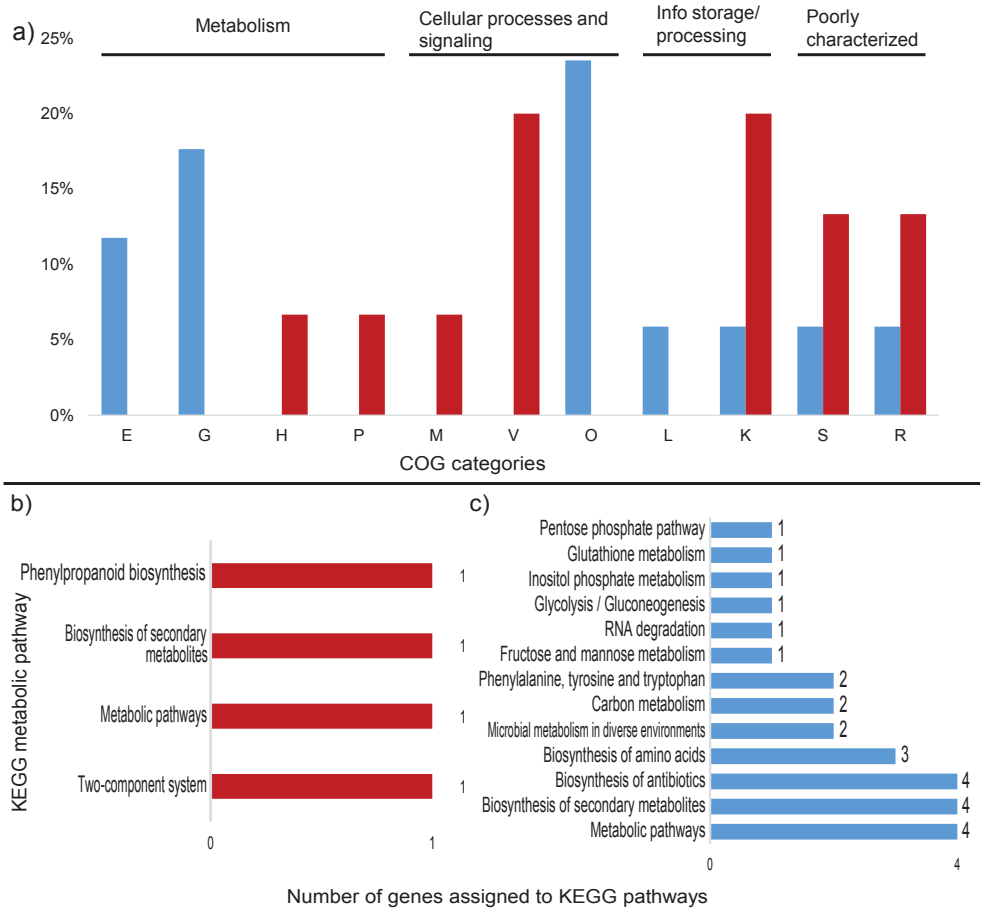

Figure 4: Differentially expressed transcripts assigned to COG categories and KEGG pathways in the transcriptomic profile of *Granulicella* strain WH15, up- and downregulated in high cellobiose treatment. a) Percentage of upregulated (red) and downregulated (blue) transcripts assigned to COG categories; b) Number of upregulated transcripts assigned to KEGG pathways c) Number of downregulated transcripts assigned to KEGG pathways. Unassigned transcripts are not shown. E-Amino acid transport and metabolism; G- Carbohydrate transport and metabolism; H-Coenzyme transport and metabolism; M-Cell wall/membrane/envelope biogenesis; V-Defense mechanisms; P-Inorganic ion transport and metabolism; O-Post-translational modification; L-Replication, recombination and repair; K-Transcription; S-Function unknown; R-General function and prediction.

Table 5: KEGG metabolic pathways assigned to significantly downregulated genes of *Granulicella* sp WH15 in high cellobiose treatment.

KEGG onthology	Gene	Related KEGG pathways
K01817	*trpF*	Biosynthesis of antibiotics, secondary metabolites and aminoacids; metabolic pathways; phenylalanine, tyrosine and tryptophan biosynthesis.
K01803	*tpiA*	Biosynthesis of antibiotics, secondary metabolites and aminoacids; metabolic pathways; phenylalanine, tyrosine and tryptophan biosynthesis; microbial metabolism in diverse environments; carbon metabolism; inositol phosphate metabolism; fructose and mannose metabolism; glycolysis / gluconeogenesis.
K04043	*dnaK*	Genetic information processing-folding, sorting and degradation.
K01609	*trpC*	Biosynthesis of antibiotics, secondary metabolites and aminoacids; metabolic pathways; phenylalanine, tyrosine and tryptophan biosynthesis.
K00033	*gndA*	Biosynthesis of antibiotics and secondary metabolites; Metabolic pathways; microbial metabolism in diverse environments; carbon metabolism; phenylalanine, tyrosine and tryptophan biosynthesis; glutathione metabolism; pentose phosphate pathway.

3.3. Proteome analysis

Qualitative analysis of the proteomic data demonstrated that the samples clustered according to treatment, i.e., high or low cellobiose condition (Figure S1). In total, 1,418 proteins could be identified in both conditions in two of three replicates each. Overall, 448 proteins showed significant differences between the low and high cellobiose conditions (t-test, p=0.01; Figure S2). In addition, 249 so-called on/off proteins that were only present in one condition were detected. The proteome patterns of WH15 under the low and high cellobiose treatments are depicted in Figure 5. In the high cellobiose condition, 121 proteins were upregulated two-fold, and 129 proteins were “on”; 78 proteins were downregulated at least twofold, and 120 proteins were “off” (Figure 5; Tables S4 and S5). A comparison with the transcriptome data revealed that 2 ORFs (*lytR* and a hypothetical protein – ORF 05985) were upregulated, and 2 ORFs (*gndA* and a hypothetical protein – ORF 08600) were downregulated in both datasets for the high cellobiose treatment. Among the differentially expressed proteins, 332 could be assigned to COG categories, and 184 were annotated to KEGG orthologs (Figure 6).

3.3.1. Upregulated proteins

COG analysis demonstrated that most of the upregulated proteins belong to the categories cell wall/membrane/envelope biogenesis (44) and defense mechanism (16) (Figure 6a). TigrFam classification grouped most of the upregulated/on proteins in the categories protein fate (27), followed by cellular processes (21), transport/binding proteins (20) and cell wall/membrane/envelope biogenesis (12) (Figure S3). Closer examination of the subroles of the most abundant protein categories demonstrated that their main functions were related to peptide secretion and trafficking, detoxification and toxin production/resistance processes, efflux pumps, transporters and TonB-dependent receptors (Table S4). In addition to the sensory transduction protein LytR (*lytR*), which was also upregulated in the transcriptome data, several cell membrane proteins could be identified, such as the outer membrane protein OprM (*oprM* 1,4. 5, 6 and 7), the LPS-assembly protein LptD (*lptD* 2), the macrolide

export proteins MacA and MacB (*macA* 1, *macB* 6 and 8), the polysialic acid transport protein KpsD (*kpsD* 2) and the multidrug resistance proteins MdtB and MdtC (*mdtB* 2 and 3, *mdtC* 3 and 6). Among the upregulated proteins, 135 were annotated to KEGG orthologs, but only 44 orthologs could be mapped to KEGG metabolic pathways (Figure 6b, Table S6), and some identifiers were assigned to more than one pathway. Most of the proteins were mapped to 'general' metabolic pathways (15), two-component systems (9), biosynthesis of secondary metabolites (6), ABC transporters (4) and bacterial secretion systems (4), indicating that no particular metabolic pathway seemed to be specifically stimulated under the high cellobiose condition (Figure S4), although the expression of many membrane proteins was enhanced (Figure 7). However, when the genes assigned to the COG carbohydrate transport and metabolism category were examined, proteins related to trehalose metabolism were observed, such as OtsA (trehalose-6-phosphate synthase), TreS (trehalose synthase) and GlgE 1 (alpha-1,4-glucan:maltose-1-phosphate maltosyltransferase), as well as a hypothetical protein similar to cellobiose phosphorylase (ORF GWH15_11910). Another interesting protein identified was CcpA (catabolite control protein A) (Table S4), which is involved maily in carbon

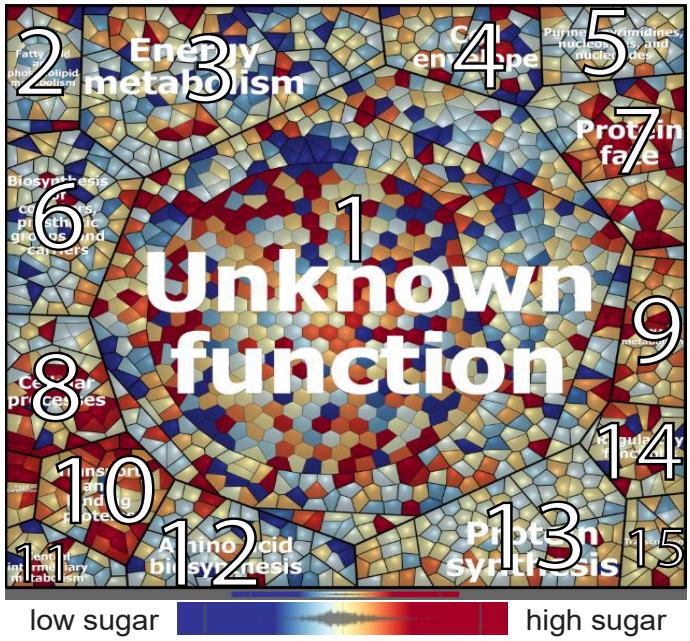

Figure 5: Voronoi treemap visualization of the protein expression patterns of *Granulicella* sp. WH15 under low and high cellobiose concentrations. Functional classification was conducted using Prophane 2.0 (www.prophane.de) and is based on TIGRFAMs, function level "main role". Each cell represents a quantified protein; proteins are clustered according to their function. Proteins with higher amount under low cellobiose conditions are depicted in blue, proteins with higher amount in high cellobiose conditions are depicted in red. 1- unknown function; 2-fatty acid and phospholipid metabolism; 3-energy metabolism; 4-cell envelope; 5-nucleotide metabolism; 6-co-factors and prosthetic groups; 7-protein fate; 8-cellular processes; 9-DNA metabolism; 10-transport and binding proteins;11-central intermediary metabolism; 12-aminoacid biosynthesis; 13-protein synthesis; 14-regulatory functions; 15-transcription.

metabolism regulation (Fujita, 2014).

3.3.2. Downregulated proteins

Within the COG categories, the downregulated proteins were mostly distributed among the carbohydrate transport and metabolism (35), amino acid transport and metabolism (20) and

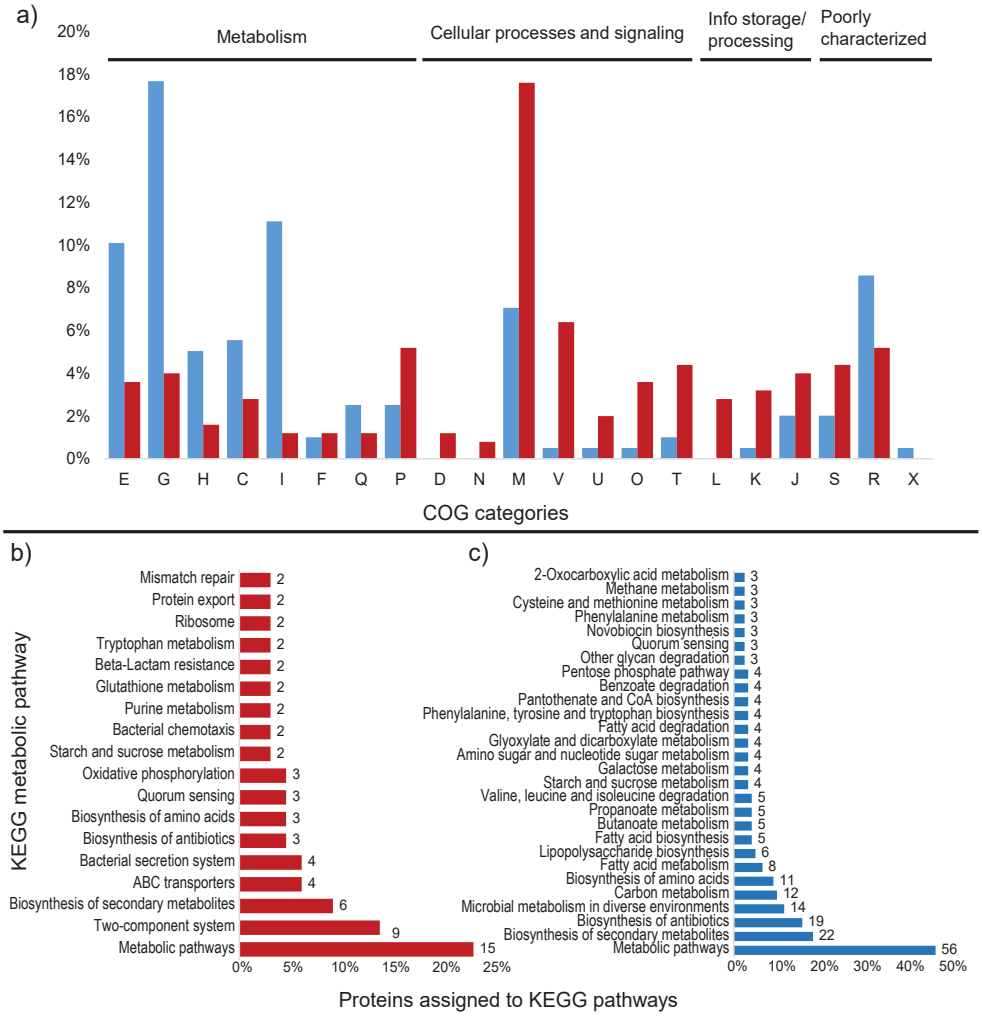

Figure 6: Differentially expressed proteins assigned to COG categories and KEGG pathways in the proteomic profile of *Granulicella* strain WH15 in high cellobiose concentration. a) Percentage of upregulated (red) and downregulated (blue) proteins assigned to COG categories; b) Percentage and number of upregulated proteins assigned to KEGG pathways; c) Percentage and number of downregulated proteins assigned to KEGG pathways. Unassigned proteins are not shown. E-Amino acid transport and metabolism; G-Carbohydrate transport and metabolism; H-Coenzyme transport and metabolism; C-Energy production and conversion; I-Lipid transport and metabolism; F-Nucleotide transport and metabolism; Q-Secondary metabolites; D-Cell cycle; N-Cell motility; M-Cell wall/membrane/ envelope biogenesis; V-Defense mechanisms; P-Inorganic ion transport and metabolism; U-Intracellular trafficking; O-Post-translational modification; T-Signal transduction mechanisms; L-Replication, recombination and repair; K-Transcription; J-Translation; S-Function unknown; R-General function and prediction; X-Mobilome.

lipid transport and metabolism (22) categories (Figure 6a). TigrFam classified most of the downregulated/off proteins in the categories of energy metabolism (29) and biosynthesis of cofactors, prosthetic groups and carriers (13) (Figure S3). Among the TigrFam subroles of the most abundant categories, most of the proteins were related to sugar metabolism and biosynthesis of vitamins. In total, 119 proteins were assigned to KEGG orthologs, and 68 orthologs were assigned to KEGG metabolic pathways. The majority of the proteins were assigned to 'general' metabolic pathways (56), biosynthesis of secondary metabolites (22), biosynthesis of antibiotics (19), microbial metabolism in diverse environments (12), carbon metabolism (12) and biosynthesis of amino acids (11) (Figure 6b). Consistent with the COG categories, downregulation of proteins linked to KEGG pathways involved in carbon, amino acid and lipid metabolism was observed (Figure S4, Table S5). In addition to general carbon metabolism, pathways related to secondary carbon sources seemed to be repressed, such as the metabolism of starch, sucrose, galactose, amino sugars, nucleotide sugars, fructose, mannose and other glycans, the pentose phosphate pathway, and glycolysis (Figure 6b, Figure 7, Table S6). Repression of pathways related to several amino acids, such as valine, leucine, phenylalanine, tyrosine and tryptophan, and the metabolism of fatty acids and lipopolysaccharides was observed (Figure 6b). The carbon-related repressed proteins

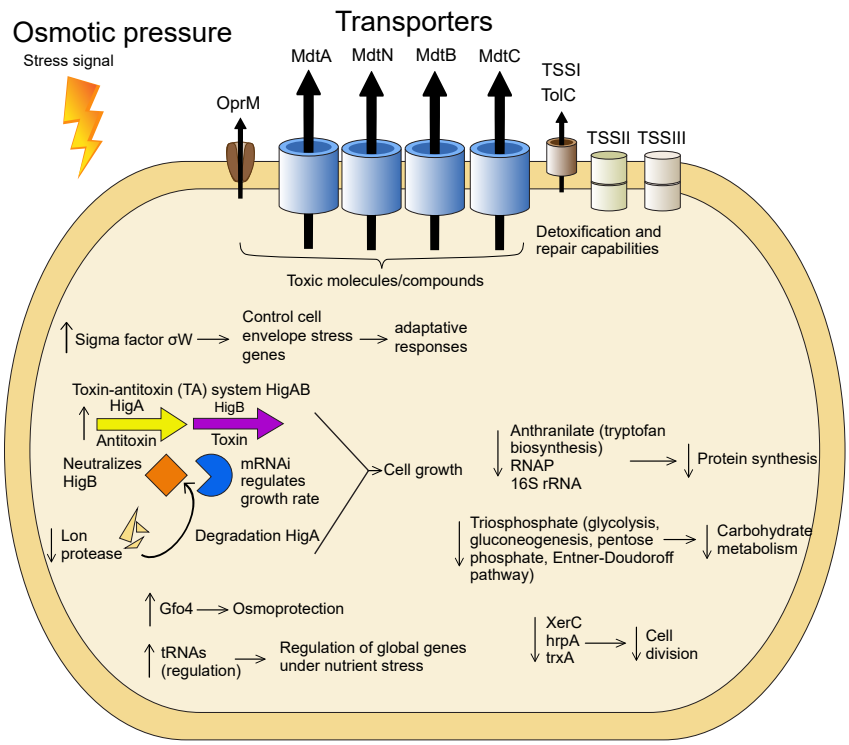

Figure 7: Model of the cellular processes involved in the adaptation to higher concentrations of cellobiose in *Granulicella* sp. WH15. The arrows depict upregulation (↑) and downregulation (↓) of transcripts and proteins.

included beta galactosidases (BgaA and BgaB), 1,4-beta-D-glucan glucohydrolase (GghA), xylan 1,4-beta-xylosidase (Xyl3A 1), beta-xylosidase (XynB), endopolygalacturonase (PehA 1), and exo-poly-alpha-D-galacturonosidase (PehX). Furthermore, 4 enzymes involved in the pentose phosphate pathway were downregulated: 6-phosphogluconate dehydrogenase, NADP (+)-dependent decarboxylase (GndA), transketolase 2 (TktB 1), gluconolactonase (Gnl 1) and KHG/KDPG aldolase (Eda). Among the proteins related to amino acid metabolism, the glutamate-pyruvate aminotransferase AlaA (alanine), histidinol-phosphate aminotransferase HisC2 (histidine), and prephenate dehydrogenase TyrA2 (tyrosine) were repressed. Enzymes involved in fatty acid degradation such as 3-ketoacyl-CoA thiolase (FadA), acyl-CoA dehydrogenase (FadE) and 3-hydroxyacyl-CoA dehydrogenase (FadN) were also downregulated (Table S5).

4. Discussion

Granulicella sp. WH15, an isolate collected from soil containing decaying wood, has a genome of 4.7 Mbp. Genome analysis revealed that a large number of genes were assigned to subsystem categories of carbohydrates (14.8%), amino acids and derivatives (12.2%) and protein metabolism (11.3%). Among other strains of *Granulicella*, 6.9% of the *G. tundricola* MP5ACTX9^{T} genome (Rawat *et al.*, 2013) and 9.1% of the *G. mallensis* MP5ACTX8^{T} genome (Rawat *et al.*, 2013) are dedicated to the transport and metabolism of carbohydrates. Similar to other *Granulicella* strains, strain WH15 possesses a wide range of glycoside hydrolases suggesting that this bacterium is well equipped for the carbon cycling process in soil as well as the hydrolysis and utilization of stored carbohydrates and the biosynthesis of exopolymers (Rawat *et al.*, 2013, Rawat *et al.*, 2013). The saccharide gene clusters revealed by ANTISMASH analysis are likely involved in the production of capsule, antigen and exopolymers. Exopolymer production by this strain has been observed under laboratory conditions, and the composition of the biopolymer has been characterized (Kielak *et al.*, 2017). However, the expression of gene clusters and proteins related to the production of EPS was not observed in this study, as the timepoint of cell collection was too early for substantial EPS production, which primarily occurs in stationary phase for both strains (Kielak *et al.*, 2017).

Members of the phylum *Acidobacteria* are generally considered slow-growing organisms that succeed in oligotrophic environments. The acidobacterium *Granulicella* sp. was isolated by employing a low nutrient culture medium (Valášková *et al.*, 2009), similar to other strains of *Granulicella* (Pankratov & Dedysh, 2010, Mannisto *et al.*, 2012). Although this strain can be cultivated using low quantities of nutrients, it also develops well in higher concentrations of sugar (Campanharo *et al.*, 2016). To better understand the behavior of *Granulicella* sp. WH15 in response to different carbon source concentrations, comparative transcriptomic and proteomic analyses were performed under 0.025% and 3% cellobiose. We used the low

cellobiose concentration as a reference, therefore evaluating up and downregulated proteins in the high cellobiose treatment. The comparative transcriptomic and proteomic profiles of strain WH15 under low and high cellobiose conditions demonstrated that the higher concentration of cellobiose triggered the expression of transcripts and proteins related to stress responses. These results suggest that the higher cellobiose concentration might have generated an osmotic stress condition for our strains. Stress conditions in bacteria induce changes at the transcriptional level that are often associated with a variety of σ factors, which bind to RNA polymerase (Helmann, 2016). Transcriptomic profiling showed that, in addition to other stress proteins, the sigma factor σW was upregulated $\log_2$ >1.0-fold. In *Bacillus subtilis*, the σW regulon controls genes related to cell envelope stress, membrane proteins and proteins involved in protection against toxins or antibiotics. Expression of this regulon has been observed under conditions of stress that affect cell wall biosynthesis or membrane integrity, such as alkali shock, salt stress, and treatments with cationic peptides and detergents (Zweers *et al.*, 2012). The cell envelope is the most external form of bacterial defense and receives stress stimuli, senses perturbations, and transmits signals that induce the transcriptional changes necessary for an adaptive response (Sacheti *et al.*, 2014). Another stress system upregulated in the high cellobiose condition was the toxin-antitoxin (TA) system HigAB. The prokaryotic TA system encodes a stable toxin (HigB) and an unstable antitoxin (HigA) that neutralizes HigB. HigB is an mRNA interferase that is believed to be involved in growth rate control (Christensen-Dalsgaard *et al.*, 2010). Expression of the system is induced by various stress stimuli, such as nutritional stress, heat shock (Tachdjian & Kelly, 2006), and exposure to chloramphenicol and chloroform (Gvakharia *et al.*, 2007, Jorgensen *et al.*, 2008). In *Escherichia coli*, amino acid starvation strongly induces the transcription of the TA system (Christensen-Dalsgaard *et al.*, 2010). The higher expression level of the antitoxin HigA compared to HigB in this study suggests that the acidobacterial cells were adapting to stress (Hayes & Low, 2009) and therefore produced more antitoxin to counterbalance the effect of HigB. In addition, these results are compatible with the downregulation of Lon protease, which contributes to the proteolytic regulation of many cell functions and is involved in the degradation of HigA (Christensen *et al.*, 2001).

The upregulated gene *gfo4* encodes glucose-fructose oxidoreductase, an enzyme involved in the production of molecules that function in intracellular osmoprotection under high cellobiose conditions (Nidetzky *et al.*, 1997). The production of this enzyme could be important to counterbalance the effects of the higher amounts of cellobiose in the culture medium. By contrast, the gene *glpF*, which encodes a glycerol uptake facilitator protein, was downregulated. This aquaglyceroporin facilitates the transport of glycerol and other linear polyalcohols across membranes, which would be redundant, as the cell is able to produce such molecules via the glucose-fructose oxidoreductase gfo4.

Another upregulated protein, LytR, is a sensory transduction protein that is part of the LytSR system, which in *Staphylococcus aureus* functions as a sensor-response system that detects

perturbations in the electrical potential of the cell membrane, such as disturbances caused by the presence of stress agents (Patel & Golemi-Kotra, 2016). Furthermore, the upregulation of tRNA genes could be involved in expression regulation. In addition to their central role in protein synthesis, tRNAs are involved in the regulation of global gene expression during nutritional stresses to adapt microbial metabolism to changes in amino acid concentrations (Li & Zhou, 2009). In the yeast *Saccharomyces cerevisiae*, for instance, glutamine tRNA is responsible for sensing of nitrogen sources in the environment and acts as a development regulator (Murray *et al.*, 1998). In addition, under different stress conditions, the relative abundances of tRNAs in *S. cerevisiae* change towards tRNAs recognizing rare codons to induce faster translation of the stress proteins necessary for adaptation (Torrent *et al.*, 2018). Although these functions of tRNAs were described in yeast, similar functions may be present in bacteria such as strain WH15.

The proteomic profile of the high cellobiose treatment was consistent with the upregulation of stress-related genes observed in the transcriptomic analyses. Most of the upregulated/on proteins were membrane proteins related to peptide secretion and trafficking, detoxification and toxin resistance, efflux pumps and transporters, and cell wall and defense mechanisms. The high abundance of membrane proteins may indicate the increased production of toxic metabolites, which can be removed from cells via pump systems (Rosner & Martin, 2009). The upregulated membrane transporters included type I (TolC family), II and III secretion system proteins, the multidrug resistance proteins MdtA, MdtN, MdtB and MdtC, and the outer membrane protein OprM. The protein TolC is important for membrane structure and function in *E. coli*, as well as in the export of toxic molecules in gram-negative bacteria, and TolC mutants are impaired in detoxification and repair (Zgurskaya *et al.*, 2011). The proteins MdtA, MdtN, MdtB and MdtC belong to the resistance-nodulation-division (RND) family of transporters, which play a role in drug resistance in gram-negative bacteria and promote the efflux of a wide range of toxic compounds, including antibiotics and detergents (Nikaido, 1996, Kim *et al.*, 2009). The outer membrane protein OprM is part of a membrane protein complex capable of actively ejecting an assortment of harmful compounds and is most notably involved in multidrug-resistance in gram-negative bacteria such as *Pseudomonas aeruginosa* (Federici *et al.*, 2005). Interestingly, genes encoding cation/multidrug efflux pumps are upregulated in cells of *Acetobacter aceti* grown in glucose, although the reasons are unknown (Sakurai *et al.*, 2010).

KEGG metabolic pathway analysis of the upregulated proteins and transcripts demonstrated that no specific metabolic pathway was enhanced by the high cellobiose treatment. Nonetheless, a predominance of categories related to membrane protein complexes was observed, such as two-component systems, ABC transporters and bacterial secretion systems. Closer examination of the proteins involved in carbon metabolism revealed the upregulation of proteins related to trehalose biosynthesis, such as trehalose-6-phosphate synthase (OtsA), trehalose synthase (TreS) and alpha-1,4-glucan:maltose-1-phosphate maltosyltransferase

(GlgE 1). Trehalose is an osmoprotectant disaccharide that accumulates in microorganisms under conditions of osmotic stress, and its role in cellular protection has been demonstrated in bacteria and yeast (Ruhal *et al.*, 2013). The expression of these enzymes indicates that the bacterial cells were reacting to the high cellobiose conditions by storing trehalose to increase their resistance to osmotic stress conditions.

The upregulation of protein GWH15_11910, which is similar to a cellobiose phosphorylase, may confirm the use of cellobiose as a carbon source, since this enzyme catalyzes the phosphorolysis of the β-1,4-glucosidic bond of the disaccharide to produce α-D-glucose-1-phosphate and D-glucose (Bianchetti *et al.*, 2011); however, no other carbon metabolism pathways were enhanced. Furthermore, the catabolic control protein CcpA was upregulated. The impact of this protein on cellular metabolism will be discussed later, as it promotes the repression of several metabolic pathways. The analysis of the downregulated transcripts indicated a decline in cell activity marked by reduced protein synthesis, carbohydrate metabolism and cell division. The downregulated proteins included the enzyme anthranilate isomerase, which is involved in tryptophan biosynthesis (Thoma *et al.*, 2000); RNase P, which is responsible for the production of the mature 5` ends of precursor tRNAs and is essential for translation (Schön, 1999); and the small subunit of the ribosome, 16S rRNA, which is fundamental for protein synthesis in prokaryotes (Singh *et al.*, 2018). As part of central carbon metabolism, the enzyme triosephosphate isomerase is vital for energy production, since it plays a key role in glycolysis, gluconeogenesis, and the pentose phosphate and Entner-Doudoroff pathways (Paterson *et al.*, 2009). The tyrosine recombinase XerC has a significant role in chromosome segregation during cell division and is important for the maintenance of replicon stability in *E. coli* (Grainge & Sherratt, 1999), and the chaperone HspA is a heat shock protein that is involved in the correct folding of proteins (Scieglinska & Krawczyk, 2014). The thioredoxin TrxA is a small redox protein that also possesses a chaperone function (Negrea *et al.*, 2009). The analysis of the downregulated proteins further revealed that the majority of the proteins were related to energy metabolism and cofactor biosynthesis, consistent with the downregulation of transcripts related to carbohydrate metabolism and cell division. Similarly, when exposed to osmotic stress, the metabolism of *E. coli* growing in minimal medium slows, and respiration and macromolecule synthesis are inhibited (Houssin *et al.*, 1991). Osmotic pressure produces deformation of the cell membrane, which could explain the expression of membrane stress signals observed in the present study. In addition, osmotic stress in *Sinorhizobium meliloti* promotes the repression of several functions of central metabolism and energy production systems (Dominguez-Ferreras *et al.*, 2006).

An important protein impacting cell metabolism is the catabolic control protein CcpA, which was upregulated under high cellobiose conditions. This protein is involved in carbon catabolite repression (CCR), a regulatory process that, in the presence of a preferable carbon source, inhibits alternative metabolic pathways (Jankovic & Bruckner, 2002). Although mainly found in Gram-positive bacteria, it has been suggested that Gram-negative bacteria might have

CcpA-dependent CCR (Warner & Lolkema, 2003). CCR regulates not only genes and operons involved in carbon metabolism but also those involved in the metabolism of amino acids and nucleotides and in the synthesis of extracellular enzymes and secondary metabolites (Fujita, 2014). Consistent with this process, we observed the downregulation of several enzymes related to secondary carbon source hydrolysis, such as beta galactosidases, xylosidases and galacturonases, as well as the pentose phosphate (PP) pathway. The PP pathway is fundamental for the generation of NADPH molecules and biosynthetic intermediates that are essential for the production of fatty acids, glutamate, purines, histidine and aromatic amino acids (Richardson *et al.*, 2015). Therefore, downregulation of the PP pathway could also contribute to the general repression of cell metabolism. Furthermore, enzymes related to amino acids, fatty acid metabolism and their corresponding KEGG pathways were repressed. Although many transcripts and proteins could not be identified due to limited information in the databases, the analysis of the present dataset suggested that the addition of a high cellobiose concentration in the culture medium of *Granulicella* sp. WH15 initially triggered a stress response. As part of this response, the expression of excretory membrane proteins was enhanced to promote the secretion of putative toxic byproducts from bacterial metabolism; energy metabolism was repressed to reallocate resources towards maintenance instead of growth; and the production of the osmoprotectant trehalose was enhanced. Data from experiments of bacteria under disturbances are important for understanding the behavior of microorganisms in response to environmental stresses, particularly for *Acidobacteria*, which is highly abundant in soil but whose soil ecology remains to be established. Further studies will improve the understanding of the mechanisms underlying the adjustment of the slow growth of *Granulicella* and other *Acidobacteria* strains to higher growth rates.

Acknowledgements

We thank Victor de Jager and Mattias de Hollander for bioinformatics assistance, Genomics Resource Center (USA) for the bacteria genome sequencing. This research was supported by (NWO-729.004.013). Ohana Y.A. Costa was supported by an SWB grant from CNPq [202496/2015-5] (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

Supplementary material

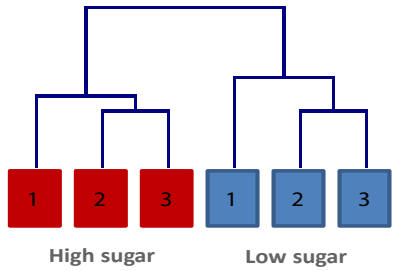

Figure S1: Cluster analysis of the proteome profile based on qualitative data in low and high cellobiose conditions. Hierarchical clustering of z-transformed normalized data was performed assuming unequal group variances (Welch approximation) and p-values based on all permutations (p=0.01) were adjusted using Bonferroni correction.

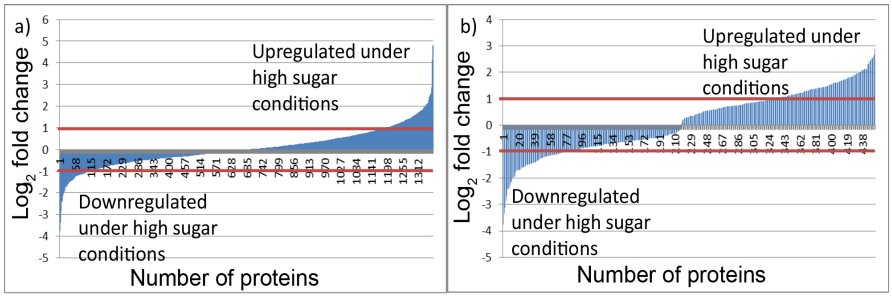

Figure S2: Expression pattern of proteins under high and low cellobiose cultivation of *Granulicella* sp. WH15 a) All proteins identified in at least two out of three replicates (excluding on/off proteins). b) Only proteins with significant change t-test p=0.01. log_2fold change is indicated by a red line.

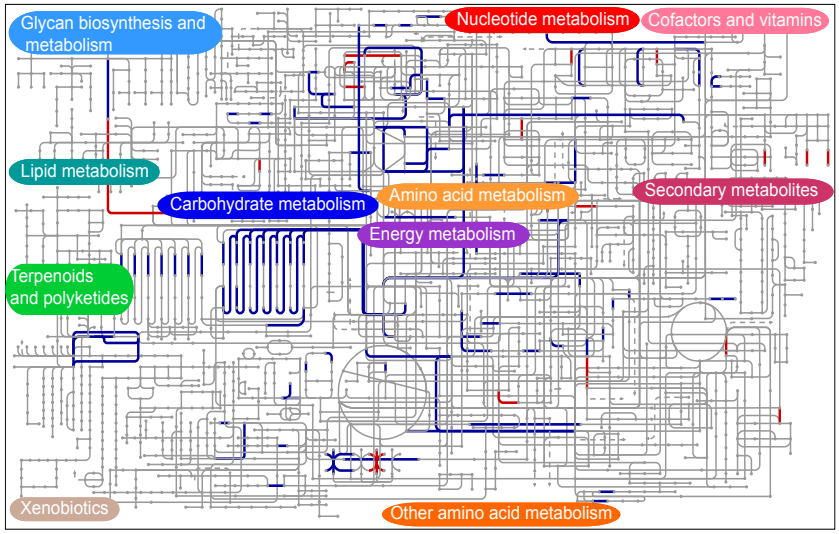

Figure S3: General overview of up (red) and downregulated (blue) metabolic pathways based on KEGG analysis of differentially expressed proteins in the proteomic profile of *Granulicella* sp WH15 in high cellobiose concentration. No particular metabolic pathway seemed to be specifically upregulated and downregulated proteins were mostly distributed among the carbohydrate transport and metabolism pathways.

Table S1: Growth of strain *Granulicella* sp. WH15 in culture media supplemented with different carbon sources.

Carbon source	Growth
Pectin	-
Glycogen	-
Glucosamine	-
Cellulose	-
D-glucose	+
D-galactose	+
D-mannose	+
D-xylose	+
L-arabinose	+
L-rhamnose	+
D-galacturonic acid	-
Cellobiose	+
D-lactose	+
Sucrose	+

+=positive growth; -=No growth

Table S2. Total number of transcripts reads per sample in the transcriptomic profile of *Granulicella* sp. WH15 under low and high cellobiose concentrations.

Sample ID	Total number of reads
Low cellobiose (1)	15,731,147
Low cellobiose (2)	12,624,878
Low cellobiose (3)	11,080,985
High cellobiose (1)	11,138,128
High cellobiose (2)	9,322,795
High cellobiose (3)	10,071,593

Table S3: Differentially expressed transcripts in the transcriptomic profile of *Granulicella* sp. WH15 in high cellobiose concentration. Only statistically significant differentially expressed genes with a fold change ≥1.0 or ≤−1.0 are shown.

ORF	Annotation	Log_2FC
GWH15_14040	hypothetical protein	3.71
GWH15_06005	hypothetical protein	3.12
GWH15_00285	tRNA-Asn(gtt)	2.74
GWH15_06010	hypothetical protein	2.70
GWH15_14055	hypothetical protein	2.66
GWH15_14060	hypothetical protein	2.18
GWH15_14100	hypothetical protein	2.00
GWH15_08445	hypothetical protein	1.94
GWH15_05955	hypothetical protein	1.93
GWH15_14070	gfo4	1.92
GWH15_16895	tRNA-Asp(gtc)	1.72
GWH15_17685	higA	1.71
GWH15_18230	hypothetical protein	1.67
GWH15_01965	lytR	1.67

ORF (continued)	Annotation	Log_2FC
GWH15_18255	hypothetical protein	1.60
GWH15_06405	hypothetical protein	1.59
GWH15_02080	hypothetical protein	1.54
GWH15_10070	hypothetical protein	1.47
GWH15_07155	hypothetical protein	1.41
GWH15_03690	hypothetical protein	1.37
GWH15_14050	sigW	1.35
GWH15_05960	hypothetical protein	1.32
GWH15_05985	hypothetical protein	1.26
GWH15_17690	higB-1	1.17
GWH15_11830	hypothetical protein	1.13
GWH15_14160	hypothetical protein	1.10
GWH15_06410	hypothetical protein	1.06
GWH15_04600	putative peroxiredoxin	1.04
GWH15_17625	hypothetical protein	-1.03
GWH15_00710	hypothetical protein	-1.06
GWH15_00505	rpoD	-1.09
GWH15_16220	hypothetical protein	-1.19
GWH15_14410	Lon protease 2	-1.20
GWH15_08600	hypothetical protein	-1.21
GWH15_14215	hypothetical protein	-1.22
GWH15_10510	hypothetical protein	-1.29
GWH15_00210	RNaseP_bact_a	-1.30
GWH15_01535	gndA	-1.33
GWH15_00720	hypothetical protein	-1.33
GWH15_01770	23S ribosomal RNA	-1.34
GWH15_19395	hypothetical protein	-1.34
GWH15_00785	trpC	-1.39
GWH15_04765	xerC	-1.44
GWH15_01625	dnaK	-1.45
GWH15_11560	hypothetical protein	-1.47
GWH15_07810	ssrA	-1.53
GWH15_15305	glpF	-1.63
GWH15_08540	hypothetical protein	-1.75
GWH15_12250	trxA	-1.80
GWH15_01755	16S ribosomal RNA	-1.94
GWH15_12725	hypothetical protein	-2.06
GWH15_08685	xerC	-2.09
GWH15_16995	tpiA	-2.10
GWH15_00715	hypothetical protein	-2.26
GWH15_06890	xerC	-2.50
GWH15_00885	hspA	-2.60
GWH15_00780	trpF	-2.80
GWH15_19400	hypothetical protein	-3.05

Table S4: Significantly upregulated proteins in the proteomic profile of *Granulicella* sp. WH15 in high cellobiose concentration. Only statistically significant differentially expressed proteins with a fold change ≥1.0 are shown.

ORF	Description	log_2 FC
GWH15_02420	Hypothetical protein	4
GWH15_16935	7-carboxy-7-deazaguanine synthase	4
GWH15_15250	ABC transporter ATP-binding protein ytrB	4

ORF (continued)	Description	log_2 FC
GWH15_08630	AI-2 transport protein tqsA	4
GWH15_07165	Aminopeptidase ypdF	4
GWH15_00460	Amylopullulanase	4
GWH15_08480	Anti-sigma-K factor rskA	4
GWH15_08415	Apolipoprotein N-acyltransferase	4
GWH15_00205	Band 7 protein	4
GWH15_05535	Beta-barrel assembly-enhancing protease	4
GWH15_17125	Carbohydrate acetyl esterase/feruloyl esterase	4
GWH15_11910	Carbohydrate transport and metabolism-Glycosyltransferase 36	4
GWH15_15690	Carboxy-terminal processing protease ctpA	4
GWH15_09810	Catabolite control protein A	4
GWH15_18425	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	4
GWH15_19395	Cell wall/membrane/envelope biogenesis asmA family	4
GWH15_01995	Cell wall/membrane/envelope biogenesis-asmA family	4
GWH15_09540	Cna B-type protein-transport	4
GWH15_18835	Cobalt-zinc-cadmium resistance protein czcA	4
GWH15_18685	Conserved protein-pfam:duf403	4
GWH15_18690	Conserved protein-pfam:duf403	4
GWH15_09565	Cyclic pyranopterin monophosphate synthase 1	4
GWH15_10075	Cytochrome c oxidase subunit 2	4
GWH15_09055	Cytochrome c-type biogenesis protein ccmF	4
GWH15_10475	Diaminopimelate decarboxylase	4
GWH15_00815	Diguanylate cyclase dosC	4
GWH15_00520	DNA primase	4
GWH15_13725	DNA-directed RNA polymerase subunit omega	4
GWH15_03780	D-ribose-binding periplasmic protein	4
GWH15_02260	Endonuclease mutS2	4
GWH15_00620	Hypothetical protein	4
GWH15_01680	Exopolyphosphatase	4
GWH15_18735	FAD-dependent decaprenylphosphoryl-beta-D-ribofuranose 2-oxidase	4
GWH15_08515	Ferrous iron permease efeU	4
GWH15_00500	Glycogen synthase	4
GWH15_06205	Guanine deaminase	4
GWH15_12695	H(+)/Cl(-) exchange transporter clcA	4
GWH15_00225	Hopanoid biosynthesis associated glycosyl transferase protein hpnI	4
GWH15_19130	HTH-type transcriptional repressor fabR	4
GWH15_12380	Hypothetical protein	4
GWH15_00410	Hypothetical protein	4
GWH15_00175	Hypothetical protein	4
GWH15_01300	Hypothetical protein	4
GWH15_03025	Hypothetical protein	4
GWH15_04195	Hypothetical protein	4
GWH15_04880	Hypothetical protein	4
GWH15_05505	Hypothetical protein	4
GWH15_05985	Hypothetical protein	4
GWH15_06025	Hypothetical protein	4
GWH15_06760	Hypothetical protein	4
GWH15_07215	Hypothetical protein	4
GWH15_08350	Hypothetical protein	4
GWH15_13120	Hypothetical protein	4
GWH15_15045	Hypothetical protein	4
GWH15_15255	Hypothetical protein	4
GWH15_16300	Hypothetical protein	4
GWH15_18130	Hypothetical protein	4
GWH15_02890	Hypothetical protein	4
GWH15_07610	Inorganic ion transport and metabolism-tonB-dependent Receptor	4
GWH15_08030	Inorganic ion transport and metabolism-tonB-dependent Receptor	4
GWH15_15415	Inorganic ion transport and metabolism-tonB-dependent Receptor	4
GWH15_00615	TonB-dependent heme/hemoglobin receptor family protein	4
GWH15_15365	Isoaspartyl dipeptidase	4
GWH15_03790	L-asparaginase 2	4

ORF (continued)	Description	log_2 FC
GWH15_11775	Lexa repressor	4
GWH15_12485	Lipid transport and metabolism-desaturase	4
GWH15_15780	LPS-assembly lipoprotein lptE	4
GWH15_13770	LPS-assembly protein lptD	4
GWH15_15340	Macrolide export ATP-binding/permease protein macB	4
GWH15_15475	Macrolide export ATP-binding/permease protein macB	4
GWH15_03815	Macrolide export protein macA	4
GWH15_12245	Magnesium and cobalt efflux protein corC	4
GWH15_06965	Mannosylfructose-phosphate synthase	4
GWH15_16655	Metallo-beta-lactamase family	4
GWH15_11300	Multidrug resistance protein mdtB	4
GWH15_12080	Multidrug resistance protein mdtB	4
GWH15_19560	Multidrug resistance protein mdtC	4
GWH15_12085	Multidrug resistance protein mdtC	4
GWH15_08245	NADH-quinone oxidoreductase subunit M	4
GWH15_15355	N-formyl-4-amino-5-aminomethyl-2-methylpyrimidine deformylase	4
GWH15_19200	N-formyl-4-amino-5-aminomethyl-2-methylpyrimidine deformylase	4
GWH15_08145	Nuclear protein SET	4
GWH15_15850	Nucleoside permease nupX	4
GWH15_13750	Nucleotide-binding protein	4
GWH15_14450	Outer membrane protein assembly factor bamA	4
GWH15_03820	Outer membrane protein oprM	4
GWH15_11310	Outer membrane protein oprM	4
GWH15_12090	Outer membrane protein oprM	4
GWH15_16030	Outer membrane protein oprM	4
GWH15_18825	Outer membrane protein oprM	4
GWH15_17110	Penicillin-binding protein 2	4
GWH15_16260	Pfam:DUF811	4
GWH15_14525	PGL/p-HBAD biosynthesis glycosyltransferase	4
GWH15_05440	Phytochrome-like protein cph1	4
GWH15_05490	Poly-beta-1,6-N-acetyl-D-glucosamine synthase	4
GWH15_06970	Polysaccharide export protein	4
GWH15_14210	Polysialic acid transport protein kpsD	4
GWH15_09965	Protein yceI	4
GWH15_15350	Putative ABC transporter ATP-binding protein yknY	4
GWH15_01450	Putative ctpA-like serine protease	4
GWH15_13740	Putative multidrug export ATP-binding/permease protein	4
GWH15_16305	Putative NTE family protein	4
GWH15_17565	Putative thiazole biosynthetic enzyme	4
GWH15_15010	Putative tonB-dependent receptor bfrD	4
GWH15_14825	Putative zinc metalloprotease	4
GWH15_13800	Putative zinc metalloprotease Rip3	4
GWH15_13795	Ribosomal RNA small subunit methyltransferase D	4
GWH15_14980	Ribosomal RNA small subunit methyltransferase H	4
GWH15_16290	RlpA-like protein	4
GWH15_09495	Sensor protein zraS	4
GWH15_01965	Sensory transduction protein lytR	4
GWH15_19165	Signal transduction mechanisms Serine Threonine protein kinase	4
GWH15_09570	Squalene-hopene cyclase	4
GWH15_12830	Thiol-disulfide oxidoreductase resA	4
GWH15_05860	TonB-dependent Receptor	4
GWH15_05950	TonB-dependent Receptor	4
GWH15_12055	TonB-dependent Receptor	4
GWH15_14280	TonB-dependent Receptor	4
GWH15_03535	TonB-dependent receptor plug	4
GWH15_19020	TonB-dependent receptor plug	4
GWH15_09325	Transcriptional activator cadC	4
GWH15_16620	Transcriptional regulatory protein ypdB	4
GWH15_14440	Translocation and assembly module tamB	4
GWH15_13920	Trehalase	4
GWH15_01390	Two component, sigma54 specific, transcriptional regulator, Fis family	4

ORF (continued)	Description	log_2 FC
GWH15_01570	Type II secretion system protein F	4
GWH15_06185	UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferase	4
GWH15_11890	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase	4
GWH15_19395	Hypothetical protein	4
GWH15_03080	Putative mycofactocin radical SAM maturase mftC	2.88
GWH15_17755	TonB-dependent receptor plug	2.68
GWH15_10545	Type II secretion system protein D	2.60
GWH15_17590	Hypothetical protein	2.54
GWH15_10175	Catalase-related peroxidase	2.52
GWH15_07490	Multidrug resistance protein mdtN	2.48
GWH15_07480	Outer membrane efflux protein bepC	2.44
GWH15_05765	TonB-dependent receptor plug	2.42
GWH15_11620	Hypothetical protein	2.32
GWH15_14285	TonB-dependent Receptor	2.28
GWH15_11915	Lipid A export ATP-binding/permease protein msbA	2.13
GWH15_15410	TonB-dependent Receptor	2.12
GWH15_13055	Putative phospholipid ABC transporter-binding protein mlaD	2.12
GWH15_16040	Multidrug resistance protein mdtC	2.11
GWH15_00900	N-formyl-4-amino-5-aminomethyl-2-methylpyrimidine deformylase	2.10
GWH15_01960	Hypothetical protein	2.08
GWH15_13535	Hypothetical protein	2.07
GWH15_01785	TM2 domain	2.02
GWH15_13930	Trehalose synthase/amylase treS	2.02
GWH15_16185	Hypothetical protein	1.99
GWH15_04655	ABC transporter, permease	1.97
GWH15_17630	BON domain	1.96
GWH15_12075	Multidrug resistance protein mdtA	1.95
GWH15_18985	Trehalose-phosphate synthase	1.95
GWH15_05290	Hypothetical protein	1.90
GWH15_15725	ATP-dependent zinc metalloprotease ftsH	1.89
GWH15_13030	Mechanosensitive ion channel	1.89
GWH15_08035	Putative membrane protein mmpL3	1.85
GWH15_00705	Outer membrane protein assembly factor bamA	1.83
GWH15_19570	Outer membrane protein tolC	1.83
GWH15_01695	Aminopeptidase N	1.83
GWH15_19225	Sporulation kinase E	1.82
GWH15_13285	Thymidylate kinase	1.81
GWH15_11185	UvrABC system protein A	1.81
GWH15_01225	Toluene efflux pump outer membrane protein ttgF	1.76
GWH15_15395	Peptidase M14, carboxypeptidase A	1.75
GWH15_02935	Multidrug resistance protein mdtB	1.73
GWH15_02030	NADH dehydrogenase-like protein yjlD	1.73
GWH15_02940	Multidrug resistance protein mdtA	1.72
GWH15_15995	DNA mismatch repair protein mutL	1.71
GWH15_03075	ABC transporter permease ytrF	1.69
GWH15_15795	Disulphide bond corrector protein dsbC	1.68
GWH15_19215	TonB-dependent receptor plug	1.67
GWH15_05980	TonB-dependent Receptor	1.66
GWH15_06975	Tyrosine-protein kinase ywqD	1.64
GWH15_04000	Tonb-dependent receptor plug	1.63
GWH15_15525	ABC transporter ATP-binding protein natA	1.62
GWH15_14150	Hypothetical protein	1.61
GWH15_01740	Succinyl-diaminopimelate desuccinylase	1.60
GWH15_06990	Hypothetical protein	1.60
GWH15_14235	TonB-dependent Receptor	1.60
GWH15_04050	Hypothetical protein	1.59
GWH15_02945	Decaprenyl-phosphate phosphoribosyltransferase	1.57
GWH15_07680	Catalase-peroxidase	1.54
GWH15_01780	Hypothetical protein	1.54
GWH15_14045	Putative oxidoreductase catD	1.53
GWH15_12475	Pca regulon regulatory protein	1.51

ORF (continued)	Description	log_2 FC
GWH15_02075	Hypothetical protein	1.48
GWH15_01335	Hypothetical protein	1.48
GWH15_12385	Ribosomal protein S12 methylthiotransferase rimO	1.47
GWH15_00530	Outer membrane protein domain-containing protein	1.47
GWH15_07050	Multidrug export protein emrA	1.45
GWH15_19055	Imidazolonepropionase	1.44
GWH15_12370	Putative efflux system component yknX	1.44
GWH15_09425	Efflux pump membrane transporter bepE	1.44
GWH15_18435	Anthranilate synthase component 1	1.43
GWH15_14490	DNA polymerase/3'-5' exonuclease polX	1.42
GWH15_05965	TonB-dependent Receptor	1.42
GWH15_12430	Putative phospholipid ABC transporter permease protein mlaE	1.41
GWH15_16035	Multidrug resistance protein mdtA	1.41
GWH15_17465	Outer membrane protein assembly factor bamD	1.40
GWH15_02400	TonB family	1.40
GWH15_02910	Hypothetical protein	1.39
GWH15_09000	Putative membrane protein	1.37
GWH15_15800	Thiol-disulfide oxidoreductase resA	1.37
GWH15_12290	Outer membrane lipoprotein Omp16	1.36
GWH15_09420	Outer membrane protein oprM	1.36
GWH15_17035	Glutathione synthase ribosomal protein s6 modification	1.36
GWH15_05770	Prolyl tripeptidyl peptidase	1.35
GWH15_13925	Alpha-1,4-glucan:maltose-1-phosphate maltosyltransferase 1	1.34
GWH15_02605	Methionine aminopeptidase 1	1.33
GWH15_08560	4,4'-diaponeurosporenoate glycosyltransferase	1.30
GWH15_11595	Putative mscs family protein ykuT	1.30
GWH15_07225	Hypothetical protein	1.29
GWH15_13275	Hypothetical protein	1.28
GWH15_07045	Outer membrane protein oprM	1.28
GWH15_11845	Vitamin B12 transporter btuB	1.27
GWH15_19245	Biopolymer transport protein exbD	1.26
GWH15_01440	Glycosyl transferase family	1.23
GWH15_12815	TonB-dependent receptor plug	1.23
GWH15_06900	BON domain	1.22
GWH15_03995	Feruloyl esterase	1.22
GWH15_12700	TonB family	1.22
GWH15_18875	TonB-dependent receptor plug	1.21
GWH15_13410	HTH-type transcriptional regulator lutR	1.19
GWH15_09020	Hypothetical protein	1.19
GWH15_04110	2-dehydro-3-deoxy-D-gluconate 5-dehydrogenase	1.18
GWH15_02670	Translation initiation factor IF-2	1.18
GWH15_18510	Biodegradative arginine decarboxylase	1.17
GWH15_05295	Dispase autolysis-inducing protein	1.15
GWH15_19675	Membrane protein insertase yidC	1.15
GWH15_07605	Response regulator protein vraR	1.15
GWH15_00730	30S ribosomal protein S1	1.14
GWH15_02465	30S ribosomal protein S12	1.13
GWH15_12305	Hypothetical protein	1.13
GWH15_19615	DNA-binding protein HU	1.12
GWH15_17135	Rod shape-determining protein mreB	1.11
GWH15_03785	L-asparaginase 2	1.11
GWH15_15505	Energy-dependent translational throttle protein ettA	1.09
GWH15_18520	Exodeoxyribonuclease 7 large subunit	1.08
GWH15_16085	Hypothetical protein	1.07
GWH15_14870	tRNA threonylcarbamoyladenosine biosynthesis protein tsaE	1.07
GWH15_03680	Vitamin B12-dependent ribonucleoside-diphosphate reductase	1.06
GWH15_05990	Hypothetical protein	1.06
GWH15_12285	Protein tolB	1.05
GWH15_11850	Hypothetical protein	1.05
GWH15_01330	Putative nicotinate-nucleotide pyrophosphorylase [carboxylating]	1.04
GWH15_11780	Hypothetical protein	1.03

ORF (continued)	Description	log_2 FC
GWH15_19025	Quinate/shikimate dehydrogenase (quinone)	1.03
GWH15_04580	Methionine synthase	1.02
GWH15_07455	Protein translocase subunit secA	1.01

Table S5: Significantly downregulated proteins in the proteomic profile of *Granulicella* sp. WH15 in high cellobiose concentration. Only statistically significant differentially expressed proteins with a fold change ≤-1.0 are shown.

ORF	Description	log2 FC
GWH15_02345	3-isopropylmalate dehydratase small subunit 1	-4.00
GWH15_14090	Formate dehydrogenase	-4.00
GWH15_00995	Glutamate synthase [NADPH] small chain	-4.00
GWH15_00610	Glycerate dehydrogenase	-4.00
GWH15_17605	Homoserine O-acetyltransferase	-4.00
GWH15_06475	1-deoxy-D-xylulose-5-phosphate synthase	-4.00
GWH15_06375	4-hydroxythreonine-4-phosphate dehydrogenase 2	-4.00
GWH15_01205	Dephospho-coA kinase	-4.00
GWH15_08325	L-2,4-diaminobutyrate decarboxylase	-4.00
GWH15_10235	Multifunctional cyclase-dehydratase-3-O-methyl transferase tcmN	-4.00
GWH15_11220	Phosphomethylpyrimidine synthase	-4.00
GWH15_18315	Phosphopantetheine adenylyltransferase	-4.00
GWH15_07695	Pyrroloquinoline-quinone synthase	-4.00
GWH15_02155	Riboflavin biosynthesis protein ribD	-4.00
GWH15_18275	3-deoxy-manno-octulosonate cytidylyltransferase	-4.00
GWH15_01640	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-Acyltransferase	-4.00
GWH15_15135	Lipopolysaccharide heptosyltransferase 1	-4.00
GWH15_15130	Tetraacyldisaccharide 4'-kinase	-4.00
GWH15_08130	UDP-3-O-acylglucosamine N-acyltransferase	-4.00
GWH15_12280	UDP-N-acetylenolpyruvoylglucosamine reductase	-4.00
GWH15_06130	Acetyl-coA:oxalate coA-transferase	-4.00
GWH15_07765	Phenylacetaldehyde dehydrogenase	-4.00
GWH15_11565	Biphenyl dioxygenase system ferredoxin subunit	-4.00
GWH15_11410	N-acetylglucosamine-6-phosphate deacetylase	-4.00
GWH15_07550	1,5-anhydro-D-fructose reductase	-4.00
GWH15_13480	2-dehydro-3-deoxy-D-gluconate 5-dehydrogenase	-4.00
GWH15_07735	2-hydroxy-3-oxopropionate reductase	-4.00
GWH15_17540	2-methylcitrate dehydratase	-4.00
GWH15_17545	2-methylcitrate synthase	-4.00
GWH15_05100	5'-nucleotidase	-4.00
GWH15_07745	Beta-galactosidase bgaB	-4.00
GWH15_13400	Beta-xylosidase	-4.00
GWH15_07310	Inosose dehydratase	-4.00
GWH15_10460	NADH-quinone oxidoreductase subunit I	-4.00
GWH15_09940	Pyrethroid hydrolase	-4.00
GWH15_06155	Ribulose bisphosphate carboxylase-like protein 2	-4.00
GWH15_08005	Sorbitol dehydrogenase	-4.00
GWH15_06480	Transketolase 2	-4.00
GWH15_18410	D-xylose 1-dehydrogenase	-4.00
GWH15_03275	Release factor glutamine methyltransferase	-4.00
GWH15_08570	6-carboxy-5,6,7,8-tetrahydropterin synthase	-4.00
GWH15_06145	Hypothetical protein	-4.00
GWH15_16100	Hypothetical protein	-4.00
GWH15_01830	Transcriptional regulatory protein yycF	-4.00
GWH15_13365	2-keto-3-deoxy-L-fuconate dehydrogenase	-4.00
GWH15_11615	3',5'-cyclic adenosine monophosphate phosphodiesterase cpdA	-4.00
GWH15_11945	4-hydroxy-4-methyl-2-oxoglutarate aldolase	-4.00
GWH15_17960	4-hydroxy-4-methyl-2-oxoglutarate aldolase	-4.00
GWH15_19220	6'''-hydroxyparomomycin C oxidase	-4.00

ORF (continued)	Description	log2 FC
GWH15_05265	Acetylxylan esterase	-4.00
GWH15_15430	Acyl-coA dehydrogenase	-4.00
GWH15_16250	Acyl-coA dehydrogenase	-4.00
GWH15_03975	Acyl-coenzyme A thioesterase paaI	-4.00
GWH15_17610	Beta-barrel assembly-enhancing protease	-4.00
GWH15_12730	Cellulose synthase operon protein C	-4.00
GWH15_06620	D-galactonate dehydratase	-4.00
GWH15_17965	D-galactonate dehydratase	-4.00
GWH15_11940	D-galactonate dehydratase family member	-4.00
GWH15_02130	Endonuclease 4	-4.00
GWH15_00665	Endo-polygalacturonase	-4.00
GWH15_17915	Exo-poly-alpha-D-galacturonosidase	-4.00
GWH15_09450	Glutamate-pyruvate aminotransferase alaA	-4.00
GWH15_02870	Hydroxymethylglutaryl-coa lyase yngG	-4.00
GWH15_08530	Hydroxypyruvate isomerase	-4.00
GWH15_00145	Hypothetical protein	-4.00
GWH15_00585	Hypothetical protein	-4.00
GWH15_02685	Hypothetical protein	-4.00
GWH15_03305	Hypothetical protein	-4.00
GWH15_03765	Hypothetical protein	-4.00
GWH15_03835	Hypothetical protein	-4.00
GWH15_03840	Hypothetical protein	-4.00
GWH15_05465	Hypothetical protein	-4.00
GWH15_06125	Hypothetical protein	-4.00
GWH15_06295	Hypothetical protein	-4.00
GWH15_06310	Hypothetical protein	-4.00
GWH15_06330	Hypothetical protein	-4.00
GWH15_06345	Hypothetical protein	-4.00
GWH15_06600	Hypothetical protein	-4.00
GWH15_07205	Hypothetical protein	-4.00
GWH15_07540	Hypothetical protein	-4.00
GWH15_07555	Hypothetical protein	-4.00
GWH15_07565	Hypothetical protein	-4.00
GWH15_07740	Hypothetical protein	-4.00
GWH15_07750	Hypothetical protein	-4.00
GWH15_07795	Hypothetical protein	-4.00
GWH15_07975	Hypothetical protein	-4.00
GWH15_08075	Hypothetical protein	-4.00
GWH15_08600	Hypothetical protein	-4.00
GWH15_08945	Hypothetical protein	-4.00
GWH15_09720	Hypothetical protein	-4.00
GWH15_10775	Hypothetical protein	-4.00
GWH15_12310	Hypothetical protein	-4.00
GWH15_12655	Hypothetical protein	-4.00
GWH15_12725	Hypothetical protein	-4.00
GWH15_13405	Hypothetical protein	-4.00
GWH15_13455	Hypothetical protein	-4.00
GWH15_13460	Hypothetical protein	-4.00
GWH15_15770	Hypothetical protein	-4.00
GWH15_16360	Hypothetical protein	-4.00
GWH15_16925	Hypothetical protein	-4.00
GWH15_17970	Hypothetical protein	-4.00
GWH15_18790	Hypothetical protein	-4.00
GWH15_19005	Hypothetical protein	-4.00
GWH15_11130	Hypothetical protein	-4.00
GWH15_09830	Hypothetical protein	-4.00
GWH15_06095	Hypothetical protein	-4.00
GWH15_17550	Methylisocitrate lyase	-4.00
GWH15_05095	Methylxanthine N1-demethylase ndmA	-4.00
GWH15_00990	NAD-dependent dihydropyrimidine dehydrogenase subunit preA	-4.00
GWH15_13220	Phosphoribosyl-AMP cyclohydrolase	-4.00

ORF (continued)	Description	log2 FC
GWH15_00845	Prephenate dehydrogenase	-4.00
GWH15_18405	Putative 3-hydroxybutyryl-coa dehydrogenase	-4.00
GWH15_02095	Putative formate dehydrogenase	-4.00
GWH15_12060	Putative N-succinyldiaminopimelate aminotransferase dapC	-4.00
GWH15_12955	Putative prophage major tail sheath protein	-4.00
GWH15_17805	Putative succinyl-coA:3-ketoacid coenzyme A transferase subunit B	-4.00
GWH15_03720	Reducing end xylose-releasing exo-oligoxylanase	-4.00
GWH15_09815	Retaining alpha-galactosidase	-4.00
GWH15_15040	Short-chain-enoyl-coa hydratase	-4.00
GWH15_13515	UDP-glucose 4-epimerase	-4.00
GWH15_07980	Hypothetical protein	-3.77
GWH15_07970	Hypothetical protein	-3.36
GWH15_07865	Putative acyl-coA dehydrogenase	-3.12
GWH15_07780	Gluconolactonase	-2.85
GWH15_07855	Putative 3-hydroxyacyl-coa dehydrogenase	-2.66
GWH15_13495	Virginiamycin B lyase	-2.42
GWH15_17530	Hypothetical protein	-2.40
GWH15_06300	Beta-galactosidase bgaA	-2.40
GWH15_07860	3-ketoacyl-coA thiolase	-2.28
GWH15_06320	Hypothetical protein	-2.16
GWH15_01715	Electron transfer flavoprotein subunit beta	-2.13
GWH15_18430	Glutamate dehydrogenase	-2.01
GWH15_07990	Extracellular exo-alpha-L-arabinofuranosidase	-1.97
GWH15_08000	Gluconate 5-dehydrogenase	-1.96
GWH15_06150	Non-reducing end beta-L-arabinofuranosidase	-1.91
GWH15_17820	Hypothetical protein	-1.77
GWH15_08445	Hypothetical protein	-1.73
GWH15_01710	Electron transfer flavoprotein subunit alpha	-1.71
GWH15_18330	Hypothetical protein	-1.71
GWH15_13235	Histidinol-phosphate aminotransferase 2	-1.70
GWH15_13420	Protein tolB	-1.70
GWH15_07965	Serine/threonine-protein kinase pknD	-1.69
GWH15_14310	Dihydroxy-acid dehydratase	-1.61
GWH15_07530	Hypothetical protein	-1.60
GWH15_03715	Hypothetical protein	-1.60
GWH15_11790	Metallo-beta-lactamase L1	-1.59
GWH15_09775	NAD-dependent malic enzyme	-1.58
GWH15_00385	Putative oxidoreductase ydbC	-1.54
GWH15_02245	1,4-beta-D-glucan glucohydrolase	-1.52
GWH15_11350	Putative KHG/KDPG aldolase	-1.51
GWH15_17080	Hypothetical protein	-1.49
GWH15_13390	L-rhamnonate dehydratase	-1.49
GWH15_01535	6-phosphogluconate dehydrogenase, NADP(+)-dependent	-1.49
GWH15_05340	General stress protein 69	-1.46
GWH15_13415	Xylitol oxidase	-1.46
GWH15_17570	D-galactarolactone cycloisomerase	-1.45
GWH15_07145	Long-chain-fatty-acid--coa ligase fadd15	-1.43
GWH15_16015	Putative propionyl-coa carboxylase beta chain 5	-1.42
GWH15_02240	Capsular glucan synthase	-1.42
GWH15_08525	Inosose dehydratase	-1.40
GWH15_01075	Catechol-2,3-dioxygenase	-1.36
GWH15_17815	Aldose 1-epimerase	-1.35
GWH15_14330	Scyllo-inositol 2-dehydrogenase (NAD(+))	-1.34
GWH15_18470	Adenine phosphoribosyltransferase	-1.32
GWH15_07615	4-hydroxy-4-methyl-2-oxoglutarate aldolase	-1.30
GWH15_02610	Monoterpene epsilon-lactone hydrolase	-1.26
GWH15_04030	Cystathionine beta-lyase metC	-1.23
GWH15_05255	Phosphoheptose isomerase	-1.23
GWH15_06910	4-alpha-glucanotransferase	-1.23
GWH15_11420	Copper-exporting P-type atpase A	-1.21
GWH15_13265	Hypothetical protein	-1.21

ORF (continued)	Description	log2 FC
GWH15_14105	Free methionine-R-sulfoxide reductase	-1.20
GWH15_13500	Methylmalonate semialdehyde dehydrogenase [acylating]	-1.19
GWH15_11370	Gluconate 5-dehydrogenase	-1.17
GWH15_09530	Riboflavin biosynthesis protein ribbA	-1.17
GWH15_18555	Ribonuclease R	-1.16
GWH15_06590	Hypothetical protein	-1.16
GWH15_14315	Hypothetical protein	-1.16
GWH15_12790	2-hydroxyhexa-2,4-dienoate hydratase	-1.14
GWH15_00640	Hypothetical protein	-1.14
GWH15_10795	Xylan 1,4-beta-xylosidase	-1.13
GWH15_02440	Acetyl-coenzyme A synthetase	-1.13
GWH15_11970	Aldehyde reductase yahK	-1.12
GWH15_06545	Putative glycerophosphoryl diester phosphodiesterase 1	-1.11
GWH15_02965	Hypothetical protein	-1.11
GWH15_13360	Ureidoglycolate lyase	-1.11
GWH15_14415	P-protein	-1.10
GWH15_11385	Hypothetical protein	-1.10
GWH15_05145	Hypothetical protein	-1.08
GWH15_15710	Putative glucose-6-phosphate 1-epimerase	-1.07
GWH15_06770	Hypothetical protein	-1.06
GWH15_14470	Hypothetical protein	-1.06
GWH15_14895	Oxalate decarboxylase oxdC	-1.03
GWH15_17295	Hypothetical protein	-1.03
GWH15_14850	Farnesyl diphosphate synthase	-1.03
GWH15_17980	Pyruvate dehydrogenase [ubiquinone]	-1.02
GWH15_08905	UDP-glucose 4-epimerase	-1.01
GWH15_10410	Trigger factor	-1.00

Table S6: Number of significantly up and downregulated proteins assigned to KEGG metabolic pathways in the proteomic profile of *Granulicella* strain WH15 in high cellobiose conditions.

KEGG metabolic pathways	Upregulated	Downregulated
Metabolic pathways	15	56
Biosynthesis of secondary metabolites	6	22
Biosynthesis of antibiotics	3	19
Microbial metabolism in diverse environments	1	14
Carbon metabolism	1	12
Biosynthesis of amino acids	3	11
Fatty acid metabolism	1	8
Lipopolysaccharide biosynthesis	0	6
Fatty acid biosynthesis	0	5
Butanoate metabolism	0	5
Propanoate metabolism	0	5
Valine, leucine and isoleucine degradation	0	5
Starch and sucrose metabolism	2	4
Galactose metabolism	0	4
Amino sugar and nucleotide sugar metabolism	0	4
Glyoxylate and dicarboxylate metabolism	1	4
Fatty acid degradation	0	4
Phenylalanine, tyrosine and tryptophan biosynthesis	1	4
Pantothenate and CoA biosynthesis	0	4
Benzoate degradation	0	4
Pentose phosphate pathway	0	4
Other glycan degradation	0	3
Quorum sensing	3	3
Novobiocin biosynthesis	0	3
Phenylalanine metabolism	1	3
Cysteine and methionine metabolism	1	3
Methane metabolism	0	3
2-Oxocarboxylic acid metabolism	0	3

KEGG metabolic pathways (continued)	Upregulated	Downregulated
Cyanoamino acid metabolism	0	2
Folate biosynthesis	0	2
Glycolysis / Gluconeogenesis	0	2
Tyrosine metabolism	0	2
Fructose and mannose metabolism	0	2
Histidine metabolism	0	2
Pentose and glucuronate interconversions	0	2
Pyruvate metabolism	0	2
Nitrogen metabolism	0	2
Arginine biosynthesis	0	2
Synthesis and degradation of ketone bodies	0	2
Valine, leucine and isoleucine biosynthesis	0	2
beta-Alanine metabolism	0	2
Riboflavin metabolism	0	2
Geraniol degradation	0	2
Alanine, aspartate and glutamate metabolism	0	2
Biotin metabolism	0	1
ABC transporters	4	1
Sulfur metabolism	0	1
Bacterial chemotaxis	2	1
Arginine and proline metabolism	1	1
Thiamine metabolism	0	1
Glycine, serine and threonine metabolism	0	1
Xylene degradation	0	1
Purine metabolism	2	1
Glutathione metabolism	2	1
Oxidative phosphorylation	3	1
Cationic antimicrobial peptide (CAMP) resistance	1	1
Degradation of aromatic compounds	0	1
Two-component system	9	1
C5-Branched dibasic acid metabolism	0	1
Pyrimidine metabolism	1	1
RNA degradation	0	1
Terpenoid backbone biosynthesis	0	1
alpha-Linolenic acid metabolism	0	1
Sphingolipid metabolism	1	1
Selenocompound metabolism	1	1
Glycerophospholipid metabolism	1	1
D-Glutamine and D-glutamate metabolism	0	1
Peptidoglycan biosynthesis	0	1
Chlorocyclohexane and chlorobenzene degradation	0	1
Inositol phosphate metabolism	0	1
Biosynthesis of unsaturated fatty acids	1	0
Lysine biosynthesis	1	0
Sesquiterpenoid and triterpenoid biosynthesis	1	0
beta-Lactam resistance	2	0
Bacterial secretion system	4	0
Tryptophan metabolism	2	0
Ribosome	2	0
Protein export	2	0
Nucleotide excision and repair	1	0
Mismatch repair	2	0
One carbon pool by folate	1	0
DNA replication	1	0
RNA polymerase	1	0
Nicotinate and nicotinamide metabolism	1	0
Phenazine biosynthesis	1	0

Chapter 4

Impact of different trace elements on the growth and proteome of two strains of Granulicella

Ohana Y.A. Costa, Chidinma Oguejiofor, Daniela Zühlke, Cristine C. Barreto, Katharina Riedel, Eiko E. Kuramae

Modified version published as: Costa OYA, Oguejiofor C, Zühlke D, Barreto CC, Riedel K, Kuramae EE (2020). Impact of Different Trace Elements on the Growth and Proteome of Two Strains of *Granulicella*, class "Acidobacteriia". **Frontiers in Microbiology**. 11:1227.

Abstract

Members of the phylum *Acidobacteria* are difficult to isolate and grow. For some isolated strains, recent studies suggested that trace elements are needed in culture media for proliferation but the impact of these trace elements on their growth and metabolism is not known to date. Here we evaluated the effect of the trace element solution SL-10 on the growth of two strains (5B5 and WH15) of *Granulicella* sp. and studied changes in the proteome in response to manganese (Mn). Growth of *Granulicella* species was enhanced in nutrient media amended with trace element solution SL-10. When trace elements were tested separately, mangenese was shown to enhance growth of *Granulicella* species which was associated with a higher tolerance to this metal compared to seven other metal ions. Variations in tolerance to metal ion concentrations among the two strains suggest different mechanisms to cope with metal ion homeostasis and stress. Comparative proteome analysis revealed different responses to manganese for the two *Granulicella* strains. Strain 5B5 had more upregulated proteins (57), while strain WH15 had more downregulated proteins (112). Further comparisons demonstrated that no upregulated or downregulated proteins were shared between the two strains. In strain 5B5, a higher number of upregulated proteins that can use Mn^{2+} as co-factor was detected. Genome analyses of the two strains also revealed that the most common transcriptional regulator of Mn homeostasis *mntR* was not present. Instead several candidate transporters were found that could be involved in Mn homeostasis of *Granulicella*. We postulate that these transporters may enhance the adaptive ability of *Granulicella* to metal-enriched environments, such as the Mn-rich decaying wood environment from which these strains were isolated.

Keywords: *Acidobacteria*, *Granulicella*, Genome, Proteome, Manganese, Metabolism.

1. Introduction

Despite being widespread and dominant in soil ecosystems (Kuramae *et al.*, 2012, Navarrete *et al.*, 2013, Pereira de Castro *et al.*, 2016), the phylum *Acidobacteria* has a low number of cultivated representatives, due to difficulties in isolation and propagation under laboratory conditions (Dedysh & Yilmaz, 2018). Most *Acidobacteria* isolates are slow growers and can take weeks to months to develop colonies (Eichorst *et al.*, 2011, de Castro *et al.*, 2013). Recently, changes in traditional culture methods and application of unconventional culture media composition have increased the number of new *Acidobacteria* isolates considerably. Currently, 62 species have been described (NCBI Resource Coordinators, 2016), while in 2011 only 14 species had been isolated and characterized (de Castro, 2011). Modifications in culture media and cultivation conditions, such as low concentration of nutrients (Janssen *et al.*, 2002, Stevenson *et al.*, 2004), higher CO_2 concentrations (Stevenson *et al.*, 2004), unusual or complex polysaccharides as carbon sources (Pankratov *et al.*, 2008, Eichorst *et al.*, 2011), longer incubation periods (de Castro *et al.*, 2013), addition of humic acids and quorum-sensing molecules (Stevenson *et al.*, 2004), employment of soil solution equivalents and inhibitors for unwanted microorganisms (de Castro *et al.*, 2013, Foesel *et al.*, 2013), are strategies that have been applied for the enrichment and isolation of new *Acidobacteria* species.

Once the isolates are obtained, better cell proliferation can be achieved with richer culture media, containing higher concentrations of nutrients (de Castro *et al.*, 2013). For instance, trace elements can be used to improve microbial growth and biomass in laboratory conditions, even though the specific requirements among strains and species are variable (Banerjee *et al.*, 2009, Merchant & Helmann, 2012). Metal ions, such as Fe, Mn, Zn, and Cu are fundamental for microbial metabolism, being required at low concentrations (Abbas & Edwards, 1990). They play an important role in biological processes, acting as co-factors of enzymes (Wintsche *et al.*, 2016), activating metalloregulators and trace element dependent proteins (Hantke, 2001, Zhang *et al.*, 2009), forming functional complexes with secondary metabolites (Morgenstern *et al.*, 2015, Locatelli *et al.*, 2016) and promoting the detoxification of reactive oxygen species (Kehres & Maguire, 2003, Locatelli *et al.*, 2016).

Although some culture media used for *Acidobacteria* growth and isolation are supplemented with trace elements (de Castro *et al.*, 2013, Navarrete *et al.*, 2013), the impact of these metals on their growth and metabolism is not yet known. Although metal ions are essential for many biological processes, they can be toxic at high concentrations (Puri *et al.*, 2010). Metal ions cannot be synthesized or degraded, therefore cellular homeostasis of metals relies mostly on transport, which involves several mechanisms that sense, uptake, immobilize or pump metals out of the cell (Chandrangsu *et al.*, 2017). In the present study, we evaluated the effect of trace elements and particularly Mn on the growth of two strains of *Granulicella* sp. WH15 and 5B5, derived from decaying wood, *Acidobacteria* subgroup 1 (class *Acidobacteriia*) (Valášková *et al.*, 2009). We used the optimized culture medium PSYL5 (Campanharo *et al.*, 2016), in order to boost the growth of the two strains, evaluated the impact of Mn through

proteome studies and performed genomic analyses on both strains.

2. Material and Methods

2.1. Acidobacteria strains

Two strains of *Acidobacteria*, 5B5 and WH15 belonging to *Granulicella* genus of subdivision 1 were used in this study. Both strains belong to the culture collection of the Netherlands institute of Ecology (NIOO-KNAW), department of Microbial Ecology. They were isolated from wood in advanced decay stage, in association with the white-rot fungus *Hypholoma fasciculare*, in the Netherlands (Valášková *et al.*, 2009). The genome of strain WH15 is deposited at NCBI with accession number CP042596 while the genome of 5B5 was sequenced in this study.

2.2. Trace elements solution (SL10) and individual trace elements effect on bacterial growth

The effect of trace element solution SL 10 (Atlas, 2010) on the growth of both bacterial strains was evaluated for two different concentrations (1 ml and 10 ml) of the solution per L of PSYL5 culture medium. PSYL5 medium was composed of (g/L): 1.8 KH_2PO_4, 0.2 $MgSO_4.7H_2O$, 30 sucrose and 1.0 yeast extract; pH was adjusted to 5.0 (Campanharo *et al.*, 2016). Culture medium without the amendment of SL10 solution was used as a control. Seven-day-old cell suspensions of both strains were inoculated in 70 ml of culture medium to an OD_{600nm} 0.01. The cultures were incubated under aeration for 7 days at 30 °C and a constant rotation rate of 50 rpm. Every 24 hours the optical density of the cultures was measured with an Eppendorf photometer at a wavelength of 600 nm (Eppendorf, Hamburg, Germany). For the evaluation of the different trace elements on the growth of both strains, individual trace element stock solutions and growth curves were executed for each metal separately, using the same growth conditions described above. The composition of the trace element solution SL10 and the final concentration of each trace element in culture medium is shown in Table 1. The metal ion that produced a significantly higher growth in comparison with the control was selected for further experiments. All experiments were executed in triplicates.

Statistical analysis was performed using Sigmaplot v14. Normality of the data was checked using Shapiro-Wilk test. Two Way Repeated Measures ANOVA was used to test the effect of SL10 solution and individual trace elements on the growth rate of WH15 and 5B5 strains.

Table1: Composition of trace element solution SL10 and final concentration (µM) of each individual metal in culture medium.

Reagents	SL 10 composition (mg/l)	Final µM concentration in culture medium (1ml/L SL10)
$FeCl_2.4H_2O$	1.500	7.54
$ZnCl_2$	0.070	0.51
$MnCl_2.4H_2O$	0.100	0.51
H_3BO_3	0.006	0.10
$CoCl_2.6H_2O$	0.190	0.80
$CuCl_2.2H_2O$	0.002	0.01
$NiCl_2.6H_2O$	0.024	0.10
$Na_2MoO_4.2H_2O$	0.036	0.15

2.3. Genome of *Granulicella* sp. 5B5

The *Granulicella* sp. 5B5 strain obtained from the collection of the Netherlands Institute of Ecology (NIOO-KNAW) was grown on 1/10 TSB agar medium (Valášková *et al.*, 2009) at pH 5.0 for 3 days at 30 °C. The bacterial cells were harvested and the genomic DNA was extracted using MasterPure™ DNA Purification Kit (Epicentre, Madison, WI) according to manufacturer's instructions. A total of 10 mg of DNA was sent to the Genomics Resource Center (Baltimore, USA) for a single long insert library (15kb-20kb), that was constructed and sequenced in one SMRTcell using the PacBio RS II (Pacific Biosciences, Inc.) sequencing platform. *De novo* assembly was performed with the help of SMRT Analysis software v2.2.0 (Pacific Biosciences) featuring HGAP 2 (Chin *et al.*, 2013), and subsequent correction with Pilon 1.16 (Walker *et al.*, 2014) to reveal a circular replicon: a 3,928,701 bp chromosome (G+C content 61,1%; 58× coverage). Automatic gene prediction and annotation was performed by using Prokka (Seemann, 2014) and RAST genome annotation server (http://rast.nmpdr.org/) (Aziz *et al.*, 2008). Genes were mapped to COG and KEGG IDs using the COG database (2014 release) (Galperin *et al.*, 2015) and KEGG database (release 2013) (Kanehisa, 2000), using eggNOG mapper. The CAZyme contents of 5B5 genome were determined by identifying genes containing CAZyme domains using the dbCAN2 meta server (cys.bios.niu.edu/dbCAN2) (Zhang *et al.*, 2018), according to the CAZy (Carbohydrate-Active Enzyme) database classification (Lombard *et al.*, 2014). Only CAZyme domains predicted by at least two of the three algorithms (DIAMOND, HMMER and Hotpep) employed by dbCAN2 were kept. Circular genome map was drawn using CGView software (Stothard & Wishart, 2004). Average Nucleotide Identity (ANI) between strains 5B5 and WH15 was calculated using the webtool ANI calculator, available at https://www.ezbiocloud.net/tools/ani (Yoon *et al.*, 2017). The *Granulicella* sp. 5B5 strain genome is deposited at NCBI with accession number CP046444.

2.4. Heavy metal resistance assays and Metal Resistance Gene (MRG) annotation

The resistance of strains 5B5 and WH15 to varied metal ion concentrations was tested in solid culture medium PSYL5 pH 5. Five concentrations (0.5 mM, 1 mM, 2 mM, 5 mM, 10 mM) of 9 metal ion sources were tested: $ZnCl_2$, $NiCl_2$, $MnCl_2$, $CoCl_2$, $CuCl_2$, $NaMoO_4$, $AlCl_2$, $CdCl_2$ and

$C_8H_4K_2O_{12}Sb_2$. As a control, an *Escherichia coli* DH5α strain, with known low metal resistance was used. Six colonies of each strain previously grown on PSYL5 solid medium without metal were inoculated in the culture media with each different metal concentration. After 7 days of growth at 30 °C, colonies were reinoculated on a new plate with the same metal concentration, in order to confirm growth. If colonies did not develop within 7 days, plates were incubated for extra 7 days. Colonies were reinoculated 3 times for confirmation. When the strains were resistant to the highest concentration of metal used (10 mM), we performed additional tests using higher metal concentrations (15 mM, 20 mM, 25 mM, 30mM, and 40 mM). In order to identify genes that could be involved in metal ion homeostasis, we searched the genomes of both strains against the experimentally confirmed and predicted BacMet databases using BacMet Scan (Pal *et al.*, 2014) with less strict parameters (40% similarity), due to the high quantity of hypothetical proteins in the genomes of both bacteria.

2.5. Acquisition of cytosolic proteome with and without manganese treatment by mass spectrometry and data analysis

For the proteome analysis, we analysed the effects of the metal ion which significantly improved the growth yield of both strains in comparison to the control without metal ion. Therefore, Mn was selected for further experiments. The growth curves of WH15 and 5B5 with added manganese ($MnCl_2$) and controls without trace elements were repeated, using the same parameters as described above. Cells were collected at day 4 of the growth curve, when the differences in the OD_{600nm} between manganese treatment and control treatment started to be statistically significant. A total of 3 ml of bacterial cells per replicate (n=6 for each strain) were harvested by centrifugation at 10,015 x g at 4 °C for 10 min. Pellets were washed twice with 1 mL of TE buffer and finally resuspended in 1 ml TE buffer. A volume of 500 µL of cell suspension was transferred into 2 mL screw cap tubes filled with 500 µL glass beads (0.1 mm in diameter; Sarstedt, Germany) and mechanically disrupted using Fastprep (MP Biomedicals) for 3 x 30 sec at 6.5 m/s; with on ice incubation for 5 min between cycles. To remove cell debris and glass beads, samples were centrifuged for 10 min at 4 °C at 21,885 x g, followed by a second centrifugation (30 min at 4 °C at 21,885 x g) to remove insoluble and aggregated proteins. The protein extracts were kept at -20 °C. Protein concentration was determined using RotiNanoquant (Carl Roth, Germany). Proteins were separated by SDS-PAGE. Protein lanes were cut into ten equidistant pieces and in-gel digested using trypsin as described earlier (Grube *et al.*, 2014). Tryptic peptides were separated on an EASY-nLC II coupled to an LTQ Orbitrap Velos using a non-linear binary 76 min gradient from 5 – 75 % buffer B (0.1 % acetic acid in acetonitrile) at a flow rate of 300 nL/min and infused into an LTQ Orbitrap Velos (Thermo Fisher Scientific, USA) mass spectrometer. Survey scans were recorded in the Orbitrap at a resolution of 60,000 in the m/z range of 300 – 1,700. The 20 most-intense peaks were selected for CID fragmentation in the LTQ. Dynamic exclusion of precursor ions was set to 30 seconds; single-charged ions and ions with unknown charge

were excluded from fragmentation; internal calibration was applied (lock mass 445.120025). For protein identification resulting MS/MS spectra were searched against a database containing protein sequences of *Granulicella* sp. strain 5B5 or *Granulicella* sp. strain WH15 and common laboratory contaminants (9,236 entries or 7,782 entries, respectively) using Sorcerer-Sequest v.27, rev. 11 (Thermo Scientific) and Scaffold v4.7 (Proteome Software, USA) as described earlier (Stopnisek *et al.*, 2016). Relative quantification of proteins is based on normalized spectrum abundance factors (NSAF (Zhang *et al.*, 2010)). The mass spectrometry proteomics data were deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE (Perez-Riverol *et al.*, 2019) partner repository with the dataset identifier PXD016551.

Statistical analysis was done using MeV (Saeed *et al.*, 2003); t-test was applied for proteins that were identified in at least two replicates of the respective condition. T-test of z-transformed normalized data were performed with the following parameters: unequal group variances were assumed (Welch approximation), P-values based on all permutation with P=0.01, significance determined by adjusted Bonferroni correction. Only significantly changed proteins showing at least 1.5-fold changes between conditions were considered for further analysis. Furthermore, so-called on/off proteins, that were only identified in one condition were analysed. Functional classification of *Granulicella* sp. strain 5B5 and WH15 proteins was carried out using eggNOGmapper (http://eggnogdb.embl.de/#/app/emapper) (Huerta-Cepas *et al.*, 2017), COG (Galperin *et al.*, 2015) and KEGG databases (Kanehisa, 2000). In order to identify proteins that could be involved in metal ion homeostasis, we searched proteins with significantly changed amounts against the experimentally confirmed and predicted BacMet databases using BacMet Scan (Pal *et al.*, 2014) with less strict parameters (40% similarity), due to the high quantity of hypothetical proteins in the genomes of both bacteria. Voronoi treemaps for visualization of proteome data were generated with Paver software (Decodon GmbH, Germany).

3. Results

3.1. Effects of the trace element solution SL10 on growth

The addition of trace element solution SL10 in liquid culture medium produced a significant effect ($p<0.001$) on the growth of both strains of *Granulicella* sp. Both concentrations (1X and 10X) of trace element solution (SL10) significantly increased 5B5 strain growth, with the highest growth recorded for 1X and 10X concentrations of SL10 (Figure 1a). In addition, the cultures showed a longer lag phase for 10X concentration of SL10 (Figure 1a).

WH15 strain had a significantly higher growth rate with the addition of 1X SL10 compared to control and 10X SL10 (Figure 1b), except at day one of incubation. Differently from strain 5B5, 10X SL10 did not enhance the growth of WH15, having instead the opposite effect (Figure 1b).

3.2. Effect of individual trace elements on growth

Of all the trace elements, manganese (Mn) and copper (Cu) significantly increased the growth of strain 5B5 compared to control starting from day three of the incubation period until the end ($p<0.001$) (Figure 1c). Iron (Fe) significantly increased the growth of 5B5 strain only at day six, when compared with the control. Boron (B) Zinc (Zn), cobalt (Co), nickel (Ni) and molydbdate (Mo) did not have any significant effect on the growth of the strain throughout the duration of the experiment (Figure 1c and d).

Throughout the incubation period, Mn was the only trace element that significantly ($p<0.001$) increased the growth of strain WH15 in comparison with the control with no metal (Figure

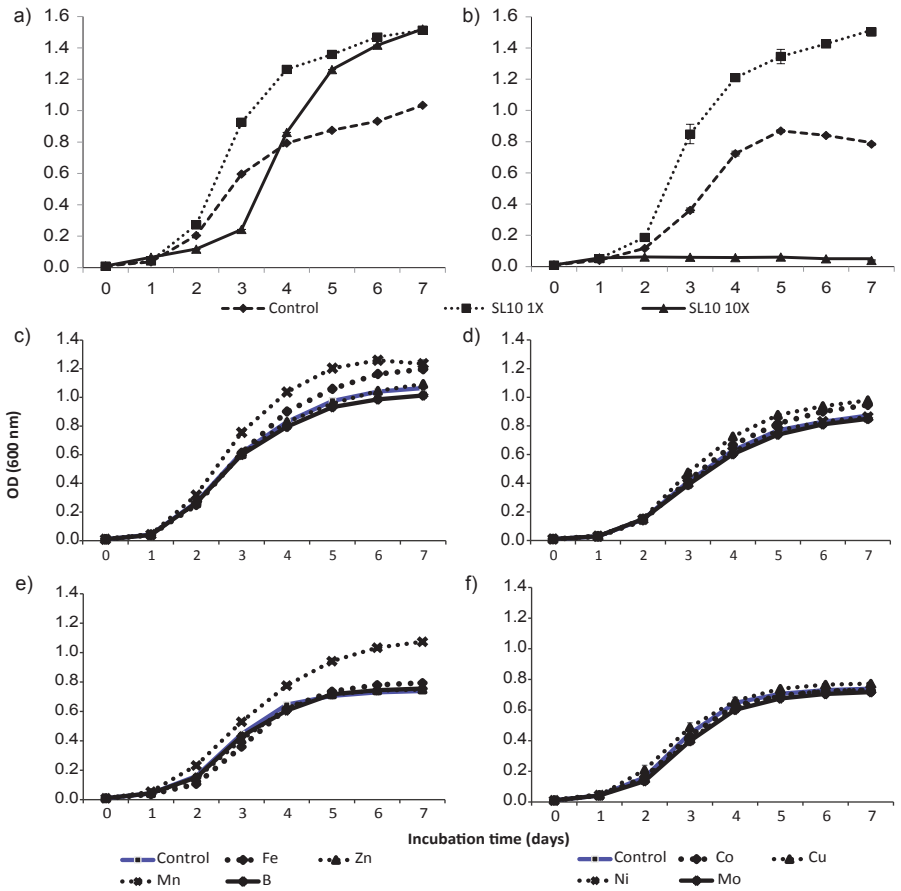

Figure 1: Growth curves of *Granulicella* sp.5B5 and WH15 strains on PSYL5 liquid culture medium with different concentrations of trace element solution (SL10) and individual metal ions. a) Strain 5B5 with SL10 solution: control (no SL10), 1X SL10 and 10X SL10; b) Strain WH15 with SL10 solution: control (no SL10), 1X SL10 and 10X SL10; c) 5B5 with individual metal ions: control (no metal), Fe, Zn, Mn, B; d) 5B5 with individual metal ions: control (no metal), Co, Cu, Ni, Mo; e) WH15 with individual metal ions: control (no metal), Fe, Zn, Mn, B; f) WH15 with individual metal ions control (no metal)), Co, Cu, Ni, Mo. The error bar is the standard error of the mean (n=3) and indicates differences in response variable between different treatments. Fe: $FeCl_2.4H_2O$; Zn $ZnCl_2$; Mn: $MnCl_2.4H_2O$; B: H_3BO_3; Co: $CoCl_2.6H_2O$; Cu: $CuCl_2.2H_2O$; Ni: $NiCl_2.6H_2O$; Mo: $NaMoO_4.2H_2O$.

1e). Fe, Zn, Co, Cu, Ni, Mo and Bo did not have any significant effect on the growth of WH15 as compared to the control (Figure 1e and 1f).

3.3. Granulicella sp. strain 5B5 genome annotation and CAZymes

The assembled genome of *Granulicella* sp. 5B5 is 3,928,701 bp, with 61.1% GC content, 3,306 proteins and only one rRNA operon. Functional annotation using COG (Cluster of Ortholog Groups) and RAST analysis resulted in the classification of 2,615 genes into 20 COG functional groups and the annotation of 1,260 genes to RAST subsystems. The properties of the genomes of strains 5B5 and also WH15 (sequenced previously (Chapter 3), Costa et al., submitted) are listed in Table 2. A circular genome map of 5B5 is depicted in Figure 2, together with that of strain WH15. The distribution of genes into COGs/RAST functional categories for strain 5B5 genome is depicted in Figure 3. Average Nucleotide Identity (ANI) (Figueras *et al.*, 2014) between strains WH15 and 5B5 was 72.75%, showing that the strains do not belong to the same species.

RAST analysis showed that only 37% of the annotated genes (1,260/3,374) could be assigned to subsystems. Among the subsystem categories present in the genome, carbohydrates, dormancy and sporulation had the highest and lowest feature counts, respectively (Figure 2b).

Table 2. Genomic features of *Granulicella* sp. strains 5B5 and WH15.

Genome	***Granulicella* sp. 5B5**	***Granulicella* sp. WH15**
Size (bp)	3,928,701	4,673,153
G+C content (%)	61.1	60.7
Number of coding sequences	3306	3,939
Number of features in Subsystems	1,260	1,496
Number of RNA genes	51	51
Number of contigs	1	1

Analysis with ANTISMASH v4.2.0 revealed the presence of 5 biosynthetic gene clusters (Table 3). The identified clusters showed potential for the production of terpenes, betalactone, type III polyketide synthases (T3PKS) and bacteriocin. Annotation with dbCAN (table 4) revealed the presence of 92 carbohydrate-associated enzymes, distributed in four classes: seven carbohydrate esterases (CE), 63 glycosyl hydrolases (GH), 20 glycosyl transferases (GT) and two polysaccharide lyases (PL), but no carbohydrate binding modules (CBM) or auxiliary activities (AA) were observed. Further evaluation of the CAZymes demonstrated the potential for the degradation of a wide range of carbohydrates, as the genome of strain 5B5 possessed CDSs for 48 CAZyme families, including as α- and β-glucosidases (GH1, GH13, GH3, GH31), α- and β-galactosidases (GH2, GH27, GH35, GH57), α- and β-mannosidases (GH1, GH38, GH125), rhamnosidases (GH28, GH106), fucosidases (GH29), xylosidases (GH39, GH43, GH54), arabinofuranosidase (GH43, GH54) and amylases (GH13, GH77). The cellulose synthase genes observed in other *Granulicella* genomes (Rawat *et al.*, 2013, Rawat *et al.*, 2013) were

not observed in the genome of strain 5B5.

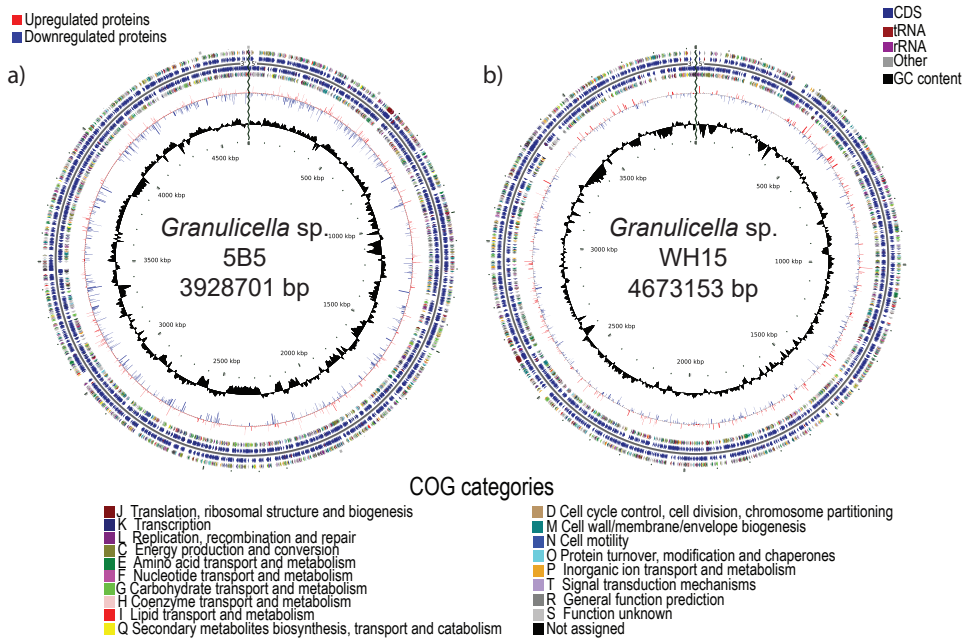

Figure 2: Graphical circular genome map of *Granulicella* sp. strains a) 5B5 and b) WH15. Rings indicate coding sequences and COG categories, GC content, upregulated (red) and downregulated (blue) proteins upon the addition of Mn to bacterial cultures.

Table 3: Biosynthetic gene clusters in *Granulicella* sp. strain 5B5 genome revealed by analysis with ANTISMASH.

Cluster	**Type**	**Most similar known cluster**
Cluster1	Terpene	Malleobactin (NRPS 11% similarity)
Cluster2	betalactone	
Cluster3	t3pks	
Cluster 4	Bacteriocin	
Cluster 5	Terpene	

Table 4: Number of genes from different CAZyme families observed in the genomes of strains 5B5 and WH15.

CAZyme family	***Granulicella* sp. 5B5**	***Granulicella* sp. WH15**
Auxiliary activity (AA)	0	13
Carbohydrate binding module (CBM)	0	22
Carbohydrate esterase (CE)	7	41
Cohesin	0	1
Glycoside hydrolase (GH)	63	86
Glycosyl transferase (GT)	20	52
Polyssacharide lyase (PL)	2	2
Total	**92**	**217**

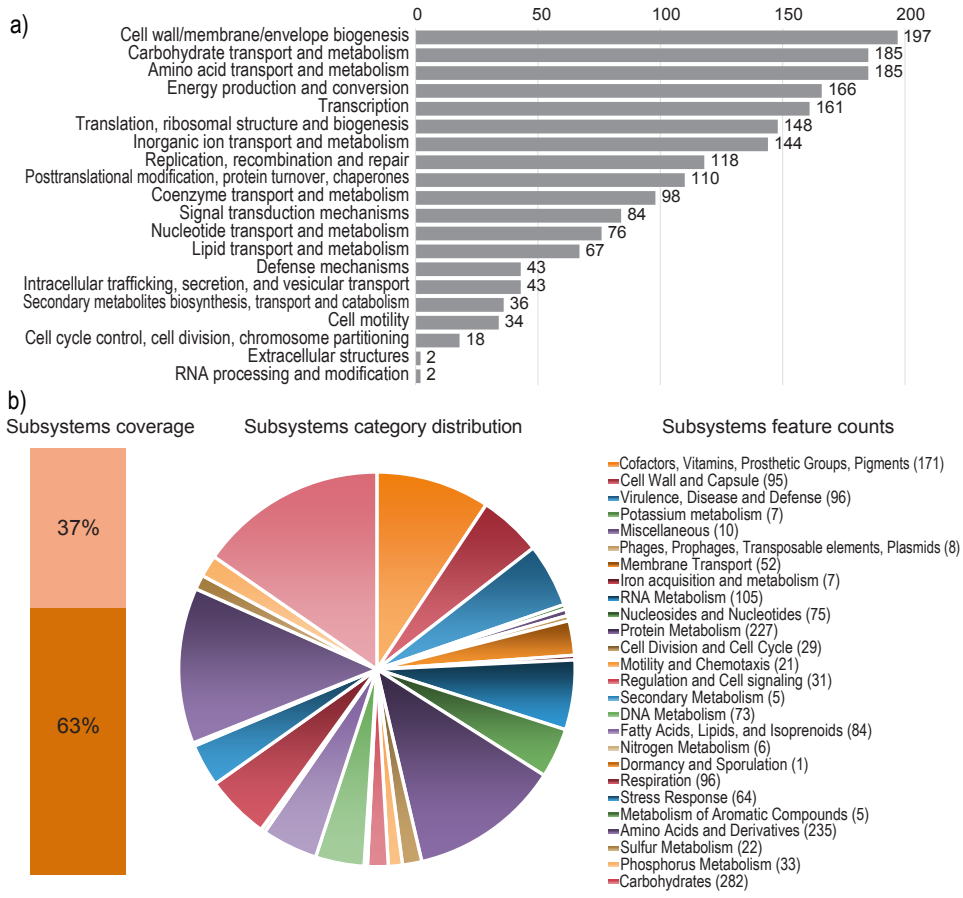

Figure 3: Statistics of COG and RAST subsystems annotations of *Granulicella* strain 5B5. a) COG categories distribution, showing number of genes annotated in each category. b) Subsystem category distribution. The light orange bar represents the percentage of proteins that could be annotated by RAST Server and the dark orange bar represents the proteins that were not annotated. The pie chart represents the percentage of proteins annotated to each subsystem category.

3.4. Metal resistance assays and Metal Resistance Gene (MRG) annotation

Metal resistance tests on agar medium demonstrated that both *Granulicella* strains were able to grow on $NiCl_2$ (max 2 mM) and $NaMoO_4$ (max 1 mM). Strain 5B5 was able to grow on 0.5 mM $ZnCl_2$, strain WH15 grew on 0.5 $AlCl_2$, and both strains could grow at concentrations of $MnCl_2$ up to 40 mM (Table 5).

WH15 genome search against BacMet experimentally-confirmed and predicted resistance genes databases revealed 28 ORFs and 78 ORFS, respectively, with hits similar (>45% identity) to genes involved in resistance of a wide range of metals such as As, Cd, Zn, Co, Cu, Fe, Mn, Mo, Ni and Zn, multidrug and metal efflux transporters and DNA binding response regulators (Table S1).

Table 5: Growth of *Granulicella* sp. strains WH15 and 5B5 in solid culture medium with different metal concentrations.

Metal source	Strain	concentration in mM					
		0.5	1	2	5	10	15-40
$ZnCl_2$	WH15	-	-	-	-	-	-
	5B5	+	-	-	-	-	-
$NiCl_2$	WH15	+	+	+	-	-	-
	5B5	+	+	+	-	-	-
$MnCl_2$	WH15	+	+	+	+	+	+
	5B5	+	+	+	+	+	+
$CoCl_2$	WH15	-	-	-	-	-	-
	5B5	-	-	-	-	-	-
$CuCl_2$	WH15	-	-	-	-	-	-
	5B5	-	-	-	-	-	-
$NaMoO_4$	WH15	+	+	-	-	-	-
	5B5	+	+	-	-	-	-
$AlCl_2$	WH15	+	-	-	-	-	-
	5B5	-	-	-	-	-	-
$CdCl_2$	WH15	-	-	-	-	-	-
	5B5	-	-	-	-	-	-
$C_8H_4K_2O_{12}Sb_2$	WH15	-	-	-	-	-	-
	5B5	-	-	-	-	-	-

+ : positive colony formation; -:no growth. Comparisons made with the control without metal.

In addition, strain WH15 possessed two copies of Mn transporter MntH, and two ORFs (GWH15_19170 and GWH15_03225), with 60.2 and 44% identity with the Mn transcriptional regulator *mntR*, respectively.

We obtained a similar profile for the 5B5 genome, with 65 ORFs that had hits higher than 45% identity against the experimentally confirmed database and 23 ORFs that had hits higher than 45% identity against the predicted database (Table S2). For both searches, genes involved in resistance to several metal ions, as well as multidrug and metal efflux transporters and transcription regulators were observed (Table S2). The genome of strain 5B5 also contains 2 copies of the *mntH* transporter, and 3 ORFs related to Mn transcriptional regulator *mntR*, as well as 3 ORFs similar to Mn ABC transporters *mntA*/*ytgA* and Mn efflux pump *mntP* (Table S2).

3.5. Manganese-responsive proteome of *Granulicella*

Since Mn had a significant effect on the growth of both *Granulicella* sp. strains, we further investigated the effects of Mn on cellular metabolism by a proteomics. At day 4, the differences in growth between control and Mn treatment started to be statistically significant for both strains (Figure 4), and therefore samples were collected at this particular time point for proteome analysis.

Proteome data for strain 5B5 showed that 1,028 proteins were detected in both treatments in at least two out of three replicates each. Overall, 216 proteins showed significantly different abundances, with 14 so-called on/off proteins, which were present in only one condition (Figure S1a). The proteome patterns of strain 5B5 under control and Mn treatments are depicted in Figure 5a. A total of 46 proteins were upregulated 1.5-fold and 11 proteins were

"on", while 43 proteins were downregulated 1.5-fold and 3 proteins were "off" in the Mn treatment. Among these differentially expressed proteins, 90 could be assigned to COG categories and 67 could be annotated to KEGG orthologs (Figure 6).

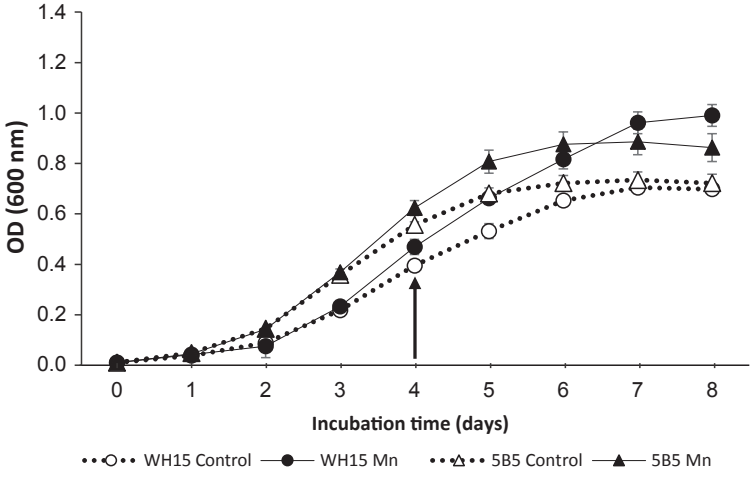

Figure 4: Growth curves of *Granulicella* sp. strains 5B5 and WH15 in PSYL 5 liquid culture medium with addition of Mn and control without addition of metal. The arrow indicates the timepoint when samples were collected for proteomics analysis. The error bar is the standard error of the mean and indicates differences in response variable between different treatments.

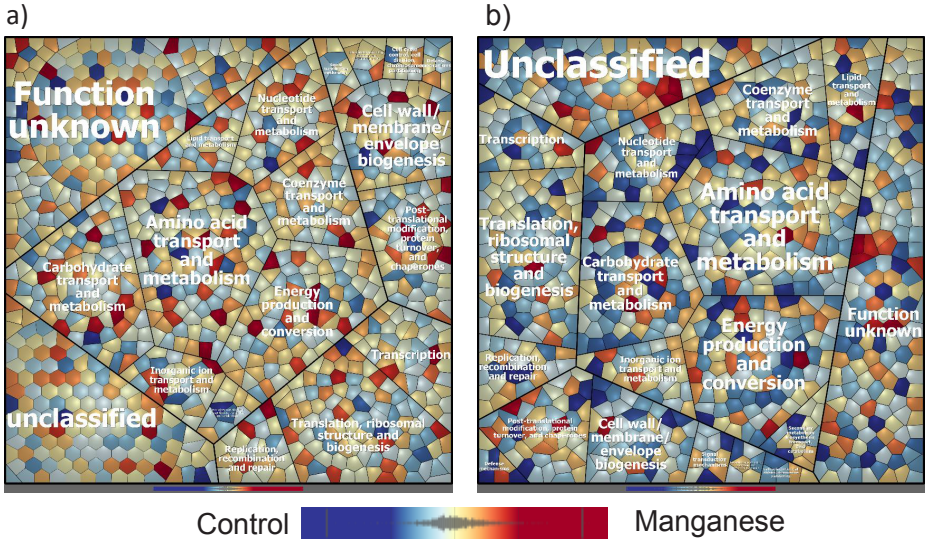

Figure 5: Voronoi treemap visualization of protein expression patterns of *Granulicella* sp. strains a) 5B5 and b) WH15 spectrum under control and manganese treatments. Functional classification was done using Prophane 2.0 (www.prophane.de) and is based on eggNOG database, B function level "subrole". Each cell represents a quantified protein; proteins are clustered according to their function. Proteins with higher amount under control conditions (no metal) are depicted in blue, proteins with higher amount in manganese treatment are depicted in red.

The qualitative analysis of the proteomic data for strain WH15 demonstrated that, overall, 909 proteins were identified in both conditions in two out of three replicates each. In total, 171 proteins showed significant differences between Mn and control conditions (t-test, p=0.01) (Figure S1b). The proteome patterns of strain WH15 under control and manganese treatments are depicted in Figure 5b.

Comparisons between treatments showed that 16 proteins were upregulated at least 1.5-fold, while 93 proteins were downregulated at least 1.5-fold. In addition, 19 proteins were

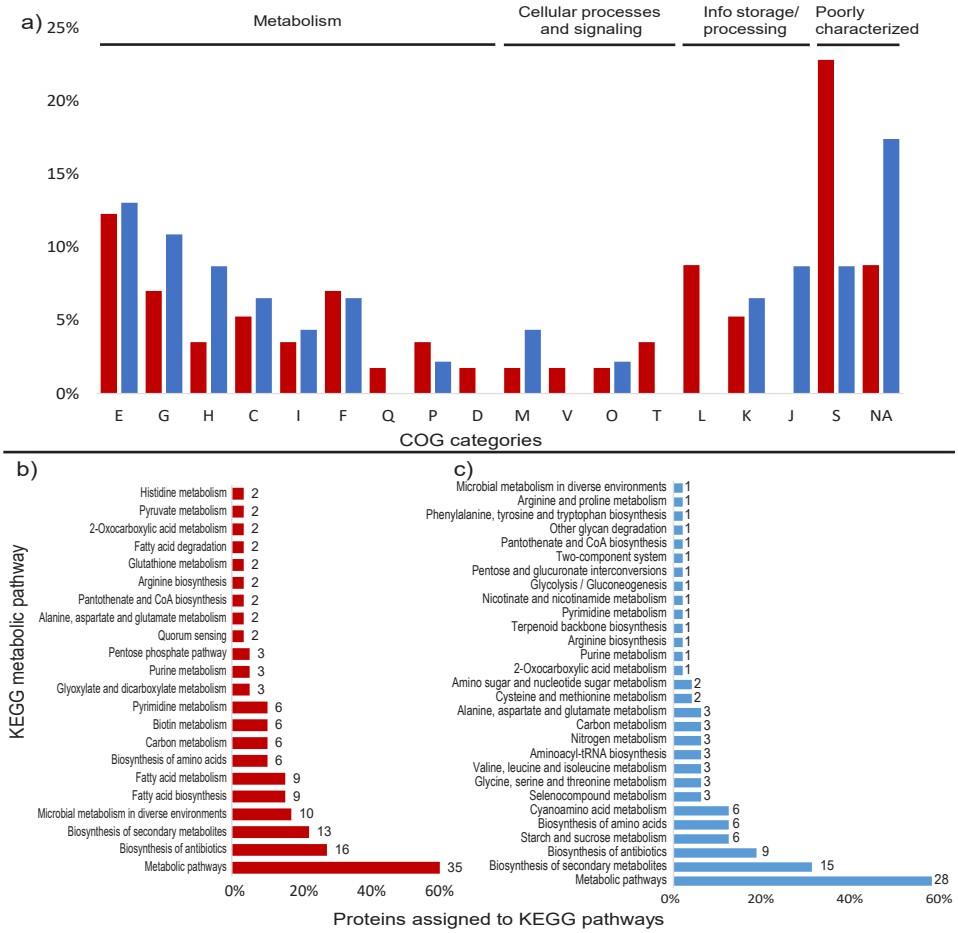

Figure 6: Differentially expressed proteins assigned to COG categories and KEGG pathways in the proteomic profile of *Granulicella* strain 5B5 with the addition of Mn. a) Percentage of upregulated (red) and downregulated (blue) proteins assigned to COG categories; b) Number of upregulated proteins assigned to KEGG pathways (only pathways with more than one protein mapped are shown); c) Number of downregulated proteins assigned to KEGG pathways. E-Amino acid transport and metabolism; G- Carbohydrate transport and metabolism; H-Coenzyme transport and metabolism; C-Energy production and conversion; I-Lipid transport and metabolism; F-Nucleotide transport and metabolism; Q- Secondary metabolites; D-Cell cycle; N-Cell motility; M-Cell wall/membrane/envelope biogenesis; V-Defence mechanisms; P-Inorganic ion transport and metabolism; U-Intracellular trafficking; O-Post translational modification; T-Signal transduction mechanisms; L-Replication, recombination and repair; K-Transcription; J-Translation; S-Function unknown; R-General function and prediction; X-Mobilome.; NA-not assigned.

"off" in the Mn treatment and present only in control conditions. Among the significantly different proteins, 112 were assigned to COG categories and 89 were annotated to KEGG orthologs (Figure 7).

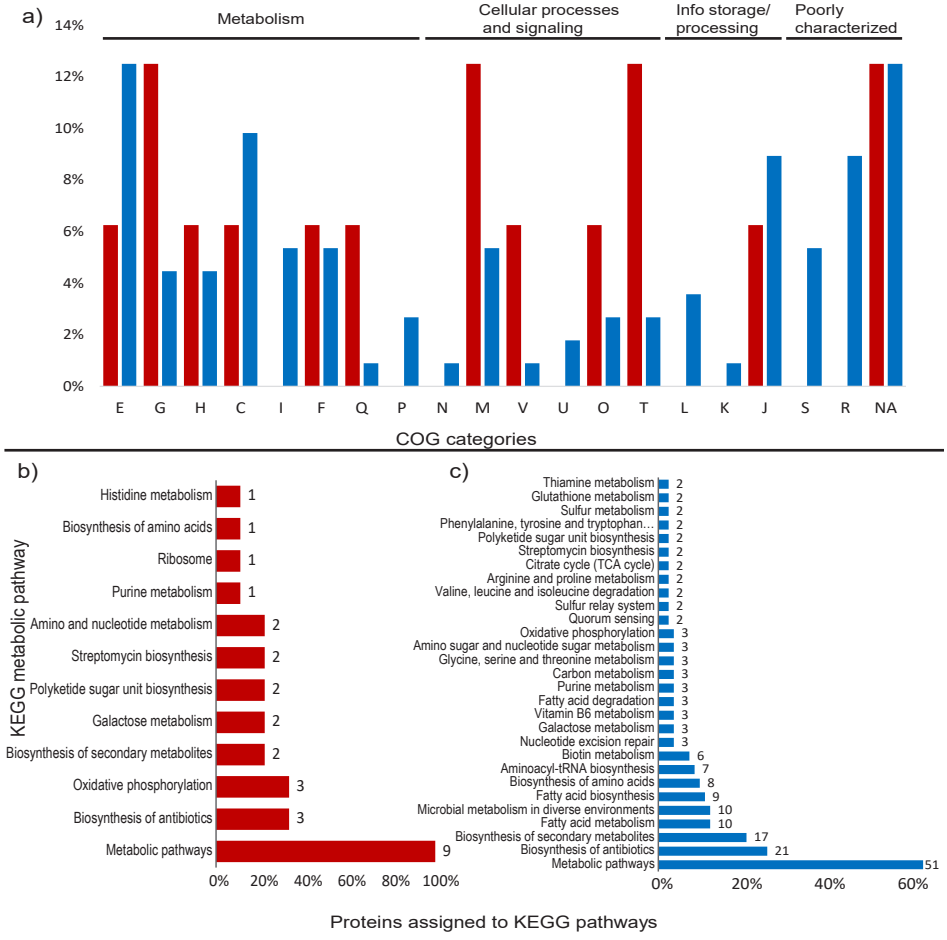

Figure 7: Differentially expressed proteins assigned to COG categories and KEGG pathways in the proteomic profile of *Granulicella* strain WH15 with the addition of Mn. a) Percentage of upregulated (red) and downregulated (blue) proteins assigned to COG categories; b) Number of upregulated proteins assigned to KEGG pathways; c) Number of downregulated proteins assigned to KEGG pathways (only pathways with more than one protein mapped are shown). E-Amino acid transport and metabolism; G- Carbohydrate transport and metabolism; H-Coenzyme transport and metabolism; C-Energy production and conversion; I-Lipid transport and metabolism; F-Nucleotide transport and metabolism; Q-Secondary metabolites; D-Cell cycle; N-Cell motility; M-Cell wall/membrane/envelope biogenesis; V-Defence mechanisms; P-Inorganic ion transport and metabolism; U-Intracellular trafficking; O-Post translational modification; T-Signal transduction mechanisms; L-Replication, recombination and repair; K-Transcription; J-Translation; S-Function unknown; R-General function and prediction; X-Mobilome.; NA-Not assigned.

Comparatively, proteome analysis revealed different responses to manganese for the two strains. Strain 5B5 had more upregulated proteins (57), while WH15 had more downregulated

proteins (112). Further comparisons demonstrated that no upregulated or downregulated proteins were shared between strains. In strain 5B5 a higher number of upregulated proteins that can use Mn^{2+} as co-factor was detected. For both strains, proteins annotated as Mn transporters were not detected in the proteomic profile.

3.5.1. Upregulated proteins

3.5.1.1. Strain 5B5

COG analysis showed that proteins were mainly distributed among the categories E-amino acid transport and metabolism (7), L-replication, recombination and repair (5), G-carbohydrate transport and metabolism (4) and F-nucleotide transport and metabolism (4) (Figure 6a). A total of 36 proteins were assigned to KEGG orthologs, that were mapped to 47 metabolic pathways, and some orthologs were mapped to more than one pathway. The majority of the proteins were mapped to 'general' metabolic pathways (35), biosynthesis of antibiotics (16), biosynthesis of secondary metabolites (13) and microbial metabolism in diverse environments (10) (Figure 6b), but no specific metabolic pathway was upregulated. Looking deeper into the upregulated proteins, we identified several enzymes that require Mn^{2+} or Mg^{2+} as cofactor, such as nucleoside diphosphate kinase Ndk, octaprenyl diphosphate synthase IspB, UDP-N-acetylmuramate--L-alanyl-gamma-D-glutamyl-meso-2,6-diaminoheptandioate ligase Mpl, adenine deaminase Ade, phosphate-specific transport system accessory protein PhoU, oxalate decarboxylase OxdD, phosphoenolpyruvate carboxykinase PckG and 3'-5' exoribonuclease YhaM (Table S3).

Search against BacMet Databases showed 25 proteins with hits (> 45% identity) against the experimentally confirmed database and 17 proteins with hits (> 45% identity) against the predicted metal resistance genes database (Table S4). The genes were mostly associated resistance/homeostasis of several metal ions, such as Fe, Cu, As, Ni, Co and Zinc. Interestingly, 4 ORFS were similar to metal ion transporters that could be involved in Mn homeostasis: ORF_05650 (hypothetical protein) had 43.5% identity with copper-translocating P-type ATPase CueA; ORF_03225 (hypothetical protein) had 31% identity with copper-translocating P-type ATPase CopA; ORF_06495 (TcrA) had 40% identity with copper-translocating P-type ATPase CopA and ORF_14875 (NatA_2) had 31% identity with metal ABC transporter ATP-binding protein TroB (Table S4).

3.5.1.2. Strain WH15

COG analysis showed that upregulated proteins distributed within several COG categories. The most common categories were: "G-carbohydrate transport and metabolism" (2), "M-cell wall/envelope/membrane biogenesis" (2) and "T-signal transduction mechanisms" (2) (Figure 7a). Within KEGG metabolic pathways, no specific pathway upregulation was observed.

A total of 9 upregulated proteins were assigned to KEGG orthologs, which were mapped to 12 KEGG metabolic pathways, since some orthologs were mapped to more than one

pathway. Most of the annotated proteins were mapped to 'general' metabolic pathways (9), biosynthesis of antibiotics (3) and oxidative phosphorylation (3) (Figure 7b).
Some of the upregulated proteins were ATP synthase subunit b, and the carbohydrate-associated enzymes putative sugar phosphate isomerase YwlF, UDP-glucose 4-epimerase GalE and dTP-4-dehydrorhamnose 3,5 epimerase RmlC (Table S5). The search against BacMet databases showed that 3 ORFs had hits against genes related to metal ion homeostasis. ORF GWH15_13825 (*cysO*) had 42% identity with predicted resistance gene *copA*, encoding a copper-exporting P-type ATPase and 31% similarity with the experimentally confirmed cation/multidrug efflux pump AdeG, which is part of AdeFGH efflux system. ORF GWH15_17845 (hypothetical protein) had 30% similarity with predicted resistance gene *rcnB/yohN*, a nickel/cobalt homeostasis protein; ORF GWH15_19690 (hypothetical protein) had 36% identity with predicted resistance gene *mtrA*, a DNA-binding response regulator.

3.5.2. Downregulated proteins

3.5.2.1. Strain 5B5

Annotation with COG database demonstrated that most proteins were distributed within the categories E –amino acid transport and metabolism (6), G-carbohydrate transport and metabolism (5), H-coenzyme transport and metabolism (4) and J-translation, ribosomal structure and biogenesis (4) (Figure 6a). Overall, 31 proteins were assigned to KEGG identifiers, which were mapped to 29 KEGG metabolic pathways. Most of the proteins were mapped to 'general' metabolic pathways (29), biosynthesis of secondary metabolites (15) and biosynthesis of antibiotics (9) (Figure 6b). Several proteins linked to general metabolism were repressed, but no specific metabolic pathway seemed to be repressed. Among the repressed proteins we observed enzymes involved in various cellular functions, such as cysteine synthase CysM, L-threonine dehydratase TdcB, ribonucleoside-diphosphate reductase subunit beta NrdB, putative glucose-6-phosphate 1-epimerase YeaD and carbonic anhydrase CynT (Table S6).

3.5.2.2. Strain WH15

Within COG categories, most of the downregulated proteins belonged to the categories 'E-aminoacid transport and metabolism' (14), 'C-energy production and conversion' (11), 'J-Translation, ribosomal structure and biogenesis' (10) and 'R-General function prediction' (10) (Figure 7a). Overall, 80 proteins were assigned to KEGG orthologs, which were mapped to 52 KEGG metabolic pathways, and some orthologs were mapped to more than one type of pathway. The majority of the proteins were mapped to 'general' metabolic pathways (51), biosynthesis of antibiotics (21), biosynthesis of secondary metabolites (17), fatty acid metabolism (10) and microbial metabolism in diverse environments (10) (Figure 7b), with no pathway specifically stimulated. Several enzymes involved in amino acid biosynthesis and metabolism were identified, such as tyrosyl, leucyl, alanyl and glycyl-tRNA synthetases

(TyrS, LeuS, AlaS and GlyS), leucyl aminopeptidase PepA, threonine synthase ThrC, xaa-pro-dipeptidyl-aminopeptidase PepQ, aminopeptidase YpdF, cysteine desulfurase IscS and prolyl tripeptidyl peptidase PtpA (Table S7).

4. Discussion

In this study, we evaluated the effect of trace element addition on the growth of two strains of *Granulicella*, belonging to phylum *Acidobacteria* subgroup 1. We observed that the growth in liquid medium of both strains was enhanced by the addition of Mn, to which the strains tolerated higher concentrations than other metal ions. Furthermore, variations in tolerance to metal ion concentrations suggest that the *Acidobacteria* strains possess different mechanisms to deal with metal stress. Strain 5B5 is likely more adapted to survive in an environment with higher concentration of several metal ions when compared to strain WH15.

When evaluated separately, Mn had a more pronounced effect on growth than other metal ions, but the mix of metals was more effective in enhancing bacterial growth, reflecting wide physiological needs and the importance of different metal ions in bacterial metabolism. For instance, *Escherichia coli* BW25113 growth was maximized with a mixture of Ni and Fe, which had a better effect than each metal separately and other metal ion mixtures (Trchounian *et al.*, 2016). The amendment of Mn to fermentation medium improved the growth of *Lactobacillus bifermentans*, increasing the production and activity of the enzyme glucose isomerase, necessary for biotechnological applications (Givry & Duchiron, 2008). Manganese is also an essential growth factor for *L. casei* and other species of lactobacilli, which is attributed to its role as a co-factor of enzyme lactate dehydrogenase, enhancing cell growth rate and biomass concentration (Fitzpatrick *et al.*, 2001, Lew *et al.*, 2013). On the other hand, Mn had no significant impact in the growth of *Halobacterium* (Joshi *et al.*, 2015).

Among the metals used for metal ion resistance testing, *Granulicella* strains WH15 and 5B5 only showed tolerance to Mn, exhibiting growth at the concentration of 40 mM Mn, which is higher than other bacterial strains. For instance, a resistant *Serratia marcescens* strain, isolated from Mn mine waters in Brazil, could grow on a maximum concentration of 6 mM Mn (Barboza *et al.*, 2017). Mn tolerance can vary widely in microorganisms, with a minimal inhibitory concentration ranging from 0.1 mM to 228.9 mM Mn in certain marine bacterial strains (Gillard *et al.*, 2019).

Manganese is essential for the growth and survival of most living organisms. It is a co-factor of a wide range of enzymes, being vital in specific metabolic pathways, such as sugar, lipid and protein catabolism (Jensen & Jensen, 2014), oxygenic photosynthesis in cyanobacteria (Kehres & Maguire, 2003, Cvetkovic *et al.*, 2010), signal transduction, stringent response, sporulation, and pathogenesis (Kehres & Maguire, 2003, Jensen & Jensen, 2014). One of the most widely known and studied Mn functions is the detoxification of reactive oxygen

species (ROS), where it is a redox-active co-factor in free radical detoxyfing enzymes, such as Mn-superoxide dismutase (MnSod) and mangani-catalase (Jakubovics & Jenkinson, 2001, Jensen & Jensen, 2014). Additionally, the detoxifying capaticities of Mn are not only enzyme-mediated, since non-protein complexes of Mn can also work as antioxidants when enzymes are not sufficient (Jensen & Jensen, 2014). Both *Granulicella* strains were isolated from decaying wood material, in association with the white rot fungus *Hypholoma fasciculare* (Valášková et al., 2009), where topsoil-litter samples have higher Mn concentrations of 101920 µg Mn/kg (Chapter 5). Since high concentrations of Mn can be observed in wood decay environments, especially when decomposition is caused by white rot fungi (Blanchette, 1984), tolerance to higher manganese concentrations could be a strategy to assure the survival of the studied *Acidobacteria* in this environment.

The evaluation of the genes in both did not reveal the presence of common genes involved in Mn regulation, the transcriptional regulator mntR (Jensen & Jensen, 2014). This result implies that the homeostasis of Mn in *Granulicella* strains is under control of another transcriptional regulator. On the other hand, both strains possessed two copies of the Mn transporter gene *mntH*. MntH seems to be the main responsible transporter in Mn influx, but it was already observed that Mn has a significant repressive effect in the expression of Mn transporters under manganese sufficiency, keeping Mn homeostasis and optimal levels of Mn inside the cells (Guedon *et al.*, 2003, Jensen & Jensen, 2014). Furthermore, the search of the genomes against BacMet databases demonstrated that both genomes possessed a wide range of transporters that are linked to the homeostasis of diverse metal ions.

Within the upregulated proteins from strain 5B5, three proteins were similar to copper P-type ATPase transporters and one protein was similar to the metal ion ATP-binding ABC transporter TroB. P-type ATPases, such as CtpC, in *Mycobacterium* species are responsible not only for Mn efflux, removing excess metal ion from the cells, but are also required for the metallation of proteins (Padilla-Benavides *et al.*, 2013). ABC transporters are as well responsible for Mn uptake, important in several bacterial species (Papp-Wallace & Maguire, 2006, Jensen & Jensen, 2014). In the proteome profile of strain WH15, we similarly observed a protein similar to the copper-exporting P-type ATPase and cation/multidrug efflux pump AdeG, which could be involved in maintaining optimal levels of Mn inside the bacterial cell. Furthermore, we found a protein similar to RcnB/YohN, which is an essential protein for Ni/Co homeostasis in *E. coli* (Bleriot *et al.*, 2011), and a protein similar to gene *MtrA*, which is involved in cell division control and cell wall metabolism in *M. tuberculosis* (Gorla *et al.*, 2018).

In addition to transporters, the proteomic analyses of both bacteria revealed other proteins which might be involved in the growth enhancement of both strains. The proteomic profile of strain WH15 had fewer upregulated enzymes in comparison with 5B5, demonstrating that Mn had more impact in transcription and protein expression in strain 5B5. Several enzymes that require Mn or Mg as cofactor were upregulated in the proteome of 5B5, but no enhancement of specific metabolic pathway was detected. Overall, the upregulated enzymes in strain 5B5

were involved in general metabolic pathways, which could be stimulated due to the higher metabolism necessary for a faster growth. For instance, nucleoside diphosphate kinase Ndk is a critical enzyme involved in the nucleotide metabolism of microorganisms, but is also part of posttranslational modification of proteins, as well as regulation of genes linked to quorum sensing, proteases and toxins (Yu *et al.*, 2017). Octaprenyl diphosphate synthase IspB is involved in the production of the lateral chain of ubiquinones, and is an essential enzyme for respiration and normal growth of *E. coli* (Okada et al., 1997). 3'-5' exoribonuclease YhaM was identified as a participating enzyme in mRNA turnover in *Bacillus subtilis* (Oussenko *et al.*, 2005). UDP-N-acetylmuramate--L-alanyl-gamma-D-glutamyl-meso-2,6-diaminoheptandioate ligase Mpl is a recycling enzyme that allows the constant remodeling of bacterial cell wall polymer occurring during cell growth and division (Herve *et al.*, 2007). Moreover, the categories replication, recombination and repair were only present within upregulated proteins, with proteins DNA helicase RecD2, DNA -binding protein Hup 2, protein RecA, metal-dependent hydrolase YcfH and UvrABC system protein UvrB. In addition, no specific pathway seemed to be repressed.

In the proteomic profile of strain WH15, we observed the upregulation enzymes involved in energy production, amino acid metabolism and transcription regulation. For instance, ATP synthase subunit β is part of the ATP synthase complex, which is involved in ATP synthesis and hydrolysis (Sokolov *et al.*, 1999). ATP phosphoribosyltransferase is an enzyme involved in histidine biosynthesis, a reaction that requires Mg, which can be substituted by Mn (Tebar *et al.*, 1977). Protein RsbV is a positive regulator of factor sigma β, which, in gram-positive bacteria is a key contributor to the resilience and survival of bacterial species to environmental conditions, such as variations in pH, osmotic stress or entry into stationary growth phase (Kullik & Giachino, 1997, Guldimann *et al.*, 2016). Phosphodiesterase CpdA is responsible for the hydrolysis of the second messenger cyclic AMP (cAMP), which controls cell responses to a variety of environmental conditions (Fuchs *et al.*, 2010). Similarly to the response of *Serratia marcescens* (Queiroz *et al*., 2018) to Mn stimulation, the limited number of upregulated proteins from strain WH15 could be attributed to the adaptation to the concentration of Mn used in the experiment, since it is an optimal growth condition. Furthermore, as a co-factor of several important enzymes (Crowley *et al.*, 2000, Jensen & Jensen, 2014), the presence of manganese might be improving bacterial growth by activating enzymes and enhancing metabolic activities involved in cell cycle and division, even when the enzymes were not differentially expressed. Among the repressed proteins, we found several amino acid synthases, that could be inhibited due to the ready availability of amino acids in the culture medium composition (Mader *et al.*, 2002), supplied by yeast extract.

The proteomic profiles of both strains were different, and did not exhibit the overexpression of specific pathways, indicating that Mn was more important in enhance enzymatic activity than to protein expression regulation. Finally, we did not find the most common transcriptional regulation of Mn homeostasis, implying that Mn regulation is performed by a

different gene or set of genes, but our analysis revealed candidate transporters that could be potentially involved in Mn homeostasis for *Granulicella* species. The presence of such type of transporters could facilitate the uptake of metal ions, improving the adaptability of bacteria to metal enriched environments (Campanharo *et al.*, 2016), promoting a tight regulation of metal ion homeostasis, as well as a resistance to high concentrations of metals.

Acknowledgments

O.Y.A. Costa was supported by an SWB grant from CNPq [202496/2015-5] (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

Supplementary Material

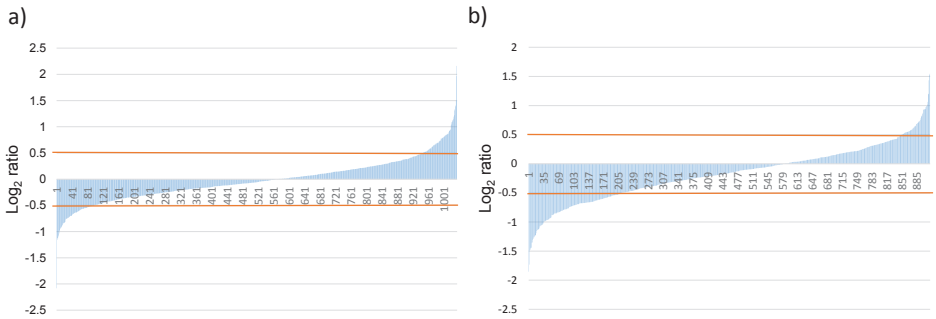

Figure S1: Expression pattern of proteins under manganese treatment of *Granulicella* sp. strain a) 5B5) and b) WH15. All proteins identified in at least two out of three replicates (excluding on/off proteins). Log_2 fold change is indicated by a red line.

Table S1: *Granulicella* strain WH15 gene annotation with BacMet databases. Annotation performed against the experimentally confirmed and predicted resistance genes databases using BacMet Scan software. Only genes with hits ≥ 40% identity in one of the databases are shown.

ORF	EXP	Compound	Identity (%)	PRE	Compound	Identity (%)
GWH15_16385	*arsB*	As	63.5	*acr3*	Arsenical-resistance protein	70.3
GWH15_11300	*mdtB*	Zn	60.5	*mdtB/yegN*	Multidrug transporter subunit MdtB	60.6
GWH15_16465	*hoxN*	Ni	60.0	*hoxN*	HoxN/HupN/NixA family Ni/Co transporter	60.0
GWH15_00760	*acn*	Fe	59.5	*acn*	Aconitate hydratase	60.1
GWH15_02935	*mdtB/yegN*	SDC, HCL	58.5	*mdtB/yegN*	Multidrug transporter subunit MdtB	58.6
GWH15_12080	*mdtB/yegN*	SDC, HCL	58.5	*mdtB*	Multidrug transporter subunit MdtB	59.2
GWH15_00470	*ruvB*	Cr, Te, Se	58.5	*ruvB*	DNA helicase RuvB	59.6
GWH15_11305	*mdtC/yegO*	SDC and other	56.5	*mdtC/yegO*	Multidrug transporter subunit MdtC	56.6
GWH15_13565	*pstB*	As	56.2	*pstB*	Phosphate ABC transporter ATP-binding protein PstB	56.6
GWH15_16395	*arsM*	As	54.4	*arsM*	Arsenite S-adenosylmethyltransferase	43.8

ORF (continued)	EXP	Compound	Identity (%)	PRE	Compound	Identity (%)
GWH15_02930	*mdtC/yegO*	SDC and other	52.9	*mdtC/yegO*	Multidrug transporter subunit MdtC	52.8
GWH15_12085	*mdtC/yegO*	SDC and other	52.5	*mdtC*	Multidrug transporter subunit MdtC	52.7
GWH15_18655	*sodA*	H_2O_2	51.3	*sodA*	Superoxide dismutase	51.3
GWH15_12800	*mdtC/yegO*	SDC and other	51.3	*mdtC*	Multidrug transporter subunit MdtC	51.2
GWH15_02405	*galE*	CTAB	50.9	*galE*	UDP-glucose 4-epimerase GalE	55.9
GWH15_16550	*pcoA*	Cu	49.8	*copA*	Cu resistance system multicopper oxidase	51.8
GWH15_04740	*arsT*	As	49.8	*arsT*	Thioredoxin-disulfide reductase	53.0
GWH15_06695	*acrR/ybaH*	Acriflavine	48.3	-	-	-
GWH15_05790	*cusA/ybdE*	Cu, Ag	48.3	*cusA/ybdE*	CusA/CzcA family heavy metal efflux RND transporter	48.3
GWH15_15485	*fabI*	Triclosan	48.2	*fabI*	Enoyl-ACP reductase FabI	50.2
GWH15_10040	*zraR/hydH*	Zn	47.7	*zraR/hydH*	Two-component system response regulator ZraR	49.2
GWH15_04095	*cpxR*	H_2O_2 and other	46.5	*cpxR*	DNA-binding response regulator	50.0
GWH15_04345	*dmeR*	Co, Ni	46.2	*dmeR*	Metal/formaldehyde-sensitive transcriptional repressor	46.2
GWH15_19170	-	-	-	*mntR*	Inner membrane protein YbiR	60.2
GWH15_14955	-	-	-	*mtrD*	Phospho-N-acetylmuramoyl-pentapeptide-transferase	58.4
GWH15_15295	-	-	-	*glpF*	Glycerol kinase	56.8
GWH15_05235	-	-	-	*merA*	Mercury(II) reductase	56.3
GWH15_15725	-	-	-	*ruvB*	DNA helicase RuvB	54.8
GWH15_14350	-	-	-	*aioE*	NAD(P)/FAD-dependent oxidoreductase	54.3
GWH15_16410	-	-	-	*arsC*	Arsenate reductase ArsC	54.0
GWH15_03545	-	-	-	*recG*	ATP-dependent DNA helicase RecG	52.8
GWH15_17515	-	-	-	*merA*	Mercury(II) reductase	52.6
GWH15_00860	-	-	-	*aioE*	NAD(P)/FAD-dependent oxidoreductase	52.3
GWH15_11710	-	-	-	*trgB*	Te resistance protein	51.6
GWH15_09530	-	-	-	*zupT/ygiE*	Zn transporter ZupT	51.1
GWH15_01540	-	-	-	*actA*	Apolipoprotein N-acyltransferase	50.0
GWH15_10285	-	-	-	*copR*	DNA-binding response regulator	50.0
GWH15_12250	-	-	-	*arsT*	Thioredoxin reductase	49.5
GWH15_16485	-	-	-	*emrBsm*	MFS transporter	48.8
GWH15_08215	-	-	-	*trgB*	Te resistance protein	48.6
GWH15_19280	-	-	-	*copG*	Uncharacterized conserved protein	48.6
GWH15_07930	-	-	-	*copC*	Cu resistance protein	48.6
GWH15_03770	-	-	-	*emhA*	Efflux RND transporter periplasmic adaptor subunit	48.6
GWH15_04310	-	-	-	*zupT/ygiE*	Zn transporter ZupT	48.6
GWH15_17565	-	-	-	*arsT*	thioredoxin-disulfide reductase	48.1
GWH15_00030	-	-	-	*modC*	ABC transporter ATP-binding protein	47.7
GWH15_07970	-	-	-	*cop-unnamed*	Cu oxidase	47.7
GWH15_10290	*copR*	Cu	44.4	*copR*	DNA-binding response regulator	47.6
GWH15_12715	-	-	-	*mdeA*	Methionine gamma-lyase	47.6
GWH15_16960	-	-	-	*fbpC*	ABC transporter ATP-binding protein	47.6
GWH15_07190	-	-	-	*gadC/xasA*	Glutamate:gamma-aminobutyrate antiporter	47.5
GWH15_00895	-	-	-	*arsT*	Thioredoxin-disulfide reductase	47.4
GWH15_08115	-	-	-	*chrA*	Chromate transporter	47.4
GWH15_13100	-	-	-	*vexD*	AcrB/AcrD/AcrF family protein	47.4
GWH15_16860	-	-	-	*copB*	Cu transporting ATPase	47.2

ORF (continued)	EXP	Compound	Identity (%)	PRE	Compound	Identity (%)
GWH15_18070	-	-	-	*copS*	Two-component sensor histidine kinase	47.2
GWH15_02325	-	-	-	*vmeG*	Efflux RND transporter periplasmic adaptor subunit	47.1
GWH15_19630	-	-	-	*arsA*	As pump-driving ATPase	47.1
GWH15_10660	*corR*	Cu	43.8	*corR*	Sigma-54-dependent Fis family transcriptional regulator	46.9
GWH15_11005	-	-	-	*abeS*	Multidrug transporter	46.9
GWH15_12005	-	-	-	*soxR*	Redox-sensitive transcriptional activator SoxR	46.9
GWH15_17680	-	-	-	*arsM*	Polymerase 2 ADP-ribosyltransferase 2	46.8
GWH15_11640	-	-	-	*abeS*	Multidrug transporter	46.7
GWH15_03945	*copR*	Cu	44.9	*copR*	DNA-binding response regulator	46.6
GWH15_05310	-	-	-	*nhlF*	NHLF_RHORH RecName: Co transport protein NhlF	46.5
GWH15_02985	-	-	-	*golT*	Cu-translocating P-type ATPase	46.3
GWH15_08255	-	-	-	*pgpA/ltpgpA*	Putative ATP-binding cassette protein subfamily C, member 1	45.9
GWH15_01935	-	-	-	*ctpV*	Metal cation transporter P-type ATPase CtpV	45.8
GWH15_03140	-	-	-	*furA*	Transcriptional repressor	45.7
GWH15_17160	*comR/ycfQ*	Cu	43.9	*comR/ycfQ*	TetR/AcrR family transcriptional regulator	45.6
GWH15_02610	-	-	-	*copB*	Cu-exporting ATPase	45.5
GWH15_13575	*pstC*	As	42.2	*pstC*	Phosphate transporter permease subunit PstC	45.3
GWH15_13260	-	-	-	*copS*	Two-component sensor histidine kinase	45.3
GWH15_14490	-	-	-	*perR*	Transcriptional repressor	45.2
GWH15_00410	-	-	-	*copA*	Putative cadmium-transporting P-type ATPase	45.2
GWH15_04685	-	-	-	*fabL/ygaA*	Enoyl-acyl-carrier-protein	45.2
GWH15_07310	-	-	-	*cutO*	Cu oxidase	45.2
GWH15_09900	-	-	-	*merA*	Mercury(II) reductase	45.2
GWH15_00360	-	-	-	*mntR*	Inner membrane protein YbiR	45.0

EXP: BacMet experimentally confimed genes database. PRE: BacMet predicted genes database; SDC - Sodium Deoxycholate; HCL- Hydrochloric Acid; CTAB- Cetrimonium bromide

Table S2: *Granulicella* strain 5B5 gene annotation with BacMet databases. Annotation performed against the experimentally confirmed and predicted resistance genes databases using BacMet Scan software. Only genes with hits ≥ 40% identity in one of the databases are shown.

ORF	EXP	Compound	Identity (%)	PRE	Compound	Identity (%)
G5B5_03005	*arsB*	As, Sb	63.7	*acr3*	Arsenical-resistance protein	70.5
G5B5_06155	*hoxN*	Ni	60.3	*hoxN*	HoxN/HupN/NixA family Ni/Co transporter	60.3
G5B5_10055	*ruvB*	Cr, Te, Se	59.6	*ruvB*	DNA helicase RuvB	59.0
G5B5_07635	*acn*	Fe	58.6	*acn*	Aconitate hydratase	59.3
G5B5_15620	*silP*	Ag	58.2	*actP*	Cu-transporting P-type ATPase	66.0
G5B5_12540	*pstB*	As	55.8	*pstB*	ABC transporter ATP-binding protein PstB	56.6
G5B5_06515	*mdtC/ yegO*	SDC	55.7	*mdtC/yegO*	Multidrug transporter subunit MdtC	55.8
G5B5_06520	*mdtB/ yegN*	SDC	55.0	*mdtB*	Multidrug transporter subunit MdtB	55.4
G5B5_16735	*merA*	Hg	52.8	*merA*	Mercury(II) reductase	36.4

ORF (continued)	EXP	Compound	Identity (%)	PRE	Compound	Identity (%)
					NAD(P)/FAD-dependent	
G5B5_10440	*aioE*	As	51.3	*aioE*	oxidoreductase	51.3
G5B5_07160	*sodA*	Se, H_2O_2	51.2	*sodA*	Superoxide dismutase	54.0
G5B5_13030	*galE*	CTAB	50.5	*galE*	UDP-glucose 4-epimerase GalE	52.3
					Metal sensitive transcriptional	
G5B5_01815	*dmeR*	Co, Ni	50.0	*dmeR*	repressor	52.0
G5B5_04025	*smdA*	DAPI	50.0	-	-	-
G5B5_06995	*copR*	Cu	50.0	*copR*	DNA-binding response regulator	52.7
G5B5_07595	*mtrD*	Triton X-100	50.0	*chrC*	hypothetical protein AYO46_00620	32.9
G5B5_08045	*arsB*	As, Sb	50.0	-	-	-
G5B5_11195	*pitA*	Zn, Te	50.0	-	-	-
G5B5_14705	*copR*	Co	49.3	*copR*	DNA-binding response regulator	51.1
G5B5_14745	*frnE*	Cd, H_2O_2	48.7	-	-	-
G5B5_01735	*pbrD*	Pb	48.6	-	-	-
G5B5_01970	*dmeF*	Co, Cd, Ni	48.5	-	-	-
G5B5_13805	*mntA/ ytgA*	Mn, Cd	48.5	-	-	-
		Sodium				
G5B5_04635	*iclR*	acetate	48.4	-	-	-
G5B5_04715	*mtrE*	Triton X-100	48.4	-	-	-
G5B5_08485	*copA*	Cu, Ag	48.4	-	-	-
G5B5_15860	*pgpA/ ltpgpA*	As, Sb	48.4	-	-	-
G5B5_15165	*sh-fabI*	Triclosan	48.0	sh-*fabI*	Enoyl-acyl-carrier-protein	50.8
	cusA/				CusA/CzcA family heavy metal efflux	
G5B5_16145	*ybdE*	Cu, Ag	47.9	*cusA/ybdE*	RND transporter	48.0
G5B5_01825	*copS*	Cu	47.5	-	-	-
G5B5_15680	*mtrD*	Triton X-100	47.5	-	-	-
G5B5_16090	*mexY*	EtBr	47.4	-	-	-
G5B5_12645	*soxS*	Zn	47.2	-	-	-
G5B5_12930	*arsT*	As	46.9	*arsT*	Thioredoxin-disulfide reductase	51.9
G5B5_06355	*modA*	W, Mo	46.9	-	-	-
G5B5_06715	*oprN*	Triclosan	46.9	-	-	-
		Sodium				
		Glycocholate			MexC family multidrug efflux RND	
G5B5_07405	*vmeC*	and others	46.9	*mexC*	transporter subunit	31.3
					Phosphate transporter permease	
G5B5_12535	*pstA*	As	46.7	*pstA*	subunit PtsA	47.2
G5B5_04020	*phoR*	BAC	46.7	-	-	-
G5B5_04280	*oqxA*	BAC	46.7	-	-	-
G5B5_08775	*ctpV*	Cu	46.5	-	-	-
G5B5_03175	*tbtB*	TBT	46.2	-	-	-
G5B5_00865	*ctpV*	Cu	45.9	-	-	-
G5B5_07315	*copA*	Cu, Ag	45.9	-	-	-
G5B5_16340	*ttgB*		45.9	-	-	-
G5B5_16800	*copB*	Cu	45.9	-	-	-
G5B5_00800	*arsH*	AS	45.7	-	-	-
G5B5_05205	*bepE*	SDC	45.7	-	-	-
					NAD(P)/FAD-dependent	
G5B5_08125	*aioE*	As	45.7	*aioE*	oxidoreductase	42.5
G5B5_14315	*czcS*	Cd, Zn, Co	45.7	-	-	-
G5B5_14920	*czrA*	Zn, Cd	45.7	-	-	-
G5B5_10760	*vmeB*	BAC	45.7	-	-	-
G5B5_11005	*cpxR*	H_2O_2	45.7	*cpxR*	DNA-binding response regulator	46.3
					Ni ABC transporter permease subunit	
G5B5_00945	*nikC*	Ni	45.6	*nikC*	NikC	46.1
G5B5_00440	*chrR*	Cr, Fe, H_2O_2	45.5	*yieF*	NAD(P)H-dependent oxidoreductase	30.5
G5B5_03750	*mtrR*	Triton X-100	45.5	-	-	-

ORF (continued)	EXP	Compound	Identity (%)	PRE	Compound	Identity (%)
	actP/	Sodium				
G5B5_05605	*yjcG*	Glycocholate	45.5	-	-	-
G5B5_05665	*fecE*	Ni, Co	45.5	-	-	-
G5B5_05745	*mexX*	EtBr	45.5	-	-	-
G5B5_06575	*fbpC*	Fe, Ga	45.5	-	-	-
G5B5_07095	*nixA*	Ni	45.5	*nrsS*	Hypothetical protein	30.4
G5B5_11940	*tbtB*	TBT	45.2	-	-	-
G5B5_09275	*cnrH*	Co, Ni	45.2	*copB*	Cu-translocating P-type ATPase	31.5
G5B5_09475	*smeE*	Triclosan	45.2	-	-	-
G5B5_12380	*wtpB*	W, Mo	45.2	-	-	-
G5B5_02995	-	-	-	*arsM*	Arsenite S-adenosylmethyltransferase	60.0
G5B5_05125	-	-	-	*mtrD*	Phospho-N-acetylmuramoyl-pentapeptide-transferase	57.6
G5B5_10215	-	-	-	*mntR*	Inner membrane protein YbiR	57.2
G5B5_09640	-	-	-	*glpF*	Glycerol kinase	56.9
G5B5_03300	*fbpC*	Fe, Ga	32.6	*ruvB*	DNA helicase RuvB	56.7
G5B5_04040	-	-	-	*wtpA*	W ABC transporter substrate-binding protein WtpA	53.3
G5B5_01300	*tbtA*	TBT	33.8	*zupT/ygiE*	Zn transporter ZupT	51.1
G5B5_16830	-	-	-	*cmtR*	ArsR family transcriptional regulator	50.0
G5B5_08815	*acrD*	Cu, Zn	32.8	*copA*	Cu-translocating P-type ATPase	50.0
G5B5_01765	*nczA*	Ni, Co, Zn	42.2	*modC*	ABC-type spermidine/putrescine transport system	50.0
G5B5_08510	*mreA*	Ni, Zn	43.8	*mreA*	Metal resistance protein	50.0
G5B5_16220	-	-	-	*zupT/ygiE*	Zn transporter ZupT	50.0
G5B5_00905	*hsmR gadW/*	EtBr	36.4	*arsT*	Thioredoxin-disulfide reductase	49.0
G5B5_11375	*yhiW*	HCl	42.4	*mtrA*	AraC family transcriptional regulator	48.8
G5B5_08230	*arsT*	As	43.8	*merA*	Mercury (II) reductase	48.6
G5B5_09395	-	-	-	*vceA*	Efflux transporter periplasmic adaptor subunit	48.4
G5B5_11185	-	-	-	*copS*	Sensor histidine kinase	47.7
G5B5_12325	*ybtP*	Fe	30.4	*ziaA*	Ca-translocating P-type ATPase	47.1
G5B5_05120	-	-	-	*aioE*	NAD(P)/FAD-dependent oxidoreductase	46.2
G5B5_04590	*irlR*	Cd, Zn	44.6	*irlR*	DNA-binding response regulator	46.2
G5B5_03145	*recG*	Cr, Te, Se	33.3	*mdtI/ydgE*	Multidrug/spermidine transporter subunit MdtI	45.9
G5B5_13015	*adeB*	Pyronin Y	40.0	*adeB*	Multidrug efflux RND transporter permease subunit	45.7
G5B5_12855	*bepE*	SDC	44.4	*bepE*	Multidrug efflux RND transporter permease subunit	45.7
G5B5_13845	*vmeV*	SDS	33.3	*trgB*	Hypothetical protein	45.6
G5B5_12165	*gadB*	HCl	36.1	*ctpV*	Metal cation transporter P-type ATPase CtpV	45.6
G5B5_04210	*comR/ ycfQ*	Cu	43.2	*comR/ycfQ*	TetR/AcrR family transcriptional regulator	45.5
G5B5_13340	-	-	-	*copD*	Hypothetical protein	45.5
G5B5_14950	-	-	-	*pbrT*	Fe permease	45.5
G5B5_15200	-	-	-	*pcm*	Protein-L-isoaspartate O-methyltransferase	45.5
G5B5_12960	*mexF*	Triclosan	31.7	*fecD*	Fe-dicitrate transporter subunit FecD	45.2

EXP: BacMet experimentally confimed genes database. PRE: BacMet predicted genes database; SDC - Sodium Deoxycholate; HCL- Hydrochloric Acid; CTAB- Cetrimonium bromide; DAPI- 4',6-diamidino-2-phenylindole; BAC- Benzylkonium Chloride; TBT-Tributyltin

Table S3: Significantly upregulated proteins in the proteomic profile of *Granulicella* sp. 5B5 with the addition of Mn. Only statistically significant differentially expressed proteins with a fold change ≥1.5 and proteins expressed only with the addition of Mn (on Mn) are shown.

ORF	Proteins	log_2 FC
G5B5_02010	D-hydantoinase	4.48
G5B5_04470	3'-5' exoribonuclease YhaM	3.91
G5B5_02385	Protein YciF	2.82
G5B5_13570	Blue-light-activated protein	2.65
G5B5_10245	Hypothetical protein	2.62
G5B5_09185	Hypothetical protein	2.52
G5B5_06075	Alpha-1,4-glucan:maltose-1-phosphate maltosyltransferase 1	2.48
G5B5_06575	Oxalate decarboxylase OxdD	2.33
G5B5_06165	Phosphoenolpyruvate carboxykinase [GTP]	2.29
G5B5_01965	Hypothetical protein	2.23
G5B5_12545	Phosphate-specific transport system accessory protein PhoU	2.17
G5B5_14775	Protein RecA	2.14
G5B5_08980	DNA-binding protein HU	2.13
G5B5_15595	Glucose-6-phosphate 1-dehydrogenase 1	2.09
G5B5_00055	ATP-dependent recd-like DNA helicase	2.04
G5B5_15800	Ribonucleoside-diphosphate reductase 1 subunit alpha	1.93
G5B5_15655	Hypothetical protein	1.92
G5B5_11545	DNA-directed RNA polymerase subunit alpha	1.91
G5B5_02760	Alpha-1,4-glucan:maltose-1-phosphate maltosyltransferase	1.90
G5B5_03225	Hypothetical protein	1.89
G5B5_03345	Hypothetical protein	1.86
G5B5_02590	Glutamate synthase [NADPH] large chain	1.82
G5B5_10000	GTP 3',8-cyclase	1.82
G5B5_11930	Hypothetical protein	1.80
G5B5_03160	Octaprenyl diphosphate synthase	1.78
G5B5_13250	Nucleoside diphosphate kinase	1.78
G5B5_02655	Uvrabc system protein B	1.75
G5B5_02610	Acetylornithine aminotransferase	1.75
G5B5_02365	UDP-N-acetylmuramate--L-alanyl-gamma-D-glutamyl-meso-2,6-diaminoheptandioate ligase	1.72
G5B5_03020	HTH-type transcriptional regulator Hpr	1.72
G5B5_03150	Putative metal-dependent hydrolase YcfH	1.72
G5B5_03725	Hypothetical protein	1.71
G5B5_08555	Hypothetical protein	1.67
G5B5_07785	Hypothetical protein	1.65
G5B5_09210	Small heat shock protein IbpB	1.65
G5B5_00640	Carbamoyl-phosphate synthase small chain	1.62
G5B5_11390	Imidazole glycerol phosphate synthase subunit HisF	1.62
G5B5_08025	Hypothetical protein	1.61
G5B5_10075	Adenine deaminase	1.60
G5B5_01745	Long-chain-fatty-acid--coa ligase Fadd15	1.58
G5B5_04740	Hypothetical protein	1.58
G5B5_08550	Regulator of RpoS	1.57
G5B5_02980	Glucose 1-dehydrogenase 1	1.56
G5B5_13190	Aminomethyltransferase	1.55
G5B5_04260	Hypothetical protein	1.51
G5B5_01715	Biotin carboxylase	1.50
G5B5_06710	Hypothetical protein	on Mn
G5B5_06495	D-threo-aldose 1-dehydrogenase	on Mn
G5B5_15040	Glutathione-independent formaldehyde dehydrogenase	on Mn
G5B5_04460	Hypothetical protein	on Mn
G5B5_04810	Hypothetical protein	on Mn
G5B5_05770	P-protein	on Mn
G5B5_07390	Nucleoid occlusion factor SlmA	on Mn
G5B5_14875	ABC transporter ATP-binding protein NatA	on Mn
G5B5_06295	Hypothetical protein	on Mn
G5B5_15170	Hypothetical protein	on Mn

ORF (continued)	Proteins	log_2 FC
G5B5_05650	Hypothetical protein	on Mn

Table S4: Significantly upregulated proteins in the proteomic profile of *Granulicella* sp. 5B5 with the addition of Mn with hits against BacMet gene databases. Annotation performed against the experimentally confirmed and predicted resistance genes databases using BacMet Scan software. Only genes with hits ≥ 40% identity in one of the databases are shown.

ORF	EXP	Compound	Identity (%)	PRE	Compound	Identity (%)
G5B5_06575	*fbpC*	Fe, Ga	45.5	-	-	-
G5B5_05650	*cueA*	Cu, Ag	43.5	*cueA*	Cu-translocating P-type ATPase	39.3
G5B5_00055	*chrA*	Cr	38.7	-	-	-
G5B5_01745	*fbpB*	Fe, Ga	38.3	-	-	-
G5B5_06295	*aioR/aoxR*	As	36.7	-	-	-
G5B5_03725	*pmrC*	Fe	36.4	*mtrA*	AraC family transcriptional regulator	32.7
G5B5_03225	*gadB*	HCl	35.9	*copA*	Cu-translocating P-type ATPase	30.9
G5B5_01965	*mexF*	Triclosan	34.7	-	-	-
G5B5_03160	*bepG*	SDC	34.3	-	-	-
G5B5_06075	*nczA*	Ni, Co, Zn	34.1	-	-	-
G5B5_08550	*actR*	Cd, Zn, HCl	33.8	*zraR/hydH*	Fis family transcriptional regulator	30.0
G5B5_06495	*tcrA*	Cu	33.3	*copA*	Cu-translocating P-type ATPase	39.6
G5B5_11930	*mexY*	EtBr	33.3	-	-	-
G5B5_07390	*bcrA*	BAC	33.0	*bcrA*	TetR/AcrR family transcriptional regulator	33.0
G5B5_13190	*adeT1*	BAC	32.6	-	-	-
G5B5_13570	*copR*	Cu	32.5	*corR*	Fis family transcriptional regulator	30.4
G5B5_09210	*ibpB*	H_2O_2	32.0	*ibpB*	Heat shock chaperone IbpB	36.7
G5B5_02980	*fabL/ygaA*	Triclosan	32.0	*fabL/ygaA*	Enoyl-acyl-carrier-protein	31.7
G5B5_02655	*copA*	Copper (Cu), Silver (Ag)	31.9	*yieF*	NAD(P)H-dependent oxidoreductase	32.1
G5B5_00640	*mtrD*	Triton X-100	31.4	-	-	-
G5B5_03020	*farR*	Palmitic acid	31.3	*farR*	Homoprotocatechuate degradation operon regulator HpaR	30.6
G5B5_15040	*mexK*	Triclosan	31.1	*arsM*	As S-adenosylmethyltransferase	33.3
G5B5_11390	*fabK*	Triclosan	30.8	*fabK*	Enoyl-acyl-carrier-protein	33.3
G5B5_04810	*copB*	Cu	30.2	-	-	-
G5B5_14875	*troB*	Zn, Mn, Fe	30.0	*troB*	Metal ABC transporter ATP-binding protein	31.2
G5B5_08555	-	-	-	*fabL/ygaA*	Enoyl-acyl-carrier-protein	36.8
G5B5_13250	-	-	-	*chrB*	Hypothetical protein	32.2
G5B5_12545	-	-	-	*pstA*	PhoU and BPD transp 1 and ABC tran domain containing protein	30.7

EXP: BacMet experimentally confimed genes database. PRE: BacMet predicted genes database; SDC - Sodium Deoxycholate; HCL- Hydrochloric Acid; BAC- Benzylkonium Chloride

Table S5: Significantly upregulated proteins in the proteomic profile of *Granulicella* sp. WH15 with the addition of Mn. Only statistically significant differentially expressed proteins with a fold change ≥1.5 are shown.

ORF	Proteins	log_2 FC
GWH15_02295	Hypothetical protein	2.92
GWH15_19690	Hypothetical protein	2.88
GWH15_18635	Putative sugar phosphate isomerase YwlF	1.92
GWH15_05130	Anti-sigma-B factor antagonist	1.91
GWH15_01045	dTDP-4-dehydrorhamnose 3,5-epimerase	1.90
GWH15_15570	ATP synthase subunit b	1.83
GWH15_13820	Hypothetical protein	1.78
GWH15_03260	3',5'-cyclic adenosine monophosphate phosphodiesterase CpdA	1.68
GWH15_00080	N5-carboxyaminoimidazole ribonucleotide mutase	1.67

ORF (continued)	Proteins	log_2 FC
GWH15_17845	Hypothetical protein	1.63
GWH15_02405	UDP-glucose 4-epimerase	1.62
GWH15_13225	ATP phosphoribosyltransferase	1.60
GWH15_19100	2-iminobutanoate/2-iminopropanoate deaminase	1.56
GWH15_13825	Sulfur carrier protein CysO	1.54
GWH15_02470	30S ribosomal protein S7	1.53
GWH15_10485	Hypothetical protein	1.50

Table S6: Significantly downregulated proteins in the proteomic profile of *Granulicella* sp. 5B5 with the addition of Mn. Only statistically significant differentially expressed proteins with a fold change ≤-1.5 and proteins expressed only in the absence of Mn (off Mn) are shown.

ORF	Proteins	log_2 FC
G5B5_06260	Transcriptional regulator YqjI	-2.08
G5B5_04865	Hypothetical protein	-1.16
G5B5_12975	Hypothetical protein	-1.13
G5B5_00540	Hypothetical protein	-1.10
G5B5_03300	ATP-dependent zinc metalloprotease FtsH	-1.08
G5B5_02930	(R)-stereoselective amidase	-0.98
G5B5_07395	Putative D-xylose utilization operon transcriptional repressor	-0.95
G5B5_12740	Glutamine--fructose-6-phosphate aminotransferase	-0.93
G5B5_06120	3-phenylpropionate-dihydrodiol/cinnamic acid-dihydrodiol dehydrogenase	-0.92
G5B5_00460	3-methyl-2-oxobutanoate hydroxymethyltransferase	-0.92
G5B5_02750	Ferredoxin-NADP reductase	-0.90
G5B5_12735	N-acetylmuramic acid/N-acetylglucosamine kinase	-0.88
G5B5_02110	Hypothetical protein	-0.85
G5B5_08120	Putative aminoacrylate peracid reductase RutC	-0.84
G5B5_02460	Hypothetical protein	-0.84
G5B5_09120	Ribosome-binding ATPase YchF	-0.83
G5B5_11445	Carbonic anhydrase 1	-0.83
G5B5_00615	Hypothetical protein	-0.82
G5B5_05430	Cysteine synthase B	-0.75
G5B5_13900	Gluconate 5-dehydrogenase	-0.75
G5B5_07595	Phospho-2-dehydro-3-deoxyheptonate aldolase	-0.74
G5B5_06545	Putative nicotinate-nucleotide pyrophosphorylase	-0.74
G5B5_09865	Beta-hexosaminidase	-0.74
G5B5_12210	Hypothetical protein	-0.73
G5B5_08645	Queuine trna-ribosyltransferase	-0.72
G5B5_04975	1-deoxy-D-xylulose 5-phosphate reductoisomerase	-0.70
G5B5_15980	Hypothetical protein	-0.70
G5B5_01030	Hypoxanthine phosphoribosyltransferase	-0.70
G5B5_12065	Hypothetical protein	-0.70
G5B5_04745	Hypothetical protein	-0.68
G5B5_15830	Hypothetical protein	-0.67
G5B5_09255	Putative glucose-6-phosphate 1-epimerase	-0.66
G5B5_02600	N-acetyl-gamma-glutamyl-phosphate reductase	-0.66
G5B5_02395	Hypothetical protein	-0.66
G5B5_09685	Hypothetical protein	-0.65
G5B5_15240	O-succinylbenzoate synthase	-0.65
G5B5_12965	2-amino-5-chloromuconic acid deaminase	-0.64
G5B5_11795	L-threonine dehydratase catabolic TdcB	-0.64
G5B5_02345	Hypothetical protein	-0.63
G5B5_15805	Ribonucleoside-diphosphate reductase subunit beta	-0.63
G5B5_13090	Aspartate carbamoyltransferase catalytic subunit	-0.62
G5B5_02400	S-methyl-5'-thioadenosine phosphorylase	-0.62
G5B5_13505	Quinone oxidoreductase 1	-0.60
G5B5_07290	Hypothetical protein	off Mn
G5B5_01820	Hypothetical protein	off Mn

ORF (continued)	Proteins	log_2 FC
G5B5_04790	Hypothetical protein	off Mn

Table S7: Significantly downregulated proteins in the proteomic profile of *Granulicella* sp. WH15 with the addition of Mn. Only statistically significant differentially expressed proteins with a fold change ≤-1.5 and proteins expressed only in the absence of Mn (off Mn) are shown.

ORF	Proteins	log_2 FC
GWH15_14260	Xyloglucanase	-1.85
GWH15_04015	Putative aliphatic sulfonates-binding protein	-1.73
GWH15_15380	N-acetyl-lysine deacetylase	-1.44
GWH15_10420	ATP-dependent Clp protease ATP-binding subunit ClpX	-1.44
GWH15_13385	Hypothetical protein	-1.36
GWH15_19375	Alanine racemase	-1.35
GWH15_11185	UvrABC system protein A	-1.31
GWH15_17800	Galactose-1-phosphate uridylyltransferase	-1.29
GWH15_15990	Hypothetical protein	-1.27
GWH15_15970	Multifunctional 2-oxoglutarate metabolism enzyme	-1.26
GWH15_04000	Hypothetical protein	-1.23
GWH15_12720	UvrABC system protein A	-1.19
GWH15_03100	Acyl carrier protein	-1.16
GWH15_14335	Alpha-ketoglutaric semialdehyde dehydrogenase	-1.14
GWH15_17270	Threonine synthase	-1.13
GWH15_06265	Glutamyl-tRNA(Gln) amidotransferase subunit A	-1.13
GWH15_14265	Xyloglucanase Xgh74a	-1.11
GWH15_11400	Beta-hexosaminidase	-1.08
GWH15_12565	Fumarate reductase flavoprotein subunit	-1.07
GWH15_02575	50S ribosomal protein L18	-1.06
GWH15_00875	Phosphoribosylformylglycinamidine synthase subunit PurL	-1.06
GWH15_12285	Protein TolB	-1.06
GWH15_17475	Hypothetical protein	-1.05
GWH15_17105	Prolyl tripeptidyl peptidase	-1.05
GWH15_03220	Phosphoribosylamine--glycine ligase	-1.03
GWH15_08475	Hypothetical protein	-1.03
GWH15_12210	GDP-L-fucose synthase	-1.01
GWH15_14365	Glycine--trna ligase alpha subunit	-0.99
GWH15_13275	Hypothetical protein	-0.99
GWH15_02875	2,3-dehydroadipyl-coa hydratase	-0.98
GWH15_13930	Trehalose synthase/amylase TreS	-0.97
GWH15_12385	Ribosomal protein S12 methylthiotransferase RimO	-0.97
GWH15_13180	Soluble aldose sugar dehydrogenase YliI	-0.97
GWH15_17185	Hypothetical protein	-0.96
GWH15_13830	Putative adenylyltransferase/sulfurtransferase MoeZ	-0.95
GWH15_00705	Outer membrane protein assembly factor BamA	-0.94
GWH15_01035	Dtdp-4-dehydrorhamnose reductase	-0.93
GWH15_02690	Inosine-5'-monophosphate dehydrogenase	-0.92
GWH15_18680	Hypothetical protein	-0.88
GWH15_08290	NADH-quinone oxidoreductase subunit 4	-0.88
GWH15_14415	P-protein	-0.88
GWH15_18875	Hypothetical protein	-0.87
GWH15_01065	Glutamyl-tRNA(Gln) amidotransferase subunit A	-0.86
GWH15_14475	Methanol dehydrogenase activator	-0.86
GWH15_19145	Cysteine desulfurase IscS	-0.85
GWH15_01945	Cyanophycinase	-0.85
GWH15_19215	Hypothetical protein	-0.84
GWH15_14360	Glycine--tRNA ligase beta subunit	-0.84
GWH15_18480	Putative formaldehyde dehydrogenase AdhA	-0.84
GWH15_18350	Thymidylate synthase ThyX	-0.83
GWH15_01270	Phosphoenolpyruvate carboxylase	-0.82
GWH15_00365	UDP-N-acetylglucosamine 4-epimerase	-0.81

ORF (continued)	Proteins	log_2 FC
GWH15_01110	CTP synthase	-0.79
GWH15_02270	3-oxoacyl-[acyl-carrier-protein] synthase 3	-0.78
GWH15_10500	Pyrroline-5-carboxylate reductase	-0.78
GWH15_17030	Putative glutamate--cysteine ligase 2	-0.77
GWH15_11850	Hypothetical protein	-0.77
GWH15_11780	Hypothetical protein	-0.74
GWH15_11475	Leucine--tRNA ligase	-0.73
GWH15_02195	Putative competence-damage inducible protein	-0.72
GWH15_07040	Putative oxidoreductase YdgJ	-0.71
GWH15_10160	3-oxoacyl-[acyl-carrier-protein] reductase FabG	-0.71
GWH15_01120	Hypothetical protein	-0.71
GWH15_08275	NADH-quinone oxidoreductase subunit F	-0.69
GWH15_14850	Farnesyl diphosphate synthase	-0.69
GWH15_14470	Hypothetical protein	-0.68
GWH15_00550	2Fe-2S ferredoxin	-0.68
GWH15_02350	3-isopropylmalate dehydratase large subunit	-0.68
GWH15_08050	Hypothetical protein	-0.68
GWH15_18880	Prolyl tripeptidyl peptidase	-0.68
GWH15_12655	Hypothetical protein	-0.68
GWH15_18355	3-oxoacyl-[acyl-carrier-protein] reductase FabG	-0.67
GWH15_07045	Outer membrane protein OprM	-0.67
GWH15_10555	Cytosol aminopeptidase	-0.67
GWH15_07840	Putative zinc protease	-0.67
GWH15_11610	Thiol:disulfide interchange protein DsbD	-0.67
GWH15_15540	Shikimate dehydrogenase (NADP(+))	-0.66
GWH15_01860	D-alanine--D-alanine ligase	-0.66
GWH15_19240	Biopolymer transport protein ExbD	-0.66
GWH15_17740	Alanine--tRNA ligase	-0.66
GWH15_11500	Sulfite reductase [NADPH] flavoprotein alpha-component	-0.66
GWH15_18340	Protein RecA	-0.65
GWH15_13090	Elongation factor Ts	-0.65
GWH15_05960	Hypothetical protein	-0.65
GWH15_00690	Chaperone SurA	-0.63
GWH15_07860	3-ketoacyl-coa thiolase	-0.63
GWH15_03065	Quercetin 2,3-dioxygenase	-0.63
GWH15_02430	6,7-dimethyl-8-ribityllumazine synthase	-0.61
GWH15_07165	Aminopeptidase YpdF	-0.61
GWH15_00040	Delta-aminolevulinic acid dehydratase	-0.61
GWH15_07145	Long-chain-fatty-acid--coa ligase Fadd15	-0.61
GWH15_06995	Pyruvate dehydrogenase E1 component subunit alpha	-0.61
GWH15_07935	Malate synthase A	-0.60
GWH15_06210	Putative FAD-linked oxidoreductase	off Mn
GWH15_09885	Putative pyridine nucleotide-disulfide oxidoreductase RclA	off Mn
GWH15_18545	Tyrosine--tRNA ligase	off Mn
GWH15_19630	Chromosome partitioning protein ParA	off Mn
GWH15_04710	ATP-dependent DNA helicase PcrA	off Mn
GWH15_15115	Acetyltransferase YpeA	off Mn
GWH15_05335	N5-carboxyaminoimidazole ribonucleotide mutase	off Mn
GWH15_00245	Putative metallo-hydrolase YflN	off Mn
GWH15_02910	Hypothetical protein	off Mn
GWH15_00345	Hypothetical protein	off Mn
GWH15_16500	Hypothetical protein	off Mn
GWH15_19325	Isoaspartyl peptidase	off Mn
GWH15_14120	Xaa-Pro dipeptidase	off Mn
GWH15_01485	Hypothetical protein	off Mn
GWH15_19600	Hypothetical protein	off Mn
GWH15_10235	Cyclase-dehydratase-3-O-methyl transferase TcmN	off Mn
GWH15_14935	Hypothetical protein	off Mn
GWH15_15315	Xyloglucanase Tgh74A	off Mn
GWH15_18775	Regulatory protein BlaR1	off Mn

Chapter 5

Identification of bacterial and fungal co-occurrence networks during assimilation of acidobacterial extracellular polymers in soil

Ohana Y. A. Costa, Agata Pijl, Eiko E. Kuramae

Submitted for publication

Abstract

Acidobacteria are one of the most abundant and ubiquitous bacterial phyla in soil but little information is available on their physiology and interactions with the other members of the soil microbial community. *Acidobacteria* produce copious amounts of extracellular polymeric substances (EPS) with unique sugar composition that could be used as a nutrient source by other microorganisms. Here, we investigated the impact of the amendment of EPS of *Granulicella* sp. strain WH15 (WH15EPS) to soil on assembly and interactions of the active bacterial and fungal communities. The addition of purified WH15EPS to the topsoil litter increased microbial activity compared with the unamended control, as measured by CO_2 respiration. Stable isotope probing (SIP) over a period of 35 days revealed that WH15EPS was assimilated by *Planctomycetes*, *Verrucomicrobia, Ascomycota* and *Basidiomycota.* Several taxa incorporated WH15EPS, including *Singulisphaera*, which was the most abundant genus in the labeled treatment. Co-inertia analysis suggested overall relationships between *Bacteria* and *Fungi*. Fungal groups were mainly connected positively to each other and negatively connected to bacterial groups, indicating competition for the carbon sources derived from the EPS. In addition, *Singulisphaera* had mostly positive connections, suggesting a potential cooperation in EPS metabolization. Our study revealed the potential interactions and structures of the co-occurrence network of active microorganisms able to metabolize WH15EPS differed from those of the control treatments, demonstrating that the analysis of co-ocurrence networks based only on total DNA may not reflect real co-occurrence among microorganisms. Furthermore, potential interactions that were not observed before, such as the connections between *Singulisphaera* and groups of *Planctomycetes* or *Acidobacteria,* can be unraveled by more specific and targeted metabolism studies.

Keywords: *Acidobacteria*, EPS, Stable Isotope Probing, *Planctomycetes*, Co-occurrence; Carbohydrates

1. Introduction

Acidobacteria are abundant bacterial phyla in soil, constituting 20–50% of the soil bacterial community (Kuramae *et al.*, 2012, Navarrete *et al.*, 2013, Pan *et al.*, 2014, Kielak *et al.*, 2016, Kielak *et al.*, 2016). However, little information is available on the physiology, ecological function, and impact of *Acidobacteria* on the soil environment (Kielak *et al.*, 2016) because their slow growth under standard laboratory conditions has resulted in a relatively small number of cultured representatives. Consequently, the factor(s) responsible for the prevalence and successful adaptation of *Acidobacteria* and their relationships with other soil-inhabiting microbes remain unknown (Kielak *et al.*, 2016).

Acidobacteria such as strains Ellin6076 and Ellin345 (Ward *et al.*, 2009), *Terriglobus tenax* (Whang *et al.*, 2014) and *Granulicella* sp. strains WH15 and 5B5 (Kielak *et al.*, 2017) produce extracellular polymeric substances (EPS). EPS are biopolymers secreted by a variety of microorganisms and mainly comprise carbohydrates, proteins and DNA. The role of EPS depends on the ecological niche and natural dwelling environment of the microorganism (Costa *et al.*, 2018). EPS production has been implicated in the long-term survival of *Acidobacteria* in several environments due to its protective properties (Kielak *et al.*, 2017). EPS also may serve as nutrient sources for other microorganisms (Flemming & Wingender, 2010), however the use of EPS from *Acidobacteria* as a nutrient has never been investigated in cross-feeding experiments.

The assimilation of EPS as a carbon source by active microorganisms can be investigated with stable isotope probing (SIP). SIP is a powerful technique that evaluates the incorporation of compounds labelled with heavy isotopes, such as ^{13}C, ^{18}O and ^{15}N, into the cell components of microorganisms metabolizing a specific substrate. Thus, SIP reflects the active microorganisms involved in the metabolism of a specific compound. SIP has been used to investigate the microorganisms responsible for the degradation of different compounds in many environments (Madsen, 2006), including labelled glucose (Verastegui *et al.*, 2014), methanol (Ginige *et al.*, 2004), phenol (Padmanabhan *et al.*, 2003), salicylate and anthracene (Singleton *et al.*, 2005). The use of SIP to investigate EPS has been limited to a single study identifying microorganisms that assimilate the EPS of *Beijerinckia indica* (Wang *et al.*, 2015). In this context, in the present study, EPS from *Granulicella* sp. strain WH15 was applied as a carbon source to soil sampled from the site where this strain was isolated, and the active bacterial and fungal community assemblages and interactions in the soil during a 35-day period were evaluated by SIP.

2. Material and Methods

2.1. Soil Sampling

Four topsoil-litter mixed samples were collected in the spring of 2017 from the Wolfheze forest in the Netherlands (Table 1). Samples were taken from topsoil (0 to 5 cm) adjacent to fallen tree trunks. The collected samples were pooled, sieved (4 mm mesh) and immediately used for SIP incubation with EPS from *Granulicella* sp. strain WH15 (WH15EPS). The physicochemical properties of the topsoil-litter samples were determined (Eurofins Agro BV, Wageningen, NL) and are presented in Table 2.

Table 1: Coordinates of the sampling sites.

Site number	Latitude	Longitude
1	51°59'14.5"N	5°47'32.7"E
2	51°59'15.9"N	5°47'29.5"E
3	51°59'15.7"N	5°47'27.7"E
4	51°59'14.8"N	5°47'23.2"E

Table 2: Physicochemical properties of topsoil-litter samples.

Component	Unit	Average (Sd)	Component	Unit	Average (Sd)
total N	mg N/kg	16535±3217	**CEC**	%	81±0.0
C/N ratio		20±6	**CEC**	mmol+/kg	214±42
Available N	kg N/ha	252±187	**B**	µg B/kg	488±4.2
pH		3.05±0.1	**Cu**	µg Cu/kg	48±9.9
OM	%	55.8±3.5	**Fe**	µg Fe/kg	3070±466.7
Na	mg Na/kg	34.5±4.9	**Mn**	µg Mn/kg	101920±9362.1
P	mg P/kg	42.65±6.4	**Zn**	µg Zn/kg	8860±693.0
K	mg K/k	218±11.3	**Clay**	%	5.5±0.7
Ca	kg Ca/ha	13±0	**Silt**	%	10±14.1
Mg	mg Mg/kg	175±7.1	**Sand**	%	27.5±12.0

Sd: standard deviation

2.2. [^{13}C]-Labeled and Unlabeled EPS Production

Granulicella sp. strain WH15 was grown on PSY5 solid medium (Campanharo *et al.*, 2016) containing 3% (wt/vol) fully ^{13}C-labeled glucose as the sole carbon source or unlabeled glucose for unlabeled control EPS production. The plates were incubated at 30 °C for 3 days and then at 20 °C for 27 days. The polysaccharide portion of EPS was extracted and purified according to the method described by Liu and Fang (2002), with modifications. EPS and cells of strain WH15 (~5 ml) were scraped from the plates into 50-ml Falcon tubes, and the volume was adjusted to 10 ml with sterile water. Sixty microliters of 36.5% formaldehyde was added to each sample and incubated at 4 °C for 1 h. Next, 4 ml of 1 M NaOH was added and incubated at 4 °C for 3 h. The cells were then pelleted by centrifugation at 9,000 xg for 40 min. The supernatants were filtered (0.2 µm membranes, Millipore) at room temperature, and monosaccharides were removed by dialysis in SnakeSkin™ Dialysis Tubing (3500 Da) (Thermo Fisher Scientific, Massachusetts, USA) against demineralized water at 4 °C for 48 h. The solutions were frozen at -80 °C for 12 h and then freeze-dried at -80 °C for 72 h.

Before freeze-drying, the DNA concentration in the EPS solution was determined in a Qubit fluorometer using a broad-range Quant-iT™ dsDNA Assay Kit (Invitrogen, Carlsbad, California, USA). EPS protein concentrations were determined by a Pierce™ Modified Lowry Protein Assay Kit (Thermo Fisher Scientific, Massachusetts, USA). The total carbohydrate content was estimated by the phenol-sulfuric acid method (DuBois *et al.*, 1956) modified for 96-well plates (Masuko *et al.*, 2005) with glucose as the standard. After purification, the EPS contained ~400 mg/ml carbohydrates, ~1% protein and undetectable amounts of DNA. The monosaccharide composition of WH15EPS was previously characterized (Kielak *et al.*, 2017) and comprises xylose (41.8%), mannose (10.25%), glucose (13.55%), galactose (26.12%), rhamnose (0.065%), glucuronic acid (8.09) and galacturonic acid (0.085%).

2.3. Stable Isotope Probing (SIP) Incubation

One milliliter of Milli-Q sterile water was added to the freeze-dried EPS immediately before inoculation in soil to facilitate a homogeneous distribution within the soil. Five grams (wet weight) of soil with 0.05% (wt/wt) WH15EPS (labeled and unlabeled controls) or without EPS was added to a 120-ml bottle, which was sealed with a butyl rubber stopper and incubated at room temperature (22 °C) in the dark. Each treatment (labeled EPS, unlabeled EPS and control without EPS) had six replicates. All vials were uncapped and aired every 4 days to maintain oxic conditions and prevent $^{13}CO_2$ cross-feeding. The vial headspace CO_2 was monitored daily via gas chromatography (GC) (Trace GC Ultra, Thermo Fisher Scientific, Massachusetts, USA). For incubations with [^{13}C]-labeled EPS, the headspace CO_2 $^{13}C/^{12}C$ ratio was monitored via GC combustion isotope ratio mass spectrometry (GC/C/IRMS) (GC IsoLink II™ IRMS System, Thermo Fisher Scientific, Massachusetts, USA). For DNA extraction, 0.5 g were sampled from the vials on days 8, 24 and 35, which corresponded to 10, 25 and 43% $^{13}CO_2$ headspace enrichment, respectively. The differences in CO_2 emissions on the different days were analyzed by analysis of variance (ANOVA one-way repeated measurements) using mixed-effects models ('lmerTest' (Kuznetsova *et al.*, 2017) and 'psycho' (Makowski, 2018) packages in R).

2.4. DNA Extraction and Fractionation

DNA was extracted from 250 mg of soil with or without ^{13}C-labeled/unlabeled substrates with the PowerSoil® DNA Isolation Kit (MO BIO Laboratories, Inc) according to the manufacturer's instructions. DNA was quantified by a spectrophotometer (NanoDrop™ 2000, Thermo Fisher Scientific, Massachusetts, USA) and visualized by 1.0% agarose gel electrophoresis and ethidium bromide staining. For gradient fractionation, 2 µg of DNA was combined with CsCl (1.72 g/ml) and gradient buffer (100 mM Tris-HCl pH 8.0, 100 mM KCl, 1mM EDTA) in an ultracentrifugation tube (PA UltraCrimp 1.8 ml, ThermoFisher Scientific, Massachusetts, USA) (Neufeld *et al.*, 2007) and ultracentrifuged at 125,395 xg (Discovery 120SE ultracentrifuge, ThermoFisher Scientific, Massachusetts, USA) under vacuum at 20 °C for 65 h. Gradient

fractionation resulted in 18 DNA fractions of approximately 100 µl each. The density of each fraction was measured with a refractometer (AR200, Reichert Technologies, New York, USA) to confirm gradient formation. DNA was precipitated from the CsCl with polyethylene glycol solution (30% PEG6000, 1.6 M NaCl) and glycogen (20 µg/µl), washed with 70% ethanol, and eluted in 30 µl of 10 mM Tris-HCl buffer, pH 8.0. The DNA concentration of each fraction was determined in a Qubit 4 Fluorometer (ThermoFisher Scientific, Massachusetts, USA) using a Quant-iT™ dsDNA HS Assay Kit (Invitrogen, Carlsbad, California, USA). The unlabeled substrate incubations were used as controls to determine the expected position of labeled soil DNA in the CsCl density gradients.

2.5. PCR, Sequencing and Sequence Processing of 16S rRNA Gene and ITS Data

Amplicon library preparation and high-throughput sequencing were performed using the "heavy" DNA fractions pooled within each sample replicate as well as the total DNA of both the amended and unamended controls. The V3-V4 region of the 16S rRNA gene and the internal transcribed spacer 2 (ITS1) region were targeted for bacteria and fungi, respectively. For bacteria, the V3-V4 region of 16S rRNA gene was targeted by using 515F (5'-GTGCCAGCMGCCGCGGTAA-3') as forward primer and 806R (5'-GGACTACHVGGGTWTCTAAT-3') (Bergmann *et al.*, 2011) as reverse primer. For fungi, the Internal Transcribed Spacer 1 (ITS1) region was targeted by using primers ITS1F (5′-CTTGGTCATTTAGAGGAAGTAA-3′) and ITS2 (5′-GCTGCGTTCTTCATCGATGC-3′) (White *et al.*, 1990). The amplification steps and Illumina MiSeq PE250 sequencing were performed at McGill University and Génome Québec Innovation Centre (Montréal, Québec, Canada). The sequences were deposited in the European Nucleotide Archive (ENA; https://www.ebi.ac.uk/ena) under the accession number PRJEB29719.

2.6. Processing and statistical analyses of 16S rRNA gene and ITS data

Raw sequencing data were processed through Hydra pipeline version 1.3.3 (Hollander, 2017) implemented in Snakemake (Koster & Rahmann, 2012). Reads were first quality filtered, by trimming adapters sequences and removing PhiX contaminants, using BBDuk2 from the BBMap tool suite (Bushnell, 2015). Paired-ends were merged using the fastq_mergepairs option from vsearch (Rognes *et al.*, 2015). The ITS1 region was extracted from ITS sequence dataset using ITSx version 1.011 (Bengtsson-Palme *et al.*, 2013). Next, vsearch was employed to cluster all reads into OTUs through the UPARSE strategy by dereplication, sorting by abundance with at least two sequences and clustering using the UCLUST smallmem algorithm (Edgar, 2010). Chimeriq sequences were detected and removed using the UCHIME algorithm in de-novo mode (Edgar *et al.*, 2011) implemented in VSEARCH. Before dereplication step, all reads were mapped to OTUs with the usearch_global method implemented in VSEARCH to generate an OTU table and converted to BIOM-Format (McDonald *et al.*, 2012). For 16S rRNA sequences, the taxonomic information for each OTU was then added to the BIOM file by aligning the sequences to the SILVA database (release 128) (McDonald *et al.*, 2012) using SINA

classifier (Pruesse *et al.*, 2012). For ITS sequences, the taxonomic information was added to the BIOM file by running the RDP Classifier re-trained on the UNITE database release 7.2 (Kõljalg *et al.*, 2013). Fungal OTUs were assigned to functional guilds using the annotation tool FUNGuild (Nguyen *et al.*, 2016).

2.7. Multivariate Analyses of 16S rRNA Gene and ITS Data

Statistical analyses were performed in RStudio version 1.1.423 running R version 3.4.4 (R Core Team, 2015). OTUs with less than 2 counts across all samples and chloroplast and mitochondrial sequences were discarded. Prior to alpha and beta diversity analyses, the data were rarefied to the size of the smallest sample (4,243 reads for the 16S rRNA gene and 7,797 for ITS region data). The 'Phyloseq' package (McMurdie & Holmes, 2013) was used to calculate the number of observed OTUs, Shannon and Inverse Simpson diversity indices, and Chao1 and ACE diversity estimators. Significant differences in the estimators between treatments were evaluated through parametric and non-parametric tests, including ANOVA, Kruskal-Wallis and Tukey's HSD tests (package 'agricolae' (Mendiburu, 2017)). Bray-Curtis distance matrices constructed using the rarefied datasets were used for principal coordinate analysis (PCoA) using the *capscale* function from the 'vegan' package (v. 2.4.6 (Oksanen *et al.*, 2018)). Group dissimilarities were tested by permutational multivariate analysis of variance (PERMANOVA) using the function *Adonis* from the 'vegan' package. To compare the structures of the bacterial and fungal communities within treatments, co-inertia (COIA) analysis of the Hellinger transformed datasets (Legendre & Gallagher, 2001) was performed using the function *coinertia* of the package 'ade4' (Dray & Dufour, 2007) as described by Schlemper et al. (Schlemper *et al.*, 2017) . Potential interactions between bacterial and fungal genera within treatments were investigated via SparCC correlation coefficients, with default parameters and 100 bootstraps (Friedman & Alm, 2012). Networks were built with significant interactions ($p<0.05$) at $|r| > 0.75$, and co-occurrence was visualized by Cytoscape (Shannon, 2003) and Gephi version 0.9.2 (https://gephi.org/). Topological parameters (numbers of nodes and edges and node degree number) were calculated using the network analyzer tool (Doncheva *et al.*, 2012) pre-installed in Cytoscape. Other plots were generated using 'ggplot2' (Wickham, 2016) and the *plot* function in R.

3. Results

3.1. CO_2 Measurements

The CO_2 emissions of the microcosms were 7,500 ppm CO_2/day initially and progressively decreased during incubation. Average CO_2 emissions were 1.3 times higher in the EPS-amended samples than in the unamended controls. At the end of the incubation, the EPS-amended microcosms produced 1.9 times more CO_2 than the unamended controls.

Headspace $^{13}CO_2$ enrichment increased throughout the incubation period, reaching 43% of the headspace in the labeled samples at day 35 (Figure 1). The differences in CO_2 emissions between treatments and p-values are shown in Table 3.

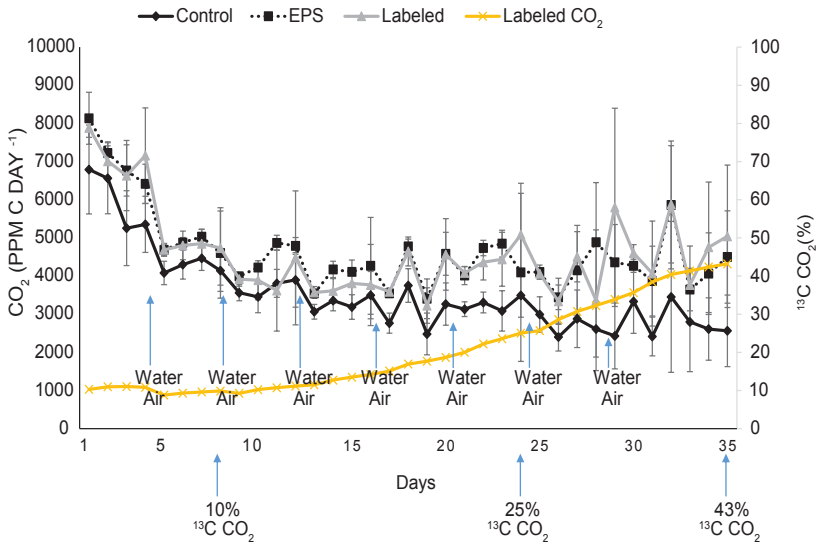

Figure 1: CO_2 emission. CO_2 production during total incubation period. Control: control without EPS; EPS: control containing ^{12}C-EPS; Labeled: incubation with ^{13}C-EPS; Labeled CO_2 percentage: $^{13}CO_2$ emitted during ^{13}C-EPS sample incubation; water: days when samples were hydrated; air: days when samples were aired.

5

Table 3: Contrast between CO_2 emission curves and statistical p-values.

Contrast	Difference	p-value
Unlabeled EPS -Control	1094.94	0.012356
Labeled EPS-Control	1058.11	0.01646
Unlabeled EPS – Labeled EPS	-36.83	0.994963

3.2. Gradient Fractionation

The density profiles of the DNA extracted from soils incubated with ^{13}C-EPS exhibited small changes compared with those of the control DNA from the ^{12}C-EPS incubations (Figure S1). “Heavy” fractions were chosen at densities at which little total DNA was detected in the control sample fractions (Figure S1). Among the 6 replicates, two from day 8 (L2D8 and L6D8) had “heavy” DNA in higher-density fractions, probably due to variations in the CsCl density centrifugation.

3.3. Sequencing and Alpha Diversity

High-throughput sequencing generated 1,250,548 and 828,045 good-quality sequences for

the bacterial 16S rRNA gene and fungal ITS region, respectively. Good's coverage (Table 4) indicated that the number of sequences reads covered 96.2%-98.4% of the bacterial diversity and 99.1-99.9% of the fungal diversity in the samples. Overall, for the 16S rRNA samples, the "heavy" fraction samples at all time points had significantly lower richness and diversity values than the unamended and ^{12}C-EPS-amended controls. Moreover, the ^{12}C-EPS-amended samples had lower richness and diversity values than the unamended samples (Figure S2). Similar to the 16S rRNA gene samples, the richness and diversity values of the ITS region were lower in the "heavy" fraction samples compared with the controls. This tendency was most obvious at days 24 and 35, whereas at day 8, the richness of the "heavy" fraction was not significantly different from that of the controls (Figure S3).

Table 4: Good's coverage for 16S rRNA gene and ITS region sequences within treatments.

16S rRNA gene sequences			**ITS region sequences**		
Treatment	**Day**	**Good`s coverage**	**Treatment**	**Day**	**Good`s coverage**
Control	8	96.2	Control	8	99.3
Unlabeled EPS	8	96.4	Unlabeled EPS	8	99.2
"Heavy" fraction	8	97.4	"Heavy" fraction	8	99.6
Control	24	96.3	Control	24	99.2
Unlabeled EPS	24	96.9	Unlabeled EPS	24	99.3
"Heavy" fraction	24	98.4	"Heavy" fraction	24	99.9
Control	35	96.3	Control	35	99.0
Unlabeled EPS	35	97.0	Unlabeled EPS	35	99.4
"Heavy" fraction	35	98.3	"Heavy" fraction	35	99.8

3.4. Community Structures

PERMANOVA (p=1.00e-04) showed that the bacterial communities differed significantly between the different treatments, days and treatment:day interactions, with the "heavy" fractions on days 24 and 35 clustering separately from the controls and the day 8 "heavy" fraction in the PCoA plot (Figure 2a). The bacterial communities clearly clustered according to treatments and days, with the first two axes of the PCoA explaining 70.5% of the variation. The fungal community patterns at different time points were less clear than those of the bacterial community. The fungal communities were more spread throughout the PCoA plot (Figure 2b). However, the heavy fractions on days 24 and 35 also clustered separately from the control samples and the day 8 "heavy" fraction. PERMANOVA (p value = 9.999^{e-05}) indicated significant differences between treatments, days and treatment:day interactions, with the first two axes of the PCoA explaining 27.8% of the variation among samples.

Co-inertia analysis revealed significant co-structures between the bacterial and fungal communities for all treatments (p<0.5). The first 2 co-inertia axes explained 80.4% ("heavy" fraction), 61.8% (EPS amended control) and 59.9% (unamended control) (cumulative projected inertia) of the total variance in the bacterial-fungal assessments (Figure 2cde). In addition, the bacterial and fungal groups at the genus level that contributed most to the co-

variance between samples were identified, which reflected the differences in abundances among time points (Figure S4).

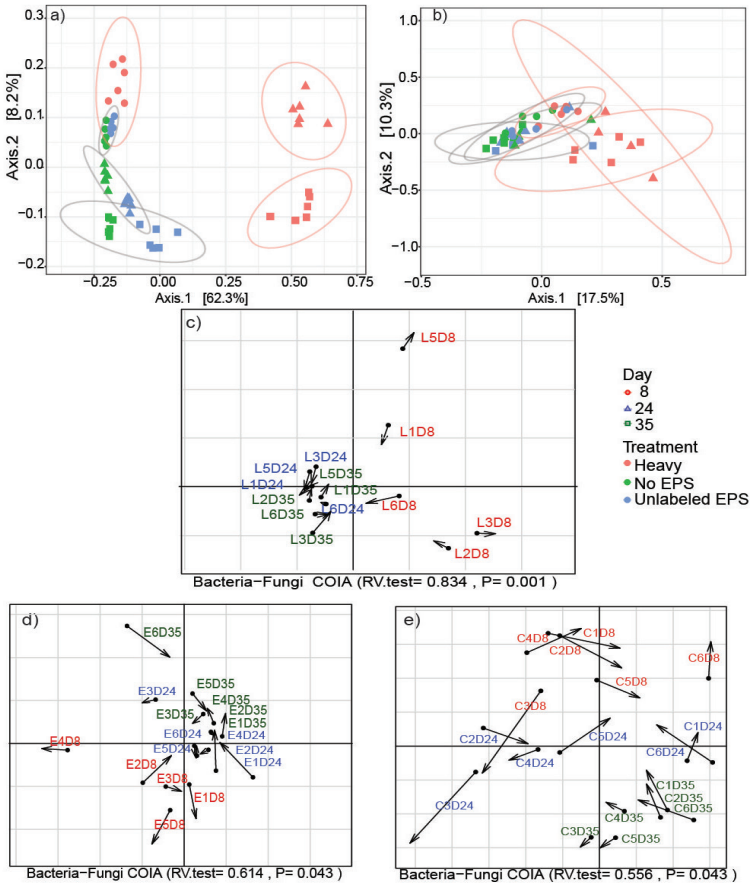

Figure 2: Co-inertia (COIA) and Principal Coordinate Analysis (PCoA) of microbial communities based on Bray-Curtis distances **a)** Bacterial community PCoA clustering of normalized and Hellinger-transformed rRNA gene 16S data **b)** Fungal community PCoA clustering of normalized and Hellinger-transformed ITS 1 region data. Co-inertia (COIA) analysis between PCoAs of **a)** bacterial and **b)** fungal community composition of samples amended with **c)** ^{13}C EPS, **d)** ^{12}C EPS and **e)** unamended samples. Samples from day 8 are colored in red, samples from day 24 are colored in blue and samples from day 35 are colored in in green.

3.5. Bacterial Community Beta Diversity

Twenty-two groups at the phylum level were observed in all samples. At day 8, the most abundant phylum in all samples was *Proteobacteria* (32.01%-39.7%) (Figure S5a). At day 24, the separation between microbial communities observed in the PCoA analysis became evident. The most abundant phylum was *Planctomycetes* in the "heavy" fraction (59.79%) but *Proteobacteria* in both control treatments (28.56%-31.23%) (Figure S5a). At day 35, *Planctomycetes* (60.79%) was still the most abundant phylum in the "heavy" fraction samples (Figure S5a), and *Proteobacteria* (25.65%-25.74%) remained the most abundant phylum in

the controls.

At the genus level, 220 groups were found in all samples. At day 8, unclassified *Pedosphaeraceae* was the most abundant group in the "heavy" fraction and the unamended control (12.34% and 13.04%, respectively). In the ^{12}C-EPS-amended treatment, the most abundant genus was *Rhodanobacter* (12.12%) (Figure 3a).

At day 24, Singulisphaera was the most abundant genus in the "heavy" fraction samples (27.79%) but represented only approximately 1.07-1.28% of the sequences in the control treatments. In both control treatments, unclassified Pedosphaeraceae (10.91%-14.36%) was the predominant group (Figure 3a).

At day 35, uncultured *Pirellulaceae* (36.21%) was the most abundant group in the "heavy"

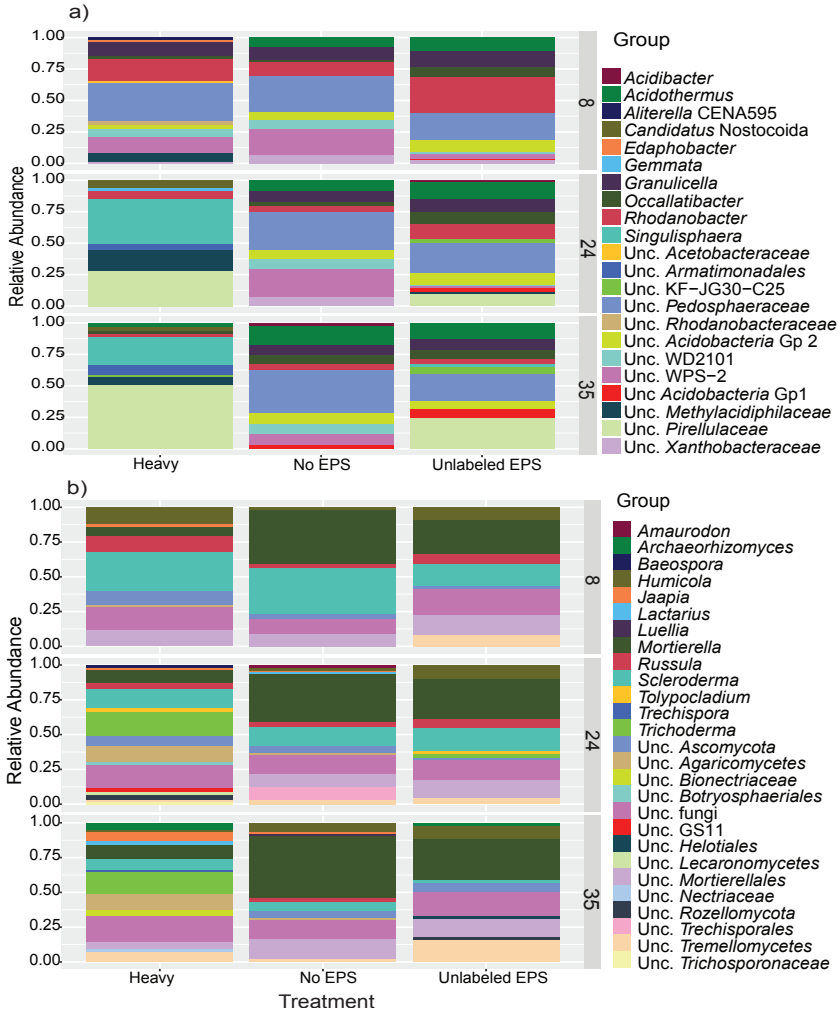

Figure 3: Relative abundance of microbial genera. Relative abundance of **a)** bacterial and **b)** fungal groups at genus level in 'heavy' fraction, EPS amended (Unlabeled EPS) and unamended (No EPS) controls at all time points (days 8, 24, 35). Only genera with >3% abundance are shown.

fraction, followed by *Singulisphaera* (16.22%). In the unamended control, the most abundant group was unclassified *Pedosphaeraceae* (16.52%), while in the amended control, the predominant group was uncultured *Pirellulaceae* (12.14%) (Figure 3a).

3.6. Fungal Community Beta Diversity

At the phylum level, 11 groups were observed in all samples. At days 8 and 24, *Basidiomycota* was the most abundant phylum in all samples (35.92%-43.16%) (Figure S5b). At day 35, the predominant group was *Basidiomycota* in the "heavy" fraction samples (38.7%) but *Mortierellomycota* in the unamended control (33.36%). In the ^{12}C-EPS amended control samples, *Ascomycota* (33.87%) was the most abundant phylum (Figure S5b).

At the genus level, 282 groups were observed in all samples. At day 8, the most abundant genus was *Scleroderma* in the "heavy" fractions (18.02%) but *Mortierella* in both control treatments (22.39%) (Figure 3b). At day 24, *Trichoderma* (15.67%) was the most abundant genus in the "heavy" fraction samples. In both control treatments, *Mortierella* was the most abundant genus (15.31%-21%) (Figure 3b). At day 35, unclassified fungi in the "heavy" fraction samples increased and were the most abundant group (15.21%), followed by *Trichoderma* (12.69%). *Mortierella* remained the most abundant genus in both control treatments (15.63%-24.96%) (Figure 3b). Analysis with FUNGuild assigned guilds to approximately 60% of the OTUs. The guild assignments demonstrated that all treatments, regardless of time point, were dominated by ectomycorrhizal fungi, soil saprotrophs and undefined saprotrophs (Figure S6).

3.7. Co-occurrence Network Analyses

3.7.1. Heavy fraction

The co-occurrence network of "heavy" fraction samples (Figure 4a) incorporated 59 nodes and 377 edges, with 51 bacterial nodes and 8 fungal nodes. At the genus level, 2 bacterial groups had connections with fungi, and 2 fungal groups were connected to bacteria. In total, 139 of the 377 connections were negative. The network contained several densely connected nodes, with 15 nodes having more than 20 neighbors each. The group with the highest number of connections was a *Planctomycetes* group, unclassified *Gemmataceae* (31 neighbors). Other groups with high connectivity were uncultured *Solirubrobacteraceae* (30 neighbors), unclassified *Armatimonadales* and *Acidothermus* (29 neighbors each). We also analyzed the bacterial genus identified as the most abundant in the "heavy" fractions, *Singulisphaera*, as well as the WH15EPS producer genus, *Granulicella*. In this treatment, *Singulisphaera* and *Granulicella* had 23 and 21 neighbors, respectively. Both genera had mainly positive connections (20 for *Singulisphaera* and 12 for *Granulicella*). No fungal group was connected to *Singulisphaera,* which was mainly connected to *Planctomycetes* and *Actinobacteria* (Figure 4d). The groups associated with *Singulisphaera* and *Granulicella* are described in Table S1. No strong associations were detected between *Granulicella* and

Singulisphaera. The most abundant fungal genus, *Trichoderma*, did not have any strong correlation.

3.7.2. ^{12}C-EPS-amended control

The network for the ^{12}C-EPS-amended control (Figure 4b) contained 29 nodes and 46 edges, with 18 bacterial nodes and 11 fungal nodes. Two bacterial groups interacted with fungi, and 3 fungal groups interacted with bacteria. No group had more than 10 neighbors, the

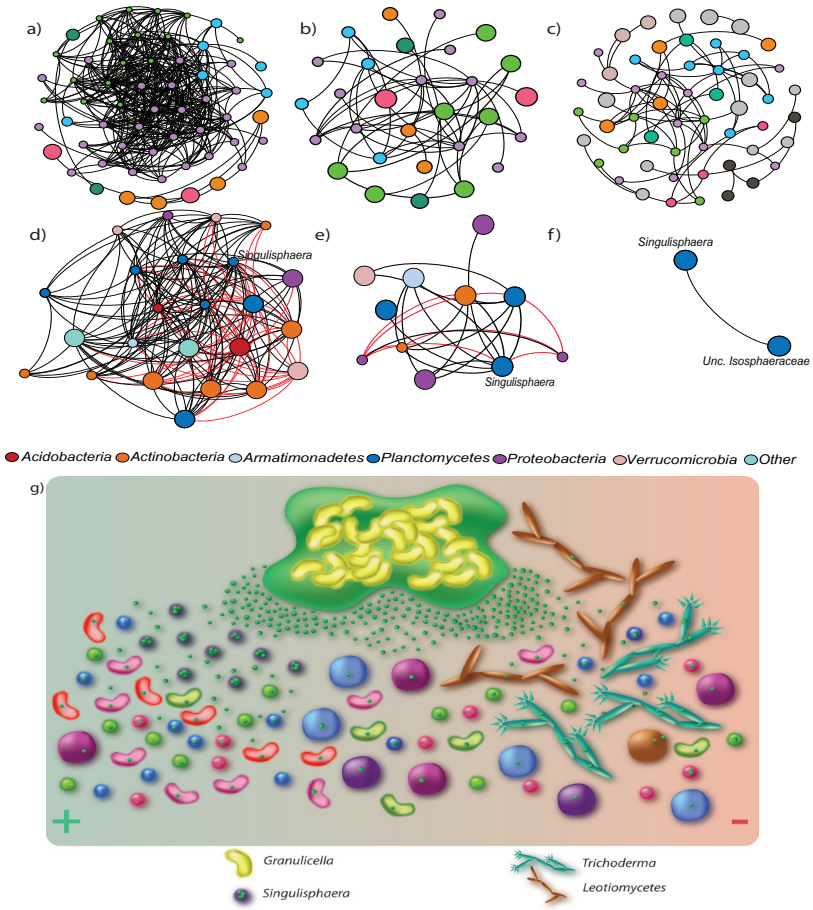

Figure 4: Co-occurrence networks and conceptual framework of the interactions of microorganism during WH15EPS assimilation. Total networks for **a)** "heavy" fraction; **b)** Unlabeled EPS control; **c)** NO EPS control (colors represent modularity classes); First neighbors of *Singulisphaera* genus in each treatment: **d)** "heavy" fraction; **e)** Unlabeled EPS control; **f)** NO EPS control (network nodes are genera). The size of the nodes represents the number of neighbors. Black edges represent positive correlations; red edges represent negative correlations. The size of each node is proportional to the number of connections (degree); **g)** Cartoon illustrating simplified hypothetical relationships among microorganisms during the metabolization of the EPS produced by *Granulicella* sp. strain WH15. *Singulisphaera* genus possessed negative (right (-)) correlations, as well as positive correlations (left (+)) with other bacterial genera. Negative associations suggest competition for the substrate, while positive associations suggest mutualism, via exchange of nutrients.

maximum number of neighbors of a node in this network. *Singulisphaera* and uncultured *Pirellulaceae* had the highest number of neighbors (10), followed by *Rhodanobacter*, Unclassified KF-JG30-C25 and uncultured *Solirubrobacteraceae*, with 6 neighbors each. In this treatment, *Granulicella* did not have any strong connections. *Singulisphaera* had only 2 negative connections, with *Dyella* and *Rhodanobacter* (Figure 4e). The associations of *Singulisphaera* are described in Table S2.

3.7.3. Unamended control

The network for the unamended control (Figure 4c) possessed 54 nodes and 66 edges, with 41 bacterial and 13 fungal nodes. Among all associations, 2 bacterial groups interacted with fungi, and 3 fungal groups interacted with bacteria. No groups had more than 10 neighbors. Genus *Burkholderia-Caballeronia-Paraburkholderia* had the highest amount of neighbours (10), followed by *Conexibacter* (7 neighbours), *Dyella* and uncultured *Acidobacteriaceae* subgroup 1 (6 neighbors each). *Singulisphaera* had only one positive connection to unclassified *Isosphaeraceae* (Figure 4f). *Granulicella* did not have any strong association.

4. Discussion

Here, we investigated the assimilation of EPS produced by *Granulicella* sp. strain WH15, a member of *Acidobacteria*. The ^{13}C-labeled biopolymer was applied to topsoil-litter samples, and high-throughput sequencing of the bacterial 16S rRNA gene and fungal ITS1 region identified the main EPS metabolizers. To increase the probability of finding true interactions, the mixed topsoil-litter was collected from the same forest site where *Granulicella* sp. strain WH15 was isolated (Valášková *et al.*, 2009).

Although the degradation of bacterial-produced biopolymers has been explored previously (cellulose produced by the bacterium *Gluconacetobacter xylinus* (Verastegui *et al.*, 2014, Wang *et al.*, 2015), indican from *Beijerinckia indica* (Wang *et al.*, 2015), and fructan from *Lactobacillus reuteri* (van Bueren *et al.*, 2015)), the metabolism of EPS produced by *Acidobacteria* and its ecological implications have not been investigated. The addition of purified WH15EPS to the topsoil-litter increased microbial activity compared with the unamended control, as measured by CO_2 respiration. The incorporation of the labeled EPS was confirmed by the increase in $^{13}CO_2$ release during incubation, which ranged from 10% to 43% of total headspace CO_2. The amount of $^{13}CO_2$ emitted varies according to the complexity of the substrate used for incubation and the capacity of the microbial community to degrade it (2012). Although the labeled material was clearly incorporated, the amount of “heavy” DNA recovered was lower than that reported in other SIP studies (Zhang *et al.*, 2016), likely due to the complexity of the substrate, a heteropolysaccharide composed of 7 different monosaccharides (Kielak *et al.*, 2017), and competition with the carbon present in the litter material (approximately 50% of

organic matter). Longer incubation times can improve the recovery of "heavy" DNA; however, a limitation of this strategy is that it also increases the possibility of cross-feeder enrichment (Verastegui *et al.*, 2014).

Alpha and beta diversity

The alpha diversity indices of the bacterial communities in the "heavy" fractions indicated lower richness and diversity than the amended and unamended controls at all sampling points, reflecting the selection of microorganisms capable of metabolizing the added EPS, particularly at day 35. These dynamics, however, were not as consistent for the fungal communities.

At day 8, the genus *Rhodanobacter* predominated in the EPS-amended control; however, the group unclassified *Pedosphaeraceae* was the most abundant in the "heavy" fraction and in the unamended control. *Rhodanobacter* is a genus of the family *Xanthomonadaceae* (*Proteobacteria*) found in soils featuring decomposition of aromatic compounds (Nalin *et al.*, 1999, Uhlik *et al.*, 2012, Song *et al.*, 2016) and forest litter (Štursová *et al.*, 2012, Verastegui *et al.*, 2014). *Pedosphaeraceae* (subdivision 3) is an uncharacterized family within the phylum *Verrucomicrobia*, which, like *Acidobacteria,* is widespread among terrestrial environments and has few cultivated representatives (Spring *et al.*, 2016).

Among the fungal communities, the genus *Scleroderma* was the most abundant in the "heavy" fractions and the second most dominant in the control samples after the genus *Mortierella* (which was also present in the heavy fraction). *Scleroderma* is a common, widespread ectomycorrhizal genus that produces macroscopic sporocarps in leaf litter, grass, bare soil or soil adjacent to forests (Jeffries, 1999). *Mortierella* is a root-colonizing endophytic fungus. Members of this globally distributed genus live as saprobes in soil on decaying organic material and dominate fungal communities in natural ecosystems (Johnson *et al.*, 2018).

The significant differences in the bacterial and fungal communities among all treatments were most evident at the later time points, and therefore enrichment of any specific microbial genus in the "heavy" fraction could not be observed at day 8. Several genera and phyla appeared to incorporate WH15EPS, possibly due to the ready hydrolysis of the biopolymer by exoenzymes already present in the litter material at the time of incubation.

At day 24, in the bacterial communities, clear enrichment of the phylum *Planctomycetes* could be observed in the "heavy" fraction samples, reaching approximately 60% of the total sequence number, with the genus *Singulisphaera* as the most abundant. At day 35, the enrichment of *Planctomycetes* persisted, with the genus *Singulisphaera* among the predominant genera but overcome by another group of *Planctomycetes*, uncultured *Pirellulaceae.* At the genus level, the total abundance of *Planctomycetes* was distributed among *Singulisphaera*, *Candidatus Nostocoida* and unclassified groups such as uncultured *Pirellulaceae* and "WD2101 soil group". *Planctomycetes* inhabit a variety of environments, including aquatic and terrestrial habitats as well as extremely acidic environments (Schlesner, 1994, Wang *et*

al., 2002, Ivanova & Dedysh, 2012, Faria *et al.*, 2018). *Planctomycetes* are highly abundant and phylogenetically diverse, especially in *Sphagnum*-dominated wetlands, where they can account for up to 54% of the total amount of 16S rRNA gene sequences (Moore *et al.*, 2015). Members of this phylum have unusual features, such as invaginations of the cytoplasmic membrane (Wiegand *et al.*, 2018). Although the functions of *Planctomycetes* are not clearly understood, they possess a wide range of hydrolytic capabilities, which would explain the promotion of their enrichment by the metabolization of the complex heteropolysaccharide WH15EPS. Using a transcriptome approach, Ivanova *et al*. (2017) observed that cellulose, xylan, pectin and chitin induced responses by different groups of *Planctomycetes*. The genus *Singulisphaera* responded significantly to pectin and xylan amendment. Members of the order *Phycisphaerales* and genus *Zavarzinella* also responded to xylan, while group WD2101 were responsive to cellulose and chitin. Another study by Ivanova *et al*. (2017) further supported the high glycolytic potential of *Planctomycetes.* In that study, a comparative genomic analysis of 4 members of the family *Isosphaeraceae*, namely *Isosphaera pallida*, *Singulisphaera acidiphila*, *Paludisphaera borealis* and the uncharacterized strain SH-PL62S, identified several CAZYmes from major families (GH5, GH13, GH57) as well as potential α-mannosidase, α-rhamnosidase and glucuronyl hydrolase activities. In addition, several CAZYmes not affiliated with currently recognized enzymes were found, demonstrating that these bacteria have the ability to use a wide range of natural carbohydrates and undiscovered glycolytic potential. Furthermore, *Singulisphaera acidiphila* is capable of hydrolyzing several polysaccharides, such as laminarin, pectin, chondroitin sulfate, aesculin, pullulan, lichenan, xylan and gelatin (Kulichevskaya *et al.*, 2008). Similar to the current study, Wang *et al*. (2015) reported enrichment of *Planctomycetes* by indican EPS of *Beijerinckia indica*. Indican is a biopolymer composed of glucuronic acid, glucose and glycerol-manno-heptose, indicating that *Planctomycetes* can decompose complex biopolymers.

Another group found in higher proportions in the "heavy" fractions at days 24 and 35 was uncultured *Methylacidiphilaceae.* This group belongs to the phylum *Verrucomicrobia*, which has few cultivated representatives but has been revealed to have hydrolytic capabilities in cultivation-independent studies. Martinez-Garcia *et al*. (2012) demonstrated that members of *Verrucomicrobia* have high hydrolytic potential and encode a wide spectrum of glycoside hydrolases, carbohydrate lyases and esterases in their genomes, indicating that they are well-equipped with enzymes for carbohydrate metabolism. Cardman *et al*. (2014) demonstrated that fluorescently labeled laminarin and xylan preferentially associated with *Verrucomicrobia* and proposed that this phylum is involved in polysaccharide hydrolysis.

Within the fungal communities, at day 24, an increase in the proportions of the genus *Trichoderma* and groups of unclassified fungi was observed in the "heavy" fraction compared with both the amended and unamended controls. *Trichoderma* is a genus of filamentous ascomycete fungi present in soils or growing on wood, bark and other fungi (mycoparasite) (Druzhinina *et al.*, 2011). This genus is highly opportunistic and adaptable to

several environments, with some strains applied for biocontrol of fungal phytopathogens. *Trichoderma reesei* is capable of decomposing woody and herbaceous materials and is an important industrial producer of hemicellulolytic enzymes (Martinez *et al.*, 2008, Druzhinina *et al.*, 2011).

As expected, the classification of fungal sequences in guilds in the current study revealed that the most abundant fungi were ectomycorrhizal (ECM) fungi and saprotrophs related to organic material decomposition (Valášková *et al.*, 2009, Urbanová *et al.*, 2015). Many species of ECM fungi act as decomposers based on the expression of extracellular enzymes (Bödeker *et al.*, 2009). Burke et al. (2014) indicated that enzyme activity varies greatly among ECM fungi, with some species producing enzymes at levels equivalent to those of saprotrophic fungi.

Co-occurrence and active interactions

Co-inertia analysis revealed significant co-variance between the fungal and bacterial communities, suggesting overall relationships between kingdoms. However, the co-occurrence networks showed that the overall number of connections between microbes was higher in the "heavy" fraction (377) than in either control treatment. The number of fungal genera connected to bacterial genera was lowest in the "heavy" fraction co-occurrence network (2), indicating that potential direct interactions between kingdoms did not increase during the assimilation of the EPS. Shorter arrows in co-inertia analysis for the heavy fraction, however, indicated a stronger relationship between bacterial and fungal communities, which could be due to indirect exchange of metabolites and use of fungal byproducts. Nonetheless, the assimilation of the EPS induced an increase in the potential interactions particularly among bacterial taxa, as observed in the number of edges in the "heavy" fraction network (377). These dynamics were not observed in the EPS-amended control network; nevertheless, not all taxa in this treatment are directly involved in the metabolism of the biopolymer. In addition, *Singulisphaera* had fewer potential interactions, with a variety of microorganisms from different phyla that have either cellulolytic capacity or the ability to remove toxic compounds (Valášková *et al.*, 2009, Kielak *et al.*, 2016). *Granulicella* did not show any strong connection with any taxa in both control treatments.

Fungi are thought to be the main players in the decomposition of recalcitrant materials such as lignocellulose, followed by bacterial decomposition of polysaccharides and polymeric compounds (Boer *et al.*, 2005, Romaní *et al.*, 2006). In the "heavy" fraction, however, only 8 fungal groups, such as Unclassified *Leotiomycetes* and the white-rot genus *Hypholoma* were observed. The fungal groups were mainly positively connected to each other, and negatively with unclassified groups of *Planctomycetes*, indicating competition for the carbon resources derived from the EPS, which hypothesis is supported by the high glycolytic capacity observed in both groups of microorganisms.

The most abundant bacterial genus in the "heavy" fraction, *Singulisphaera*, had mostly positive connections, especially with other groups of *Planctomycetes*, *Actinobacteria* and

Verrucomicrobia, suggesting potential cooperation for metabolizing WH15EPS. Several studies demonstrate that those groups of bacteria have glycolytic and detoxifying capacities, producing several enzymes that could be involved in the degradation of the EPS and resource sharing (Martinez-Garcia *et al.*, 2012, Uhlik *et al.*, 2012, Ivanova *et al.*, 2017). In addition, no direct correlation was found between *Singulisphaera* and the EPS producer *Granulicella,* which suggests that the metabolism of WH15EPS by *Singulisphaera* and enrichment of the genus did not negatively impact the abundance of *Granulicella*. This might be because of high amount of EPS added in the microcosm experiment. By contrast, *Granulicella* had a negative correlation with groups of *Actinobacteria,* demonstrating a potential competition, where *Actinobacteria* are metabolizing WH15EPS and impacting negatively *Granulicella*. *Actinobacteria* are widely spread in the environment and play an essential role in carbon cycling, presenting a wide range of extracellular enzymes (Lacombe-Harvey *et al.*, 2018). Furthermore, the presence of *Granulicella* in the heavy fraction demonstrates the capacity of the genus to use EPS as a carbon source; however, there is no experimental evidence that the producer strain is able to metabolize its own EPS.

It must be noted that the analysis of co-occurrence networks derived from labeled and unlabeled treatments demonstrated that the treatment based on unlabeled total DNA may not reflect real co-occurrence among microorganisms, suggesting this that type of analysis can yield misleading ecological inferences.

Several bacterial and fungal taxa have the ability and potential to metabolize the EPS of *Granulicella* sp. strain WH15. It is not possible to quantify the amount of EPS produced only by this genus in the natural environment; however, our study revealed active interactions between microorganisms in their natural habitat when EPS as added as carbon source. In addition, the structure of the co-occurrence network of microorganisms able to metabolize EPS differed from those in the control treatments, particularly in unlabeled EPS control demonstrating that hidden potential interactions can be unraveled by more specific metabolism studies. Finally, the number of potential associations with uncultured and unclassified taxa reinforces that further efforts are needed to characterize these groups to better understand their functions in the environment.

Acknowledgments

We would like to thank Wietse de Boer for helping with the sampling local, Hans Zweers for the GC measurements, Késia Lourenço for helping with soil analyses and Jos Raaijmakers and Marcio Leite for comments on the manuscript. O.Y.A. Costa was supported by an SWB grant from CNPq [202496/2015-5] (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

Supplementary Material

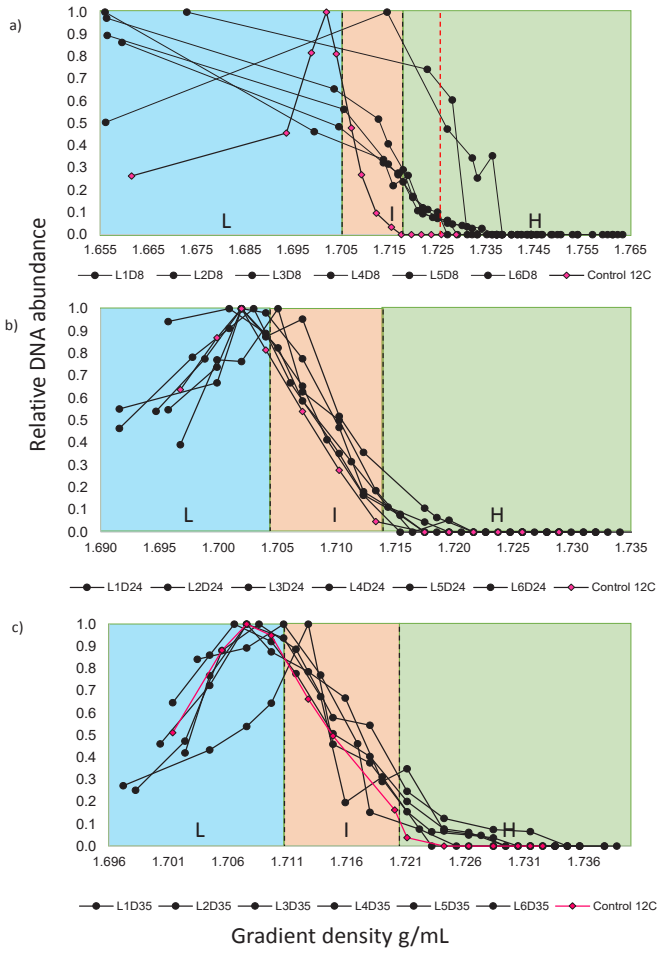

Figure S1: Gradient fractionation. Relative DNA concentrations in CsCl gradient fractions from total DNA extracted at **a)** day 8; **b)** day 24; **c)** day 35. The Y axis indicates the relative DNA concentration recovered from each fraction, with the highest concentration set equal to 1. L (blue) = Light fraction, I (orange) = intermediate fraction, H (green) = heavy fraction. The red dotted line in figure a) indicates the start of the heavy fraction for replicates L2D8 and L6D8.

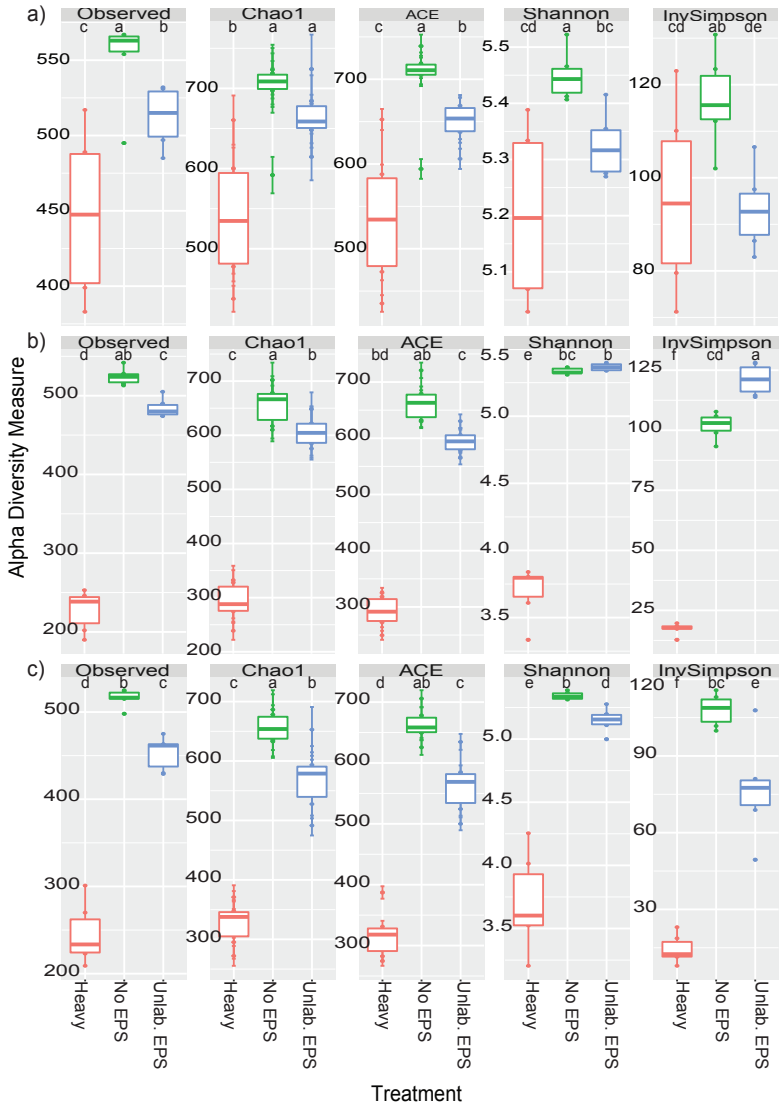

Figure S2: Box-plot comparisons of alpha-diversity assessment by richness estimators (number of observed OTUs, Chao1, ACE) and diversity indices (Shannon, Inverse Simpson) for bacterial 16S rRNA gene samples at different time points. a) day 8; b) day 24; c) day 35. Within each index, identical letters mean no significant difference (p-value < 0.05). Comparisons performed across treatments using ANOVA test and Tukey`s HSD post-hoc test. Data rarefied to the minimum sampling depth. Unlab. EPS-incubation containing ^{12}C-WH15EPS. Heavy – 'heavy fraction' of incubations containing ^{13}C-WH15EPS.

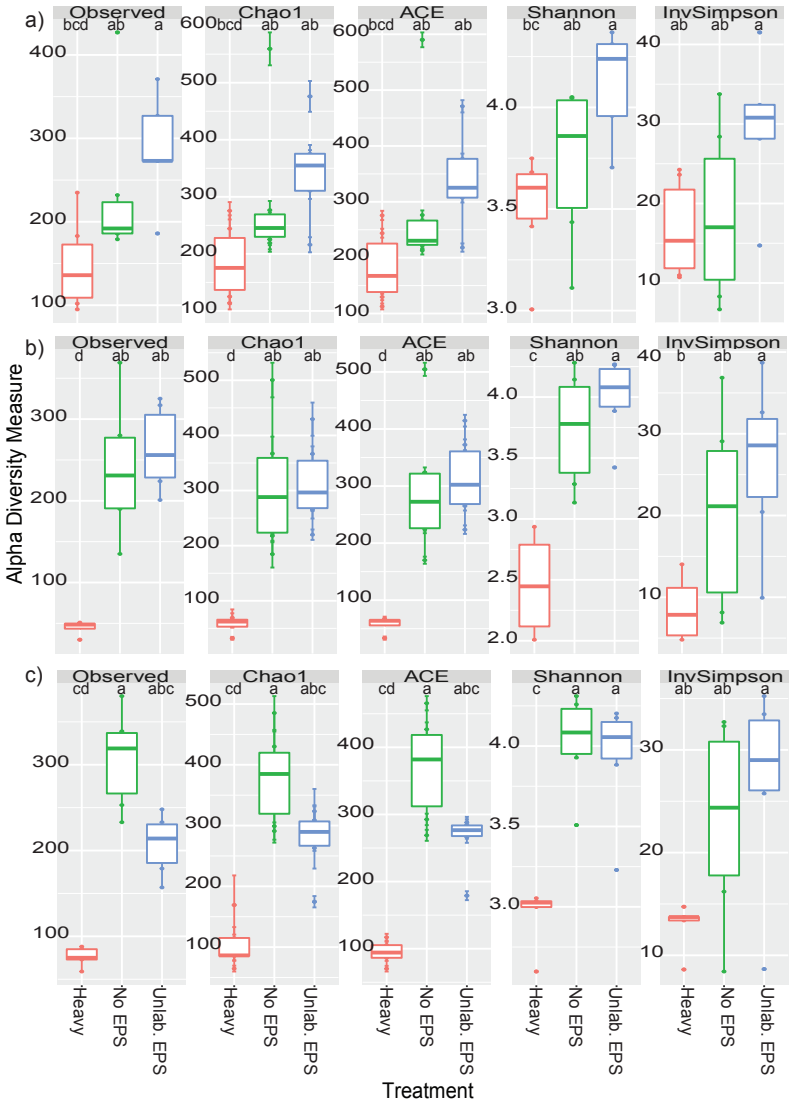

Figure S3: Box-plot comparisons of alpha-diversity assessment by richness estimators (number of observed OTUs, Chao1, ACE) and diversity indices (Shannon, Inverse Simpson) for fungal ITS 1 region gene samples at different time points. a) day 8; b) day 24; c) day 35. Within each index, identical letters mean no significant difference (p-value < 0.05). Comparisons performed across treatments using ANOVA test and Tukey`s HSD post-hoc test. Data rarefied to the minimum sampling depth. Unlab. EPS-incubation containing ^{12}C-WH15EPS. Heavy – 'heavy fraction' of incubations containing ^{13}C-WH15EPS.

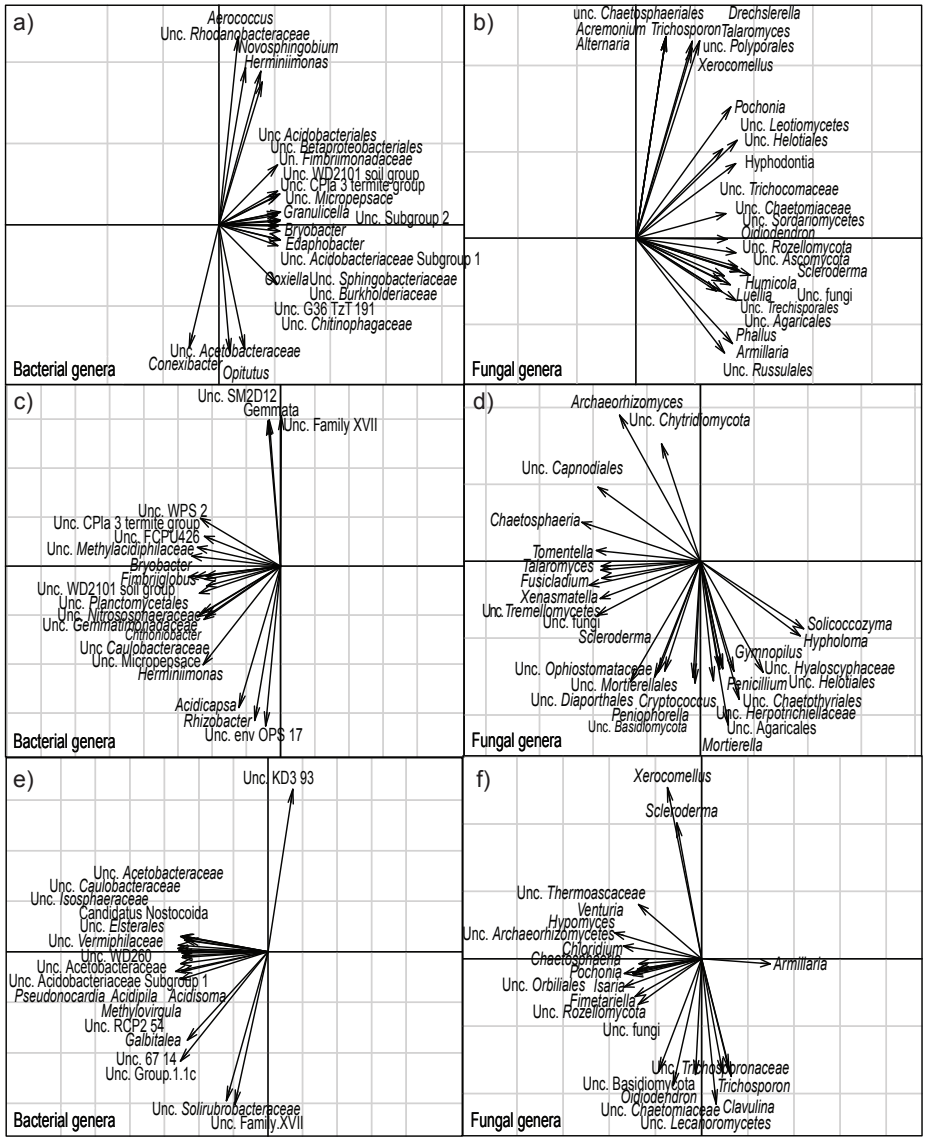

Figure S4: Co-inertia`s (COIA) bacterial and fungal important genera. Top 20 genera of a), c), e) bacteria and b), d), f) fungi that significantly correlated with co-variation of the samples in ^{13}C EPS (a and b), ^{12}C EPS (c and d) and unamended samples (e and f).

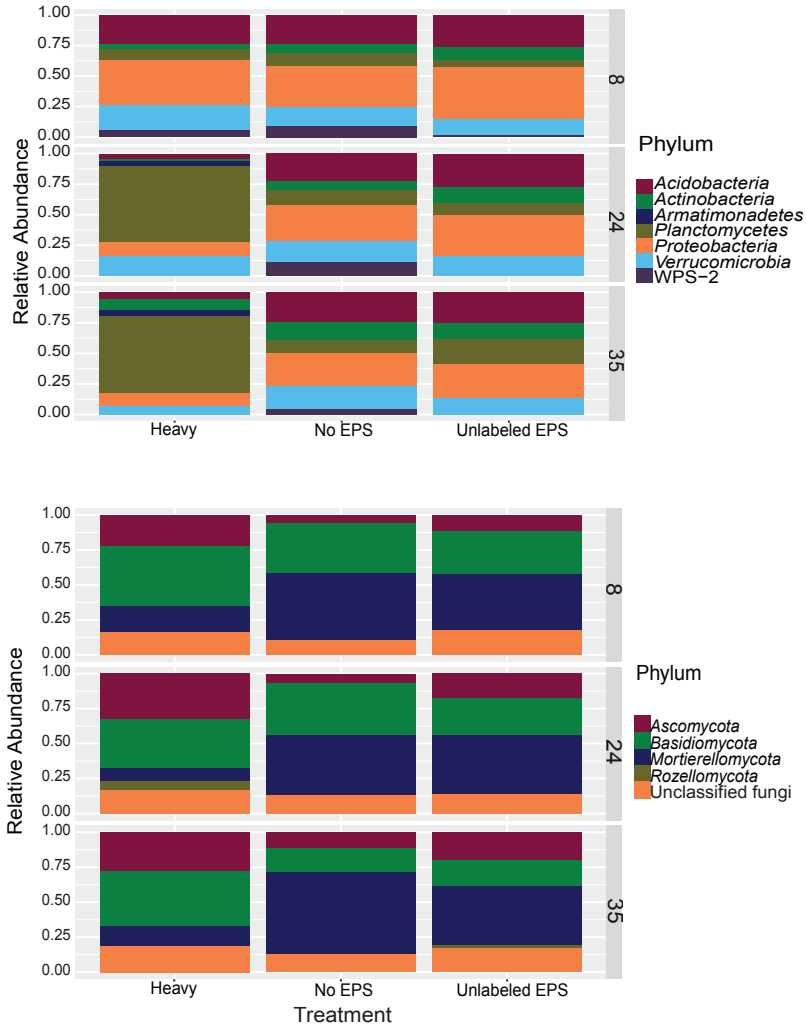

Figure S5: Relative abundance of microbial phyla. Relative abundance of a) bacterial and b) fungal groups at phylum level in heavy fraction, EPS amended (Unlabeled EPS) and unamended (No EPS) controls at all time points (8, 24, 35). Only phyla with >3% abundance are shown.

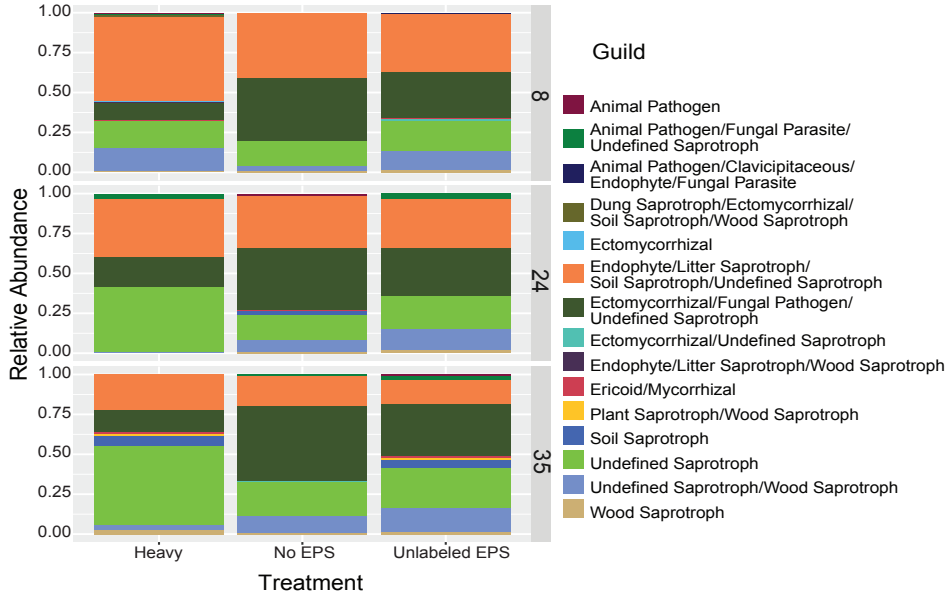

Figure S6: Fungal guilds. Relative abundance of fungal guilds in heavy fraction, EPS amended (Unlabeled EPS) and unamended (No EPS) controls at all time points, based only in the OTUs with assigned guilds (60% of total sequences). Only guilds with >1% abundance are shown.

Table S1: Neighbors of genus *Singulisphaera* and *Granulicella* in co-occurrence network of "heavy" fraction samples.

Group	Phylum	Type of interaction	r-value
Singulisphaera			
Unclassified *Armatimonadales*	*Armatimonadetes*	positive	0.97
Occallatibacter	*Acidobacteria*	positive	0.92
Unclassified Subgroup 2	*Acidobacteria*	negative	-0.81
Conexibacter	*Actinobacteria*	positive	0.90
Unclassified *Solirubrobacterales*	*Actinobacteria*	positive	0.87
Uncultured *Solirubrobacteraceae*	*Actinobacteria*	positive	0.82
Unclassified IMCC26256	*Actinobacteria*	positive	0.82
Uncultured *Acidimicrobiia*	*Actinobacteria*	positive	0.81
Acidothermus	*Actinobacteria*	positive	0.80
Actinospica	*Actinobacteria*	positive	0.75
Unclassified Lineage IV	*Elusimicrobia*	positive	0.82
Uncultured Family XVII	*Firmicutes*	positive	0.86
Candidatus Nostocoida	*Planctomycetes*	positive	0.97
Gemmata	*Planctomycetes*	positive	0.96
Unclassified *Isosphaeraceae*	*Planctomycetes*	positive	0.92
Uncultured *Gemmataceae*	*Planctomycetes*	positive	0.85
Unclassified WD2101 soil group	*Planctomycetes*	negative	-0.82
Uncultured *Pirellulaceae*	*Planctomycetes*	positive	0.97
Anaeromyxobacter	*Proteobacteria*	positive	0.83

Group (continued)	Phylum	Type of interaction	r-value
Unclassified KF JG30 C25	*Proteobacteria*	positive	0.78
Unclassified *Chthoniobacteraceae*	*Verrucomicrobia*	positive	0.93
Uncultured *Methylacidiphilaceae*	*Verrucomicrobia*	positive	0.88
Unclassified *Pedosphaeraceae*	*Verrucomicrobia*	negative	-0.76
Granulicella			
Unclassified *Pedosphaeraceae*	*Verrucomicrobia*	positive	0.88
Unclassified Subgroup 2	*Acidobacteria*	positive	0.85
Edaphobacter	*Acidobacteria*	positive	0.80
Unclassified *Acidobacteriaceae* Subgroup 1	*Acidobacteria*	positive	0.79
Candidatus Koribacter	*Acidobacteria*	positive	0.76
Uncultured_*Acidobacteriales*	*Acidobacteria*	positive	0.75
Acidothermus	*Actinobacteria*	negative	-0.79
Uncultured *Solirubrobacteraceae*	*Actinobacteria*	negative	-0.89
Unclassified 67 14	*Actinobacteria*	negative	-0.82
Uncultured *Acidimicrobiia*	*Actinobacteria*	negative	-0.79
Conexibacter	*Actinobacteria*	negative	-0.75
Mucilaginibacter	*Bacteroidetes*	positive	0.86
Uncultured Family XVII	*Firmicutes*	negative	-0.80
Uncultured *Gemmataceae*	*Planctomycetes*	negative	-0.85
Unclassified WD2101 soil group	*Planctomycetes*	positive	0.76
Uncultured *Xanthobacteraceae*	*Proteobacteria*	positive	0.90
Burkholderia Caballeronia Paraburkholderia	*Proteobacteria*	positive	0.86
Unclassified *WD260*	*Proteobacteria*	positive	0.83
Rhodovastum	*Proteobacteria*	negative	-0.79
Bradyrhizobium	*Proteobacteria*	positive	0.76
Unclassified WPS 2	WPS 2	positive	0.83

Table S2: Neighbors of genus *Singulisphaera* and *Granulicella* in co-occurrence network of unlabeled EPS samples.

Group	Phylum	Type of interaction	r-value
Singulisphaera			
Acidothermus	*Actinobacteria*	positive	0.88
Dyella	*Proteobacteria*	negative	-0.76
Gemmata	*Planctomycetes*	positive	0.76
Rhodanobacter	*Proteobacteria*	negative	-0.76
Unclassified *Armatimonadales*	*Armatimonadetes*	positive	0.79
Unclassified G36 TzT 191	*Proteobacteria*	positive	0.75
Unclassified KF JG30 C25	*Proteobacteria*	positive	0.84
Unclassified *Pedosphaeraceae*	*Verrucomicrobia*	positive	0.83
uncultured *Pirellulaceae*	*Planctomycetes*	positive	0.93
uncultured *Solirubrobacteraceae*	*Actinobacteria*	positive	0.84

Chapter 6

Genetic potential of microbial communities involved in the degradation of a complex acidobacterial extracellular polymer

Ohana Y. A. Costa, Mattias de Hollander, Agata Pijl, Eiko E. Kuramae

Modified version published as: Costa OYA, De Hollander M, Pijl A, Liu BB, Kuramae EE (2020). Cultivation-independent and cultivation-dependent metagenomes reveal genetic and enzymatic potential of microbial community involved in the degradation of a complex microbial polymer. **Microbiome**, 8(76).

Abstract

Polysaccharides are the main components of extracellular polymeric substances (EPS), biopolymers synthesized by a wide range of strains of microorganisms. EPS are the constituents that preserve the tridimensional structure of biofilms, maintaining internal cohesion and promoting adhesion to surfaces. The elimination of biofilms is important for human health because those structures cause problems in hospitals and in food processing industries. To identify EPS degrading microorganisms, we performed Stable Isotope Probing (SIP) combined with metagenomics on topsoil litter amended with EPS of *Granulicella* sp. WH15 (WH15EPS). In addition, we coupled solid culture medium with metagenomics to detect and cultivate potential GH producers. Among all carbohydrate-active enzymes (CAZymes) detected, the most abundant families belonged to Glycoside Transferase (GT) families. Among the glycoside hydrolases (GH), the most abundant family in the metagenomics datasets was amylase family GH13. In the "heavy" fraction of the metagenomics SIP dataset, GH109 (α-N-acetylgalactosaminidases), GH117 (agarases,), GH50 (agarases), GH32 (invertases and inulinases), GH17 (endoglucanases), GH71 (Mutanases) families were more abundant in comparison with the controls. Those GH familes originated from microorganisms that are believed to be able to degrade WH15EPS and potentially applicable for biofilm deconstruction. Subsequent assembly of 4 metagenome-assembled genomes (MAGs) (unclassified *Proteobacteria*) also contained GH families of interest, involving mannosidases, lysozymes, galactosidases and chitinases. We demonstrated that functional diversity induced by the presence of WH15EPS in both culture-dependent and culture independent approaches was enriched in GHs, such as amylases and endoglucanases that could be applied in chemical, pharmaceutical and food industrial sectors. Furthermore, WH15EPS may be used for the investigation and isolation of yet unknown taxa, such as unclassified *Proteobacteria* and *Planctomycetes*, increasing the number of current cultured bacterial representatives.

Keywords: *Acidobacteria*, EPS, Metagenonics, Stable Isotope Probing, *Planctomycetes*, Carbohydrates

1. Introduction

Exopolysaccharides are the main and most studied components of extracellular polymeric substances (EPS), biopolymers synthesized by a wide range of strains of microorganisms (Flemming & Wingender, 2010, Costa *et al.*, 2018). EPS are the constituents that preserve the tridimensional structure of biofilms, maintaining internal cohesion and promoting adhesion to surfaces (Flemming & Wingender, 2010). The elimination of biofilms is important for human health in general, because those structures are implicated in several diseases, causing problems for instance in hospitals and in food processing industries (Hunter, 2008, Nahar *et al.*, 2018). Enzymatic removal of biofilms is superior to the use of conventional cleaning agents, which are not eco-friendly, producing toxic residues and erosion of equipment (Nahar *et al.*, 2018). Enzymes are an environmentally friendly alternative due to their biodegradable nature (Liu & Kokare, 2017). EPS and biofilms are complex, requiring a wide range of enzymes for a complete degradation (Flemming & Wingender, 2010).

Glycoside hydrolases (GHs) are hydrolytic enzymes which can be applied for the degradation of EPS polysaccharides for the removal of biofilms (Fleming *et al.*, 2016). Furthermore, GHs are among the industrially important enzymes that are extensively searched through metagenomics, as they are extremely desired and important in food and other industrial sectors. Those enzymes are employed for brewing, baking, production of syrups, food processing, texture, flavoring, as well as the production of dairy and fermented foods (Coughlan *et al.*, 2015). GHs are also necessary for the production of biofuels, by converting cellulose and lignocellulosic biomass into sugars that can be fermented into bioethanol by microorganisms (Ezeilo *et al.*, 2017). More than 50% of the current industrial enzymes are produced by microorganisms, such as strains of *Bacillus* and *Aspergillus*, while around 15% are derived from plants (Liu & Kokare, 2017). In addition, microbial enzymes with potential applications were obtained from habitats such as hydrothermal vents (Legin *et al.*, 1997), arctic tundra (Oh *et al.*, 2019), cow rumen (Hess *et al.*, 2011) and termite guts (Warnecke *et al.*, 2007).

The main goal of our study was to use a microbial EPS to find microbes and functions involved in EPS degradation. We used topsoil-plant litter as a source of microbes. Plant litter is mostly composed of recalcitrant biopolymers, which are sources of carbon, energy and nutrients for microbial communities living in litters layers (Urbanová *et al.*, 2015, Vivanco *et al.*, 2018). Due to their complexity, the breakdown of plant cell wall components requires a wide range of enzymes, produced by the microorganisms during litter decomposition process (Schneider *et al.*, 2012, Chen *et al.*, 2018). Therefore, it is an interesting environment for the retrieval of complex polysaccharide-degrading enzymes. We applied the EPS of the *Acidobacteria Granulicella* sp. strain WH15 (WH15EPS). The unique composition of its EPS may be a source to retrieval of a wide range of novel glycoside hydrolase genes that could be employed in the industry for several processes. WH15EPS has a more complex composition than most commercially available microbial polymers. It is composed of 7 monosaccharides

(mannose, glucose, galactose, xylose, rhamnose, glucuronic and galacturonic acids) (Kielak *et al.*, 2017), while other known EPS are composed of maximum 4 different monosaccharides (Rehm, 2010). The degradation of WH15EPS would require a broader range of enzymes than other EPS, therefore we hypothesized that the application of WH15EPS to topsoil-litter samples would promote the enrichment of a wider range of GHs. We performed Stable Isotope Probing (SIP) combined with metagenomics, and coupled solid culture medium with metagenomics using WH15EPS as an enrichment factor to detect and cultivate potential GH producers and find GH genes with biotechnological potential.

2. Material and Methods

2.1. Soil samples

Four topsoil-litter mixed samples were collected in the spring of 2017 from the Wolfheze forest in the Netherlands (Table S1). Samples were taken from topsoil (0 to 5 cm) adjacent to fallen tree trunks. The collected samples were pooled, sieved (2 mm mesh) and immediately used for SIP incubation with EPS from *Granulicella* sp. strain WH15 (WH15EPS). The physicochemical properties of the topsoil-litter samples were determined (Eurofins Agro BV, Wageningen, NL) and are presented in Table S2.

2.2. SIP metagenome

The methods used for WH15EPS labeling, purification, sample collection, incubation and DNA extraction and fractionation are described in chapter 5. Library preparation and high-throughput shotgun sequencing were performed using the "heavy" DNA fractions pooled within each sample replicate as well as the total DNA of both the ^{12}C-EPS-amended and unamended controls. Shotgun sequencing was performed in 4 replicates of samples collected after 35 days of incubation (latest timepoint). Library preparation and Illumina MiSeq PE250 shotgun sequencing were performed at McGill University and Génome Québec Innovation Centre (Montréal, Québec, Canada). The sequences were deposited in the European Nucleotide Archive (ENA; https://www.ebi.ac.uk/ena) under the accession number PRJEB31257.

2.3. Metagenome of cultivated microorganisms grown in culture media with WH15EPS as sole carbon source

For evaluation of the metagenome of microorganisms that were able to grow in culture medium with WH15EPS as a sole carbon source, ten grams of fresh topsoil-litter sample were mixed with 100 ml of 100 mM MES buffer (2-[N-morpholino]ethanesulphonic acid, 1.95 g/l, pH 5.5), agitated for 30 min at room temperature on a vortex and decanted for 30 min. Dilutions (10^{-3} to 10^{-6}) were prepared in sterile MES buffer and 200 µl of the dilutions were plated in

quadruplicate. Diluted culture medium DNMS [$MgSO_4.7H_2O$ 0.2g/l, $CaCl_2.2H_2O$ 0.053g/l, chelated iron solution 0.2 ml/l (ferric III ammonium citrate 0.1g/100 ml, EDTA 0.2g/100ml, HCl 0.3ml/100ml) trace element solution SL10 1ml/L (Atlas, 2010), NH_4Cl 0.1g/l, agar 20g/l] with added WH15EPS (Kielak *et al.*, 2017) (0.05%) pH 5.5 and 40ng/µl (40 mg/l) cicloheximide to prevent growth of fungi was used for plating. To prevent caramelization, the freeze-dried purified WH15EPS was hydrated with milli-Q water, sterilized by filtration through a 0.2 µm membrane (Millipore) and added to the culture medium after autoclaving. Chelated iron solution and trace element solution SL10 were added after autoclaving and cooling of the culture medium. The plates inoculated with the soil suspension were incubated at room temperature for 1 month. The dilution 10^{-3} was chosen for sequencing. After incubation, colonies were scraped and used for total DNA extraction with PowerSoil® DNA Isolation Kit (MO BIO Laboratories, Inc). Following the first DNA extraction, a second round of DNA extraction was performed for each sample, according to Dimitrov et al. (2017). The total DNA extracted from the plates was used for metagenome shotgun sequencing. Library preparation and Illumina HiSeq XTen sequencing were performed at Genewiz (Suzhou, China). The sequences were deposited in the European Nucleotide Archive (ENA; https://www.ebi.ac.uk/ena) under the accession number PRJEB24069.

2.4. Bioinformatics and statistical analyses of metagenome data

2.4.1. SIP metagenome

SIP metagenome sequences were processed using EBI MGnify (Mitchell *et al.*, 2017) pipeline and SqueezeMeta (Tamames & Puente-Sánchez, 2019) pipeline in sequential mode. Briefly, in the SqueezeMeta pipeline, trimming and quality filtering were performed using Trimmomatic (Bolger *et al.*, 2014), assembly for each sample separately was done using Megahit (Li *et al.*, 2015); Prodigal (Hyatt *et al.*, 2010) was used for Open Reading Frame (ORF) prediction, and barrnap (Seemann, 2018) was employed for small subunit (SSU) rRNA gene sequence retrieval, which were classified using RDP classifier (Wang *et al.*, 2007). Diamond (Buchfink *et al.*, 2014) software was used for taxonomic classification of the ORFs against Genbank nr database and functional annotation with eggNOG database, for KO and COG numbers (Huerta-Cepas *et al.*, 2016). eggNOG-mapper (Huerta-Cepas *et al.*, 2017) was employed for carbohydrate-active enzymes annotation with against dbCAN database (Yin *et al.*, 2012). SqueezeMeta script SQM2tables.py was used to compute the average coverage and normalized TPM (transcripts per million) values for information on gene and function abundances. Normalized TPM SqueezeMeta ORF dataset and SSU rRNA gene data recovered from MGNify analysis were used for statistical analyses, performed in RStudio version 1.1.423 running R version 3.5.1 (R Core Team, 2015). For the SSU rRNA gene-based analysis, OTUs with less than 1 count across all the samples, chloroplast and mitochondrial sequences were discarded; prior to alpha diversity analyses, the data were rarefied to the size of the smallest sample (175 reads). For both ORF-based and SSU rRNA gene-based taxonomy datasets, 'Phyloseq' package

(McMurdie & Holmes, 2013) was used to calculate the number of observed OTUs, Shannon and Inverse Simpson diversity indices, and Chao1 and ACE diversity estimators. Bray-Curtis distance matrices constructed using the Hellinger transformed (Legendre & Gallagher, 2001) datasets were used for principal coordinate analysis (PCoA) using the capscale function from the 'vegan' package v. 2.4.6 (Oksanen *et al.*, 2018). Group dissimilarities were tested by permutational multivariate analysis of variance (PERMANOVA) using the function Adonis from the 'vegan' package. CANOCO (version5) (Braak & Smilauer, 2012) was employed to explore the relationship between sample treatments and taxa abundance through redundancy analysis (RDA) in the Hellinger transformed datasets. The statistical significance (p-value <0.05) of eigenvalues and treatment-taxa abundance correlations were tested using Monte Carlo permutation test at 499 permutations and the top 20 taxa associated with the dispersion of the treatments were displayed in RDA graphs.

In order to identify predicted functions (COG, KEGG and CAZymes) responsible for the observed clustering patterns, we performed a feature selection using a 'random forest' algorithm using the R package Boruta (Kursa & Rudnicki, 2010) (1,000 trees, p-value <0.05). Boruta tests if the importance of each individual variable is significantly higher that the importance of a random variable by fitting random forest models iteratively until all predictor variables are classified as "confirmed" or "rejected" at the 0.05 alpha level (Leutner *et al.*, 2012). The heatmaps for each function were constructed with pheamap (Kolde, 2019) R package, based on z-score transformed TPM (transcripts per million) abundances to improve normality and homogeneity of the variances. Sequences were submitted to the European Nucleotide Archive (ENA) and are available under the accession number PRJEB31257.

2.4.2. Metagenome of cultivated microorganisms

The DNA of the cultivated microorganisms were shotgun metagenome sequenced and the sequences were processed using EBI MGnify (Mitchell *et al.*, 2017) pipeline and ATLAS (Automatic Tool for Local Assembly Structures) (White III *et al.*, 2017) pipeline. For ATLAS, quality filtering was performed using BBDuk2 and cross-assembly was done with Megahit (Li *et al.*, 2015); functional and taxonomic analysis were performed at ORF level for the assembled contigs. Prodigal (Hyatt *et al.*, 2010) was used for ORF prediction and eggNOG database (Huerta-Cepas *et al.*, 2016) was used for functional annotation (COG and KO numbers) using DIAMOND software (Buchfink *et al.*, 2014). EggNOG-mapper (Huerta-Cepas *et al.*, 2017) was used for functional annotation of CAZymes with dbCAN database (Yin *et al.*, 2012). Kaiju software (Menzel *et al.*, 2016) was used for ORF taxonomy assignment against NCBI RefSeq database. Custom scripts were used to generate tables containing information of taxonomy and function abundance of the ORFs in all samples. Quality controlled contigs >1000 kb were used for binning using Concoct (Alneberg *et al.*, 2014), Maxbin (Wu *et al.*, 2014) and Metabat (Kang *et al.*, 2015); resulting bins were refined using DAS tool(Sieber *et al.*, 2018) and genome dereplication was performed with dRep (Olm *et al.*, 2017). Completeness and

contamination of the assembled genomes were checked using CheckM (Parks *et al.*, 2015), as well as taxonomy assignment. The ORFs of the genomes were predicted using Prodigal (Hyatt *et al.*, 2010), and DIAMOND software (Buchfink *et al.*, 2014) was used for functional annotation with eggNOG (COG and KO numbers) (Huerta-Cepas *et al.*, 2016). The annotation of CAZymes was performed with EggNOG-mapper (Huerta-Cepas *et al.*, 2017) against dbCAN database (Yin *et al.*, 2012). Sequences were submitted to the European Nucleotide Archive (ENA) and are available under the accession number PRJEB24069.

3. Results

3.1. SIP metagenome

3.1.1. Overview of the metagenome data

After quality control filtering, a total of 18,762,958 reads were maintained for further analysis, with an average of 1,563,580 reads per sample. A total of 1,209,745 ORFs were predicted for functional annotation, and approximately 50% of these ORFs were classified using KEGG and COG databases. The sequencing statistics are in Table 1.

Table1: SIP shotgun metagenomics sequencing statistics for each treatment. Average from 4 replicates.

Sequence statistics	Unamended Control	EPS amended	"Heavy" fraction
Number of reads	1,590,046	1,591,447	1,509,247
Number of contigs	82,370.25	78,474.5	92,521.75
Longest contig (bp)	2,651	3,645.25	9,414.75
N50	466	473	542
Mapping (%)	18.4	19.6	32.9
Number of ORFs	96,141	91,272	115,023.3

3.1.2. Community composition based on SSU rRNA and ORF classification

Taxonomic annotation based on SSU rRNA annotation demonstrated that bacteria, fungi and archaea accounted for approximately 84%, 4% and 2% of the sequences, respectively. At phylum level, 17 bacterial groups, 5 fungal groups and 3 archaeal groups were observed in all the samples. The most abundant groups at phylum level belonged to kingdom *Bacteria* (Figure S1a). *Proteobacteria* was the most abundant phylum in all treatments (26.4-28% of the sequences), followed by *Actinobacteria* (14.5-17.5% of the sequences). In both unamended and ^{12}C-EPS-amended control treatments, *Acidobacteria* was the third most abundant group (14.5-15.8% of the sequences), while in the "heavy" fraction samples, *Planctomycetes* was the third most abundant phylum (16.45% of the sequences) (Figure S1a). At genus level, we observed 167 groups in all samples, of which 110 were unclassified groups. "Unclassified *Bacteria*" was the most abundant group in the unamended control (3.5% of the sequences),

while "unclassified *Acidobacteriaceae*" (2.6% of the sequences) was the most abundant in the ^{12}C-EPS-amended control (Figure 1a). In labeled samples, the predominant group was "unclassified *Planctomycetes*" (3.2% of the sequences) (Figure 1a). Among the 10 most abundant groups, only 2 classified genera were observed: *Acidothermus* (1.8-2.9% of the sequences) and *Singulisphaera* (0.2-2.6% of the sequences) (Figure 1a). Similarly, the taxonomic composition of the ORF-based analysis was dominated by kingdom *Bacteria*, with an average of 82% of the ORFs belonging to bacteria and approximately 18% of the ORFs

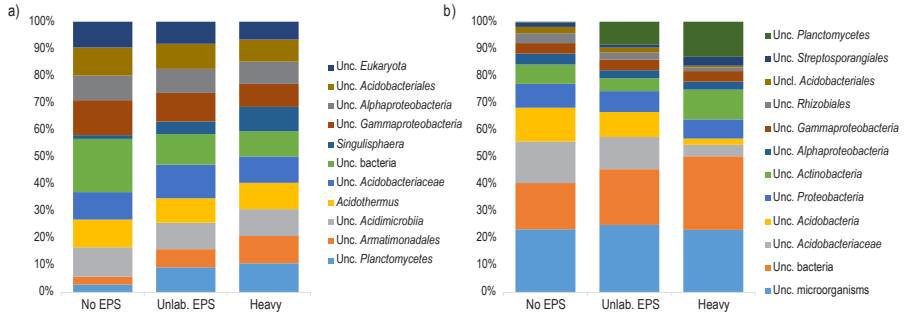

Figure 1: Taxonomic composition and relative abundance of microbial groups at genus level in SIP metagenome treatments based on a) SSU rRNA gene taxonomic classification b) ORF taxonomic classification. Only the ten most abundant groups for each treatment are displayed. Average abundances of 4 replicates. Unc: unclassified. No EPS – incubation without WH15EPS. Unlab. EPS-incubation containing ^{12}C-WH15EPS. Heavy – 'heavy fraction' of incubations containing ^{13}C-WH15EPS

originating from unclassified organisms, in all the samples (Figure S1b). At phylum level, we observed, in total, 103 bacterial groups, 6 fungal groups and 11 archaeal groups in all the samples. *Acidobacteria* (20.1-25.3% of the sequences) was the most abundant phylum in unamended and ^{12}C-EPS-amended control samples, while *Actinobacteria* (26% of the sequences) was the predominant group in "heavy" fraction samples (Figure S1b). At genus level, we found 1541 groups, of which 667 were unclassified. The top three most abundant groups in both control treatments were "unclassified microorganisms" (17.3-19.4% of the ORFs), "unclassified *Bacteria*" (12.7-16% of the ORFs) and "unclassified *Acidobacteriaceae*" (9.3-11.5%), while the predominant groups in "heavy" fraction samples were "unclassified microorganisms" (16.1% of the ORFs), "unclassified *Bacteria*" (18.9% of the ORFs) and "unclassified *Planctomycetes*" (9% of the ORFs) (Figure 1b).

PERMANOVA (p-values<0.001) showed that, for both SSU rRNA data and ORF based analysis, the microbial communities were different between treatments, with both control treatments closer to each other, and "heavy" fraction samples separated from both control treatments in PCoA graphs (Figure 2). For SSU rRNA communities, the first two axes of PCoA explained 43.3% of the variation, while for ORF based data, 90.6% of the variation was explained. RDA analysis for both datasets showed that mainly groups of *Planctomycetes,* such as "unclassified *Planctomycetes*", "unclassified *Planctomycetales"*, "unclassified *Planctomycetia"* and *Singulisphaera*, were driving the dispersion of the microbial communities between "heavy"

fraction and both control treatments (Figure S2), consistently with the higher abundance of *Planctomycetes* in labeled samples. Alpha diversity indices showed that richness and diversity indices were lower for "heavy" fraction samples in comparison with both controls (Figure S3), supported by ANOVA test (p-value < 0.05).

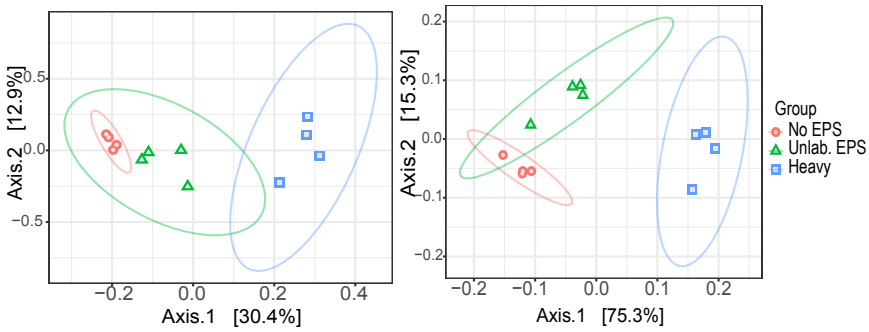

Figure 2:. Principal Coordinate Analysis (PCoA) clustering of normalized and Hellinger-transformed SIP metagenome sequencing data based on Bray-Curtis distances of a) SSU rRNA gene taxonomic classification and b) ORF taxonomic classification. No EPS – incubation without WH15EPS. Unlab. EPS-incubation containing ^{12}C-WH15EPS. Heavy – 'heavy fraction' of incubations containing ^{13}C-WH15EPS

3.1.3. Functional profile

KEGG, COG and dbCAN databases were employed for functional gene annotation to explore the functional characteristics of the microbial communities. Approximately 60% of the ORFs were assigned to COGs, matching in total to 20,644 COGs. The most abundant COG categories in all the samples were "R-General function prediction" (10.8-11.6% of the ORFs), "E-Aminoacid transport and metabolism" (10-5-11% of the ORFs), "C-Energy production and conversion" (8.1-8.6%), "L-Replication, recombination and repair" (7.1-7.6% of the ORFs) and "G-Carbohydrate transport and metabolism" (6.4-6.8%) (Figure S4a). Boruta feature selection "random forest" analysis was used to identify feature annotations that segregated significantly (p-value <0.05) between treatments. A total of 32 COGs were selected by Boruta algorithm. 13 among the identified COGs were more abundant in the unamended control samples, while 19 were more abundant in the labeled samples (Figure 3a). However, most of the features identified by the analysis belonged to the category unknown function. Some of the unknown COGs abundant in the labeled treatment, though, were associated mostly to phyla *Planctomycetes* and *Acidobacteria*, according to eggNOG database v 4.5 (Table S3).
KEGG analysis demonstrated that about 50% of the ORFs were assigned to 7,343 KEGG functional orthologs. The 17 most abundant KEGGs in all samples were assigned to three categories: signaling and cellular processes (8 KEGGs – 0.16% of the total ORFs), genetic information and processing (6 KEGGs – 0.14% of the total ORFs) and metabolism (3 – 0.21% of the total ORFs) (Figure S4b). Boruta feature selection identified 40 KEGGs that influenced the dispersion of the samples, of which 26 were more abundant in the labeled treatment and 14 were more abundant in the unamended control (Figure 3b). Among the KEGGs more

abundant in the labeled treatment, 13 could be assigned to KEGG pathways, mostly related to “metabolic pathways” and “microbial metabolism in diverse environments” (Table S4). Within the KEGGs more abundant in the unamended control treatment, 8 could be assigned to KEGG pathways, the majority related to “metabolic pathways” (Figure 3b, Table S4).

Annotation using dbCAN database showed that families GT41 (8.4%-11% of the CAZYmes), AA3 (4.4%-5%), GT4 (3.4%-4.7%), GT2 (4.1%-4.3%) and CE10 (3.5%-4.2%) were among the most predominant in all the treatments (Figure S4c). Boruta feature selection identified 27 CAZyme families affecting the dispersion of the sample treatments (Figure 3c), the vast majority belonging to the category glycoside hydrolase (GH). Among the selected families, 15 were more abundant in the labeled treatment and 12 were more abundant in the unamended control. The categories abundant in the labeled treatment involved xylan

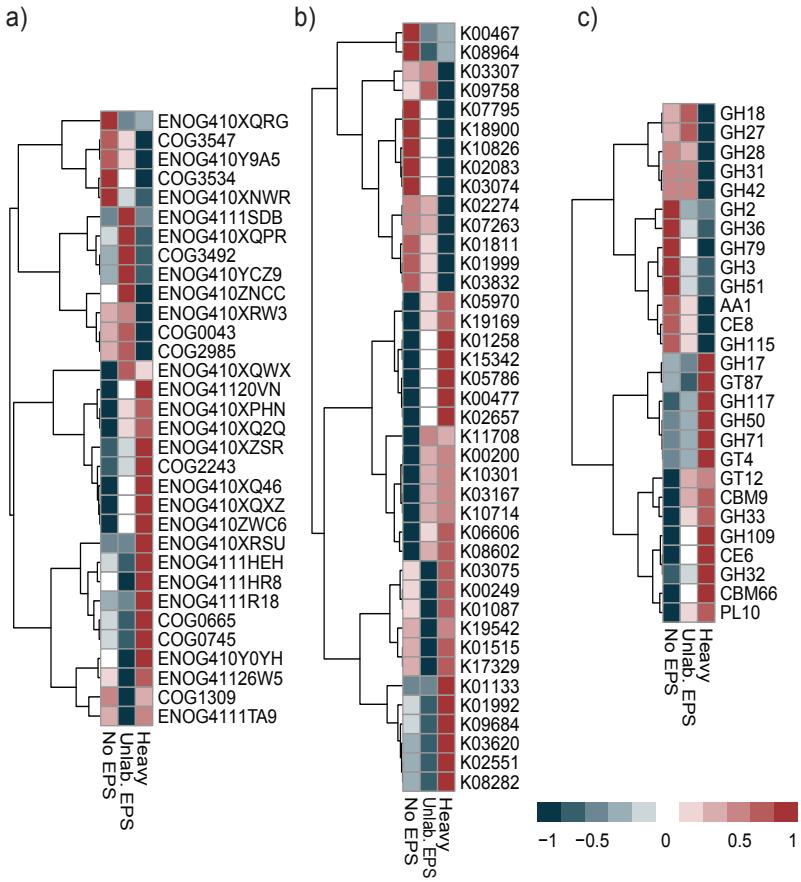

Figure 3: Boruta random forest feature selection of functions that significantly segregated across treatments based on 1000 permutations (p-value <0.05) for a) COG annotation, b) KEGG annotation and c) dbCAN annotation. Heatmaps based on the z-scored TPM normalized relative abundances of annotated ORFs from SIP metagenome samples. The description of the functions displayed in the heatmap are detailed in Table S3 (COG), Table S4 (KEGG) and Table S5 (dbCAN). No EPS – incubation without WH15EPS. Unlab. EPS-incubation containing ^{12}C-WH15EPS. Heavy – ‘heavy fraction’ of incubations containing ^{13}C-WH15EPS.

and fructan modules, xylanases, mannosytransferases and agarases, while the categories abundant in the unamended controls are mostly α and β galactosidases and glucosidases (Table S5). PERMANOVA (p-values<0.001) demonstrated that for KEGG, COG and dbCAN data, the functional gene compositions were different between treatments, similarly to taxonomic analysis, with control treatments grouping together and separated from "heavy" fraction samples (Figure S5).

3.2. Metagenome of cultivated microorganisms

3.2.1. Overview of the metagenomics data

A total of 422,735,048 reads were obtained after sequence quality 01 filtering, with an average of 80% of the ORFs classified with KEGG and COG databases. The sequencing statistics are described in Table 2.

Table 2: Sequencing statistics for shotgun metagenome of cultivated microorganisms. Average from 2 replicates per plate of culture medium.

Sequence statistics	Plate1	Plate2	Plate3	Plate4
Number of reads	49,148,370	54,247,258.5	58,397,852	49,574,043.5
Assembled reads	1.47^{E+10}	1.6224^{E+10}	1.75^{E+10}	1.4796^{E+10}
Number of contigs	67,980,868	76,125,070.5	82,202,170	68,767,888
Number of predicted genes	159,832	254,727	535,677	479,683
KEGG (% classified ORFs)	66.2	67.2	65.2	63.9
COG (%classified ORFs)	94.6	94.5	94.1	94.0
CAZYmes (%)	4.5	4.7	4.7	4.5
GC content (%)	59.6	60.3	59.0	58.2

3.2.2. Community composition based on SSU rRNA and ORF classification

Analysis of the taxonomic composition based on SSU rRNA showed an average of 73% of the sequences belonged to kingdom *Bacteria*, 20% to kingdom *Fungi* and 7% were derived from other Eukaryotes (Figure S6a). At phylum level, 17 bacterial groups, 7 fungal groups and 14 eukaryotic groups were identified. The most abundant group was the bacterial phylum *Proteobacteria*, with ~47.9% of the sequences, followed by fungal phylum *Ascomycota*, with ~ 14.5% of the sequences (Figure S6b). At genus level, 450 groups in total were observed, with the most abundant groups being bacterial groups. The predominant groups were "unclassified *Bacteria"* (~2.2% of the sequences) and *Dyella* (~1.5% of the sequences) (Figure 4a). *Silvimonas* and *Burkholderia* were also among the top 10 most abundant genera (~1.4 and 1.3% of the sequences, respectively). Similarly, for the ORF based data, the most abundant groups at genus level belonged to kingdom *Bacteria,* revealing the presence of 1930 groups at genus level. "Unclassified microbes" was the most abundant group, followed by genera *Caballeronia* (15.4% of the ORFs) and *Paraburkholderia* (15.1% of the ORFs) (Figure 4b). Other genera, such as *Burkholderia*, *Rhodanobacter* and *Dyella* were also among the predominant groups (7.8, 7.1 and 4.9% of the ORFs) (Figure 4b).

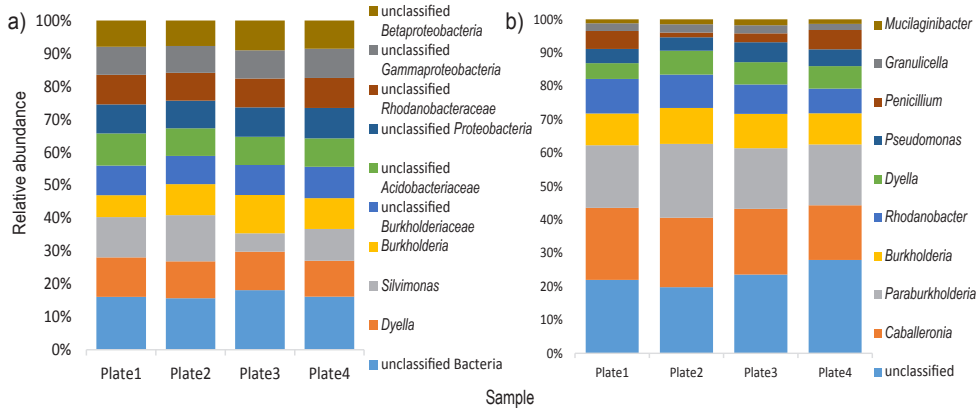

Figure 4: Taxonomic composition and relative abundance of microbial groups at genus level in samples from the metagenome shotgun of cultivated microrganims based on a) SSU rRNA gene taxonomic classification b) ORF taxonomic classification. Only the ten most abundant groups are displayed. Average from 2 replicates per plate of culture medium.

3.2.3. Functional profile

The functional profile of the cultivated microbes' metagenome was explored through the annotation with KEGG, COG and dbCAN databases. COG analysis demonstrated that approximately 20.6% of the annotated COGs were assigned to category "S-function unknown". Among the classified COGs, similarly to SIP metagenome, the predominant categories involved "E-Aminoacid transport and metabolism" (~8.6% of the ORFs), "G-Carbohydrate transport and metabolism" (~8.0% of the ORFs) and "C-Energy production and conversion" (~7.3% of the ORFs) (Figure 5a).

KEGG pathway analysis showed that around 65% of the ORFs were assigned to 9945 KEGG orthologs. The 20 most abundant KEGGs were distributed in the categories "Genetic information processing" (1 KEGG- ~0.24% of the total ORFs), "Metabolism" (4 KEGGs - ~1.18% of the total ORFs) and "Signaling and cellular processes" (15 –KEGGs -4.54% of the ORFs), of which 13 KEGGs were classified as transporters (Figure 5b).

The analysis of the carbohydrate-active enzymes with dbCAN database demonstrated the presence of 298 CAZyme families. Twenty-three families were predominant, which abundance was above 1%. Within the most abundant families, we observed 2 AA families (7.75% of the CAZymes), 1 CBM family, 4 CE families, 10 GH families and 6 GT families (Figure 5c). Those CAZyme families comprise mostly enzymes with cellulolytic (alpha-glucosidases, alpha fucosidases), hemicellulolytic (alpha-rhamnosidases, alpha-xylosidases, alpha-mannosidases, beta-galactosidases) and cell wall metabolism activities (N-acetylglucosaminyltransferases, alpha-N-acetylgalactosaminidases and peptidoglycan lyases) (Table S6). The most abundant family was GT 41 (Figure 5c), which encompasses UDP-GlcNAc: peptide β-N-acetylglucosaminyltransferases and UDP-Glc: peptide N-β-glucosyltransferases, enzymes envolved in protein glycosilation. Among the GH families, the

most abundant was GH13.

Among all 127 GH families found in both metagenome datasets, 114 families were observed in both datasets, while 5 families were exclusive from the SIP dataset (GH112, GH48, GH52, GH86, GH98) and 8 were exclusive from the cultivated microbial dataset (GH111, GH131, GH132, GH134, GH45, GH7, GH80, GH85) (Figure S7).

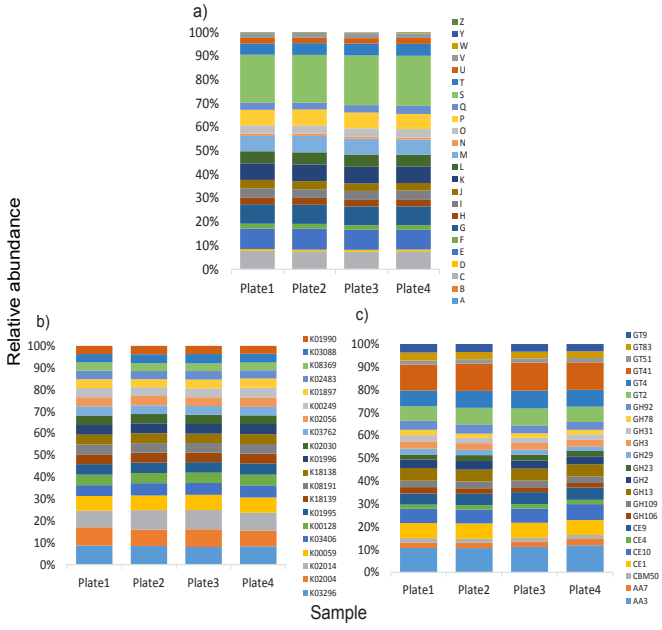

Figure 5: Relative abundance distribution of the most abundant functional categories in TMM-normalized metagenome sequencing data from the shotgun metagenome of cultivated microorganisms. a) COG annotation (all categories); b) KEGG annotation (above 0.2 % abundance); c) dbCAN annotation (above 1% abundance). The description of the functions displayed in b) and c) are detailed in Table S6. Average from 2 replicates per plate of culture medium. E-Amino acid transport and metabolism; G- Carbohydrate transport and metabolism; H-Coenzyme transport and metabolism; C-Energy production and conversion; I-Lipid transport and metabolism; F-Nucleotide transport and metabolism; Q- Secondary metabolites; D-Cell cycle; N-Cell motility; M-Cell wall/membrane/envelope biogenesis; V-Defence mechanisms; P-Inorganic ion transport and metabolism; U-Intracellular trafficking; O-Post translational modification; T-Signal transduction mechanisms; L-Replication, recombination and repair; K-Transcription; J-Translation; S-Function unknown; R-General function and prediction; X-Mobilome.

3.2.4. Taxonomy of the enriched glycoside hydrolase families

Taxonomic analysis of the most abundant GH family in both metagenome datasets, GH13, demonstrated that the majority of the sequences of GH13 in the cultivated microbes dataset belonged to phyla *Proteobacteria* (66.8% of the GH sequences) and *Acidobacteria* (21.8% of the GH sequences), while in the SIP dataset the most abundant phyla for GH13 were *Actinobacteria* (20.4-45.7% of the GH sequences), *Acidobacteria* (4-24.7% of the sequences) and other phyla (27-34% of the GH sequences) (Table3).

Table 3: Taxonomy associated to sequences of glycoside hydrolases belonging to GH13 family (most abundant) and the enriched GH families in heavy fraction samples from SIP metagenome.*

GH families	Sample	*Proteobacteria*	*Acidobacteria*	*Actinobacteria*	*Planctomycetes*	Others
GH13	Cultivated	66.8 (2605)	21.8 (827)	0.9 (36)	0.09 (1)	10.4 (51.4)
GH13_SIP	Control	20.1(128)	23.4 (148)	25.6 (162)	0.1(4)	31(196)
	EPS	17.9(89)	24.7(125)	20.4(95)	2.9(13)	34(172)
	Labeled	17.2(127)	4(30)	45.7(333)	5.7(42)	27(201)
GH109	Control	11(22)	45(92)	12(26)	2(4)	31(68)
	EPS	7(19)	26(76)	7(18)	21(60.8)	40(111.6)
	Labeled	7(34)	9(48)	13(62)	29(150)	42(217)
GH117	Control	0(0)	33(1)	33(1)	0(0)	33(1)
	EPS	0(0)	0(0)	17(1)	0(0)	38(5)
	Labeled	9(1)	0(0)	27(3)	0(0)	64(7)
GH50	Control	100(3)	0(0)	0(0)	0(0)	0(0)
	EPS	100(2)	0(0)	0(0)	0(0)	0(0)
	Labeled	8(2)	0(0)	0(0)	0(0)	92(24)
GH32	Control	0(0)	44(4)	11(1)	0(0)	44(4)
	EPS	0(0)	11(2)	5(1)	5(1)	79(15)
	Labeled	6(2)	14(5)	3(1)	9(3)	69(24)
GH17	Control	75(9)	0(0)	0(0)	0(0)	25(3)
	EPS	43(3)	0(0)	0(0)	0(0)	57(4)
	Labeled	44(8)	0(0)	0(0)	0(0)	56(10)
GH71	Control	0(0)	0(0)	100(1)	0(0)	0(0)
	EPS	0(0)	25(1)	50(2)	0(0)	25(1)
	Labeled	22(5)	0(0)	35(8)	0(0)	43(10)

*average percentage from 4 replicates (total number of sequences).

Within GH families that were more abundant in the SIP "heavy" fraction (Figure 4c), sequences of GH109 belonged mainly go *Acidobacteria* (45% of the GH sequences), other phyla (31-42% of the GH sequences) and *Planctomycetes* (2-29% of the GH sequences). GH117 family sequences belonged predominantly to *Actinobacteria* 17-33% of the sequences, *Acidobacteria* (0-33% of the GH sequences) and other phyla (33-64% of the GH sequences). Family GH50 sequences belonged mainly to *Proteobacteria* (8-100% of the GH sequences) and other phyla (0-92% of the GH sequences). GH 32 sequences were affiliated mainly to *Acidobacteria* (11-44% of the GH sequences) and other phyla (44-79% of the GH sequences). GH17 sequences belonged to phylum *Proteobacteria* (44-75% of the GH sequences) and other phyla (25-57% of the GH sequences). GH71 sequences were affiliated to phyla *Actinobacteria* (35-100% of the GH sequences), *Proteobacteria* (0-25% of the sequences), *Acidobacteria* (0-25% of the sequences) and other phyla (0-43% of the sequences).

3.2.5. Metagenome Assembled Genomes (MAGs) assembled from the metagenome of cultivated microorganisms

The binning process using contigs longer than 5 kb generated, after curation and quality filtering, 4 draft genomes. The genome length ranged from 3.0 to 6.3 Mb and the GC content ranged from 57 to 62%. All MAGs belonged to phylum *Proteobacteria*. None of the MAGs was classified to genus level, however the genomes were closer to genera *Paraburkholderia* (MAG1) and *Amantichitinum* (MAG2 and MAG4). MAG 3 closest classification was to family *Rhodanobacteraceae*. The characteristics of the genomes are described in Table 4. The

coverage of the genomes is described in Table S7.

Table 4: Genome characteristics for the 4 metagenome-assembled genomes (MAGs) obtained in this study.

Genome	MAG1	MAG2	MAG3	MAG4
Taxonomy (closest hit)	*Burkholderiaceae* 95% (*Paraburkholderia* 86%)	*Neisseriaceae* 42% (*Amantichitinum*: 42%)	*Rhodanobacteraceae* 77%	*Neisseriaceae*: 42% (*Amantichitinum* 42%)
Length (Mb)	6.3	3.0	4.8	3.7
Contigs	1997	997	80	1482
Completeness (%)	83.2	79.6	99.7	87.5
Contamination (%)	4.76	3.92	2.44	5
GC(%)	62	57	59	57
Number of predicted genes	7,126	3,580	4,280	4,552
Hits to protein database				
KEGG %	90.1	96.8	79.2	93.8
COG%	85.3	85.1	80.2	84
DBcan n(%)	279 (3.9)	141(3.9)	210 (4.9)	180 (4.0)

Approximately 83.7% of the ORFs predicted for the MAGs could be assigned to COGs. The analysis showed that most of the COG assigned ORFs fell on the category “S-function unknown” (16.4-18.4% of the ORFs). Among the classified COGs, however the most abundant categories were “K-transcription” (5.9-9% of the ORFs), “E-Amino acid metabolism” (4.8-8.1% of the ORFs), “G-Carbohydrate metabolism” (3.32-7.2%), “C-Energy production” (4.2-5.9%), “P-Inorganic ion metabolism” (4.65-6.3%) and “M-cell wall/membrane biogenesis” (5.2-5.9%) (Figure 6a).

KEGG pathway analysis demonstrated that around 90% of the predicted ORFs could be assigned to KEGG orthologs. The majority of the most abundant KEGG orthologs in all the MAGs were related to several types of transporter functions (Figure 6b and Table S8). In order to evaluate the features of the MAGs that could be involved in the uptake of the WH15EPS sugar units, we decided to look deeper into the transporters. Twenty-four of the KEGG orthologs observed in MAG1 genome were associated to the transport of several sugars, such as sorbitol, ribose, arabinose, xylose, fructose, rhamnose, glucose, mannose and multiple sugars (Table S9). Among the KEGG orthologs observed in MAG 2 genome, 62 were related to sugar transport, such as maltose, raffinose, lactose, glucosides, cellobiose, xylose, fructose, rhamnose, glucose, mannose and multiple sugars (Table S10). MAG3 did not exhibit sugar specific transporters within the 60 KEGGs related to transport function, however we observed some general type transporters (Table S11). In MAG4, 61 KEGG orthologs related to sugar transport were observed, such as maltose, raffinose, lactose, sorbitol, cellobiose, arabinose, xylose, fructose, rhamnose, glucose, mannose and multiple sugars (Table S12). We also performed the analysis of the CAZYmes with dbCAN database, in order to find enzymes that could be in associated the breakdown of the WH15EPS. MAG1 possessed 279 CAZymes

distributed in 90 families, of which the most abundant were CE1, GT4, GT42, CE10 and AA3 (Figure 6c). The seventy-six glycoside hydrolases observed were distributed in 43 families, including a wide range of activities, such as endo and exo-mannosidases, alpha and beta glucosidases and galactosidases, xylosidases, fucosidases and rhamnosidases (Table S13). MAG2 possessed 141 CAZymes distributed in 65 families and GT41, GT2 and CE1 were the most abundant families (Figure 6c). A total of 51 glycoside hydrolases from 30 families were observed, with activities such as alpha and beta glucosidases, beta galactosidases, mannanases and mannosidases, xylanases and polygalacturonases (Table S13). In MAG3, 210 cazymes distributed in 81 families were observed, and GT41, GT2, CE1 and CE10 were the most abundant (Figure 6c). Sixty-four glycosil-hidrolases distributed in 37 families were detected. The activities included alpha and beta galactosidases, alpha glucosidases, mannosidases, mannanases, rhamnosidases, arabinosidades, chitinases and trehalases

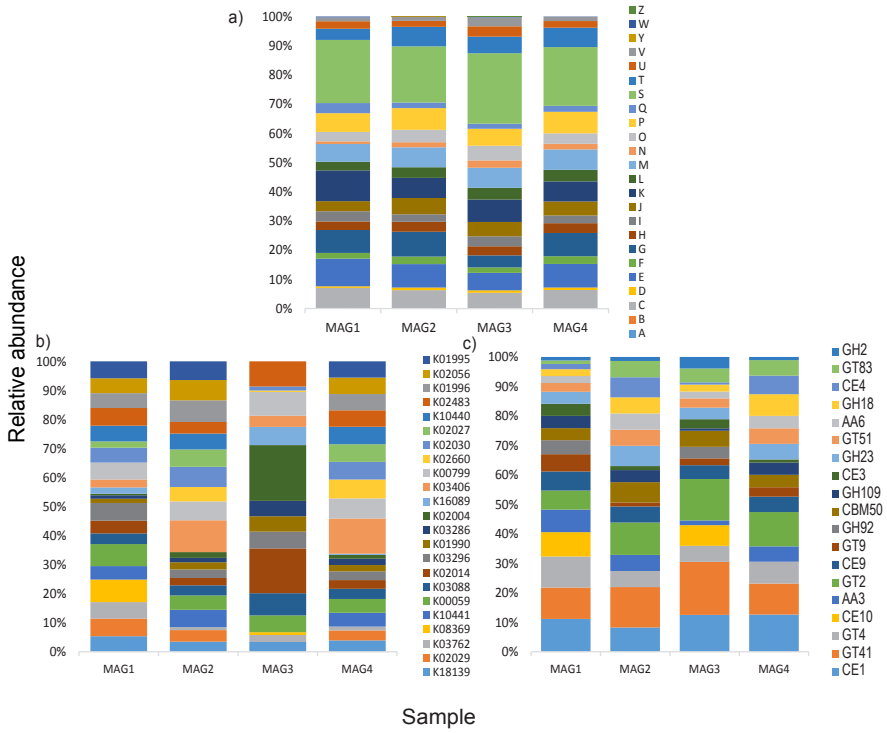

Figure 6: Relative abundance distribution of the most abundant functional categories in metagenome assembled genomes (MAGs) assembled from the shotgun metagenome of cultivated microorganisms sequencing data. a) COG annotation (all categories); b) KEGG annotation (top 10 for each genome); c) dbCAN annotation (top 10 for each genome). The description of the functions displayed in b) and c) are detailed in Table S8 and Table S13, respectively. E-Amino acid transport and metabolism; G- Carbohydrate transport and metabolism; H-Coenzyme transport and metabolism; C-Energy production and conversion; I-Lipid transport and metabolism; F-Nucleotide transport and metabolism; Q- Secondary metabolites; D-Cell cycle; N-Cell motility; M-Cell wall/membrane/envelope biogenesis; V-Defence mechanisms; P-Inorganic ion transport and metabolism; U-Intracellular trafficking; O-Post translational modification; T-Signal transduction mechanisms; L-Replication, recombination and repair; K-Transcription; J-Translation; S-Function unknown; R-General function and prediction; X-Mobilome.

(Table S13). The genome of MAG4 displayed 180 CAZymes distributed in 73 families, of which the most abundant were CE1, GT2 and GT41 (Figure 6c). The 64 glycoside hidrolases were spread among 34 families, including activities such as chitinases, arabinofuranosidases, alpha and beta glycosidases, mannosidases, cellulases, xylanases and polygaracturonases (Table S13). The distribution of most abundant CAZYmes and GH families in both metagenomics datasets and MAGs is depicted in Figure S8.

4. Discussion

In the present study, we applied culture-independent and culture-dependent techniques to evaluate microbial diversity and functions involved in the degradation of a microbial biopolymer, WH15EPS, focusing on enzymes of biotechnological interest. First, we compared the functional potential of the environment with and without the presence of WH15EPS, evaluating the taxonomic and functional enrichment induced by the addition of the biopolymer. To this end we used Stable Isotope Probing (SIP), which was followed by metagenomics to evaluate the functional potential of the microorganisms grown in culture medium with WH15EPS as the sole carbon source.

It was already observed that specific functions and activities can be selectively targeted through enrichment and specific techniques such as SIP. Enrichment with specific substrates allows the manipulation of the local microbial community prior to the metagenomic DNA extraction, increasing the prevalence of target functions (Ekkers *et al.*, 2011). For instance, Verastegui *et al.* (2014) incubated five ^{13}C labeled plant-derived carbon substrates (glucose, cellobiose, xylose, arabinose and cellulose) to diverse soils for the characterization of active soil bacterial communities and their glycosyl hydrolases. Furthermore, SIP has been used for bioremediation studies, involving the breakdown and metabolization of compounds such as methanol (Ginige *et al.*, 2004), phenol (Padmanabhan *et al.*, 2003) and biphenyl (Lee *et al.*, 2011).

SIP analysis demonstrated that in both SSU rRNA and ORF-based characterization, the phyla *Proteobacteria*, *Actinobacteria*, *Acidobacteria* and *Planctomycetes* were the most abundant in WH15EPS amended and unamended treatments. However, the addition of WH15EPS to the litter samples promoted an increase in the abundance of the phylum *Planctomycetes*, which was more evident in 'heavy' fraction samples, showing that *Planctomycetes* also play an active part in the degradation of WH15EPS. Furthermore, at genus level in the SSU rRNA based analysis, "unclassified *Planctomycetes*" and *Singulisphaera*, which belong to the same phylum, are the most abundant groups in the labeled treatment, while "unclassified *Planctomycetes*" is also among the most abundant in the ORF based analysis.

Proteobacteria, *Actinobacteria* and *Acidobacteria* are widely known to be involved in carbon-degradation processes. Several *Proteobacteria*, such as *Dyella*, *Mesorhizobium* and

Sphingomonas have been linked to glucose (Pinnell *et al.*, 2014) and cellulose assimilation (Haichar *et al.*, 2007), while the genus *Rhodanobacter* was implicated in the degradation of aromatic compounds (Song *et al.*, 2016). Members of *Actinobacteria* were already associated with cellulose, cellobiose and glucose hydrolysis (Schellenberger *et al.*, 2009, Bao *et al.*, 2019). *Acidobacteria* have mainly been linked to the degradation of hemicellulose, especially xylan, and genomic analyses showed the presence of genes associated to the degradation of a wide variety of polysaccharides (de Castro *et al.*, 2013, Kielak *et al.*, 2016). The glycolytic potential of the phylum *Planctomycetes* was demonstrated recently by Ivanova et al. (2017), using transcriptomics to evaluate their response to cellulose, xylan, pectin and chitin. The authors observed that each polymer induced an increase in abundance of transcripts belonging to different genera. Transcripts belonging to genus *Singulisphaera*, for instance, had their abundance increased significantly in response to pectin and xylan amendments.

The cultivation-dependent approach demonstrated, as expected, a lower taxonomic diversity, in which the widely studied *Proteobacteria* were among the most abundant. However, we also observed a wide diversity of unclassified bacteria and other unclassified microorganisms in both SSU and ORF-based analyses. The discrepancy between the diversity of taxa, especially the most abundant groups, observed in cultured and uncultured-based techniques is defined as “The Great Plate Count Anomaly” (Staley & Konopka, 1985). The cultivability of microorganisms in laboratory depends of many factors, such as nutrients, oxygen level, temperature, pH and growing factors (Vester *et al.*, 2015), limiting the total assortment of taxa that can be actually recovered in culture media. Nevertheless, adding WH15EPS as an alternative carbon source allowed us to demonstrate that several still unknown microorganisms can be grown in laboratorial conditions if unusual compounds are explored. The lower diversity in the culture media plates permitted the assembly of 4 draft genomes related to the most abundant *Proteobacteria*, which classification until genus level was not possible, once more demonstrating the enrichment and potential for isolation of previously unknown microbes.

In order to find potential enzymes of biotechnological interest we investigated the diversity of CAZymes in both culture-independent and culture-dependent generated datasets, due to their importance in almost all industrial sectors, such as chemical, pharmaceutical and food industries, as well as production of detergents, textiles, leather, paper and bioenergy (Berini *et al.*, 2017). Furthermore, we also investigated the presence of enzymes that could be employed for biofilm removal.

Among all CAZymes observed, the most abundant families belonged to Glycoside Transferase families, such as GT41, GT2 and GT4, either in culture based or in culture independent datasets. GTs are known to catalyze the formation of glycosidic bonds by transferring a sugar residue from a donor to an acceptor, which could be carbohydrates, proteins lipids, DNA and other molecules (Schmid *et al.*, 2016). Even though a large proportion of genes of microorganism`s genomes in general encode for GTs (abour 1-2% of the total number of genes) (Lairson *et al.*,

2008), those enzymes are still not as well explored as GHs (Schmid *et al.*, 2016). Glycosilated compounds play a wide range of roles, such as energy storage, cell integrity and signaling, among others, and the glycosilation of natural products is important in the exploration of bioactive compounds (Liang *et al.*, 2015). GTs are involved in the production of antibiotics, such as chloroeremomycin (Mulichak *et al.*, 2003), vancomycin (Mulichak *et al.*, 2004) and erythromycin D (Moncrieffe *et al.*, 2012), therefore they might be of interest especially for the pharmaceutical industry.

Within glycoside hydrolases, the most abundant family in both metagenomics datasets was GH13 (from *Proteobacteria*), which encompasses starch and pullulan modifying enzymes, including α-amylases, pullulanases, α-1,6-glucosidases, branching enzymes, maltogenic amylases, neopullulanases, and cyclodextrinases (Labes *et al.*, 2008). Amylases are among the most important enzymes for food industry, where they can be employed for production of glucose and maltose syrups, reduction of viscosity of syrups, production of clarified fruit juices, solubilization of starch for brewing processes and manufacture of baked products (Liu & Kokare, 2017). Furthermore, the application of α-amylases for the inhibition of biofilm formation has been investigated. In the study of Fleming *et al*. (2016) the use of amylase (from *Bacillus subtilis*) and cellulose (from *Aspergillus niger*) solutions to biofilms of *S. aureus* and *P. aeruginosa* decreased biomass significantly, increasing the effectiveness of antibiotics treatments. A similar effect was observed in the study of Craigen *et al*. (2011), where a commercially available α-amylase detached the aggregates produced by *S. aureus* and inhibited biofilm production.

Notwithstanding, feature selection with Boruta package revealed the differential abundance of GH families in "heavy" fraction SIP samples, originated from microorganisms that are believed to be able to degrade WH15EPS. These microorganisms belonged mainly to phyla *Proteobacteria*, *Acidobacteria*, *Actinobacteria*, *Planctomycetes*, as well as high proportion of unknown microorganisms. GH109 (*Acidobacteria* and *Planctomycetes*) contains α-N-acetylgalactosaminidases, which might be employed in the development of universal red blood cells, through the enzymatic removal of monosaccharides from red blood cells' membranes, and improvement of blood supply in hospitals (Liu *et al.*, 2007). Furthermore, those enzymes can be involved in the deconstruction of WH15EPS, since it contains units of xylose, glucose and arabinose (Kielak *et al.*, 2017). Families GH117 (*Acidobacteria* and *Actinobacteria*) and GH50 (*Proteobacteria*) contain agarases, which can be used for the production of oligosaccharides with antioxidant activities for applications in food, pharmaceutical and cosmetic industries (Fu & Kim, 2010). Family GH32 (*Acidobacteria*) comprises invertases and inulinases, enzymes that can be applied in food and fermentation processes (Khan *et al.*, 2013, Mohan *et al.*, 2018). GH17 (*Proteobacteria*) is composed of endoglucanases with activity against β-glucan and laminarin, effective additives for the degradation of polysaccharides for animal feed (Mohan *et al.*, 2018). Mutanases belonging to GH71 (*Actinobacteria*) family already showed activity against glucans present in dental

plaque (Wiater *et al.*, 2005).
Interestingly, sixteen of the most abundant GH families in the culture-independent dataset were found to be the predominant in the culture-dependent approach, and all the GH families with higher abundances in the labeled SIP samples were also observed in the culture-dependent dataset. Furthermore, the MAGs also contained GH families of interest. MAG 1 (similar to *Paraburkholderia*) contained 8 ORFs belonging to family GH92, which encompasses alpha mannosidases with applications in food and pharmaceultical industries, for the production of juices, degradation of plant material or coffee extraction (Konan *et al.*, 2016). In MAG2 (similar to *Amantichitinum*), five ORFs were classified as GH23, which contains lysozymes that can be used as polysaccharide hydrolysers for biofilm breakdown (Hukić *et al.*, 2018, Nahar *et al.*, 2018). MAG3 (*Rhodanobacteraceae*) is abundant in GH92 and GH23 but also GH2 family ORFs, which comprises several enzymes. Within the best characterized ones there are β-galactosidases employed for the production of lactose-free milk products and other galactooligosaccharides (Mallela *et al.*, 2016). MAG4 (similar to *Amantichitinum*) is rich in GH18 enzymes, envolving chitinases, that, for instance are important agents with applications for fungal biological control and bioremediation processes (Dahiya *et al.*, 2005). Our study showed that, using SIP and a complex EPS (WH15EPS), we could detect the subset of the total microbial community that was capable of incorporating the biopolymer. Among those we observed members of *Planctomycetes* as an interesting target for biotechnological studies and heterologous expression. In addition, we demonstrated that functional diversity induced by the presence of WH15EPS in both culture-dependent and culture independent approaches was enriched in genes coding for GHs, for instance, amylases, chitinases, agarases and endoglucanases and that could be applied in chemical, pharmaceutical and food industries. Furthermore, the use of WH15EPS may be employed for the investigation and isolation of yet unknown taxa, such as unclassified *Proteobacteria* and *Planctomycetes*, increasing the number of current cultured bacterial representatives.

Acknowledgements

We would like to thank Wietse de Boer for helping with the sampling local. O.Y.A. Costa was supported by an SWB grant from CNPq [202496/2015-5] (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

Supplementary material

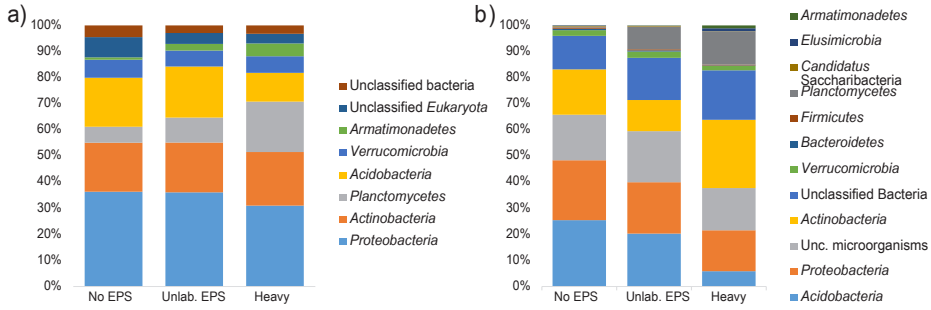

Figure S1: Taxonomic composition and relative abundance of microbial groups at phylym level in SIP metagenome treatments based on a) SSU rRNA gene sequence classification (>2.2 % abundance) b) ORF taxonomic classification (>0.1% abundance). Average abundances of 4 replicates. Unc: unclassified. No EPS – incubation without WH15EPS. Unlab EPS-incubation containing ^{12}C-WH15EPS. Heavy – 'heavy fraction' of incubations containing ^{13}C-WH15EPS.

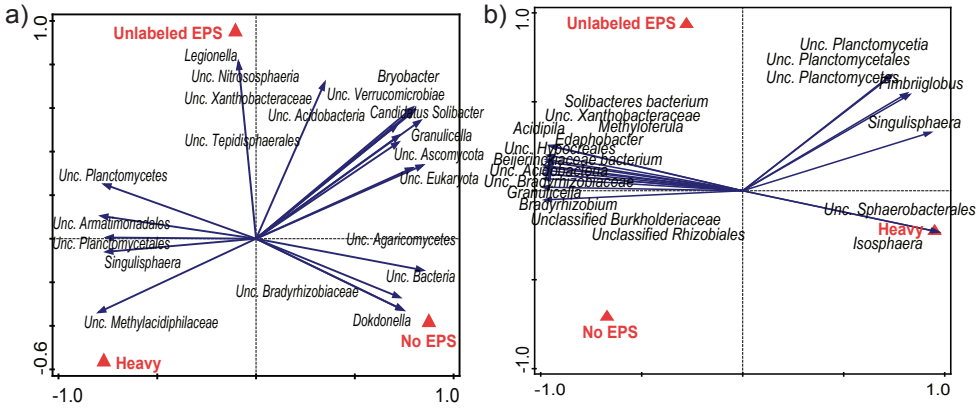

Figure S2: Biplot of the Redundancy analysis (RDA) based on normalized and Hellinger-transformed abundances of a) SSU rRNA gene taxonomy classification and b) ORF taxonomic classification. Only the best 20 fitting groups are displayed. Unc: unclassified. No EPS – incubation without WH15EPS. Unlab EPS-incubation containing ^{12}C-WH15EPS. Heavy – 'heavy fraction' of incubations containing ^{13}C-WH15EP.S

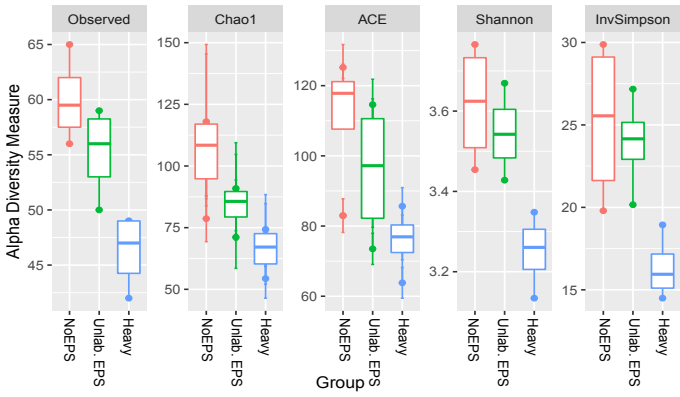

Figure S3: Box-plot comparisons of alpha-diversity assessment by richness estimators (number of observed OTUs, Chao1, ACE) and diversity indices (Shannon, Inverse Simpson) for SIP SSU rRNA gene samples. 'Heavy fraction' values are significantly lower in comparison with both controls for all comparisons (p-value < 0.05). Comparisons performed across treatments using ANOVA test and Tukey`s HSD post-hoc test. Data rarefied to the minimum sampling depth. Unlab. EPS-incubation containing ^{12}C-WH15EPS. Heavy – 'heavy fraction' of incubations containing ^{13}C-WH15EPS.

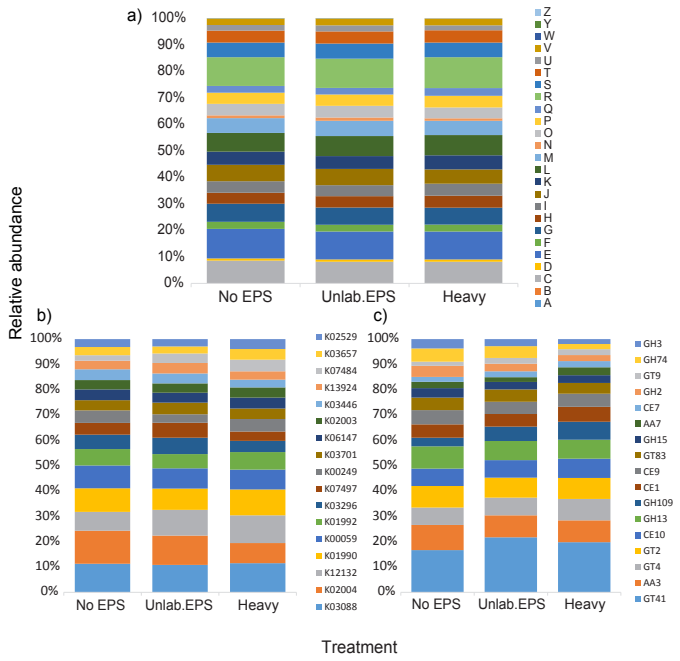

Figure S4: Relative abundance distribution of the most abundant functional categories in TPM-normalized metagenome sequencing data from the SIP metagenome. a) COG annotation (all categories); b) KEGG annotation (above 0.1 % abundance); c) dbCAN annotation (above 1% abundance). E-Amino acid transport and metabolism; G- Carbohydrate transport and metabolism; H-Coenzyme transport and metabolism; C-Energy production and conversion; I-Lipid transport and metabolism; F-Nucleotide transport and metabolism; Q- Secondary metabolites; D-Cell acycle; N-Cell motility; M-Cell wall/membrane/envelope biogenesis; V-Defence mechanisms; P-Inorganic ion transport and metabolism; U-Intracellular trafficking; O-Post translational modification; T-Signal transduction mechanisms; L-Replication, recombination and repair; K-Transcription; J-Translation; S-Function unknown; R-General function and prediction; X-Mobilome.

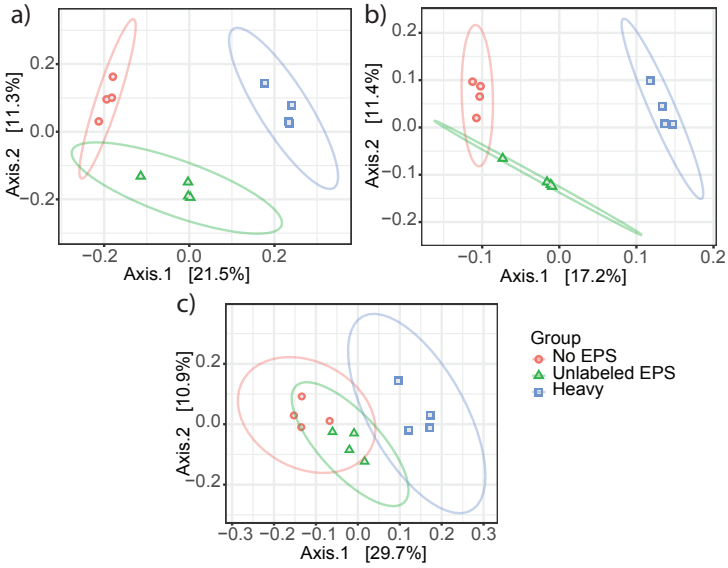

Figure S5: Principal Coordinate Analysis (PCoA) clustering of normalized and Hellinger-transformed SIP metagenome sequencing data based on Bray-Curtis distances of a) COG annotation, b) KEGG annotation and c) dbCAN annotation. No EPS – incubation without WH15EPS. Unlabeled EPS-incubation containing ^{12}C-WH15EPS. Heavy – 'heavy fraction' of incubations containing ^{13}C-WH15EPS.

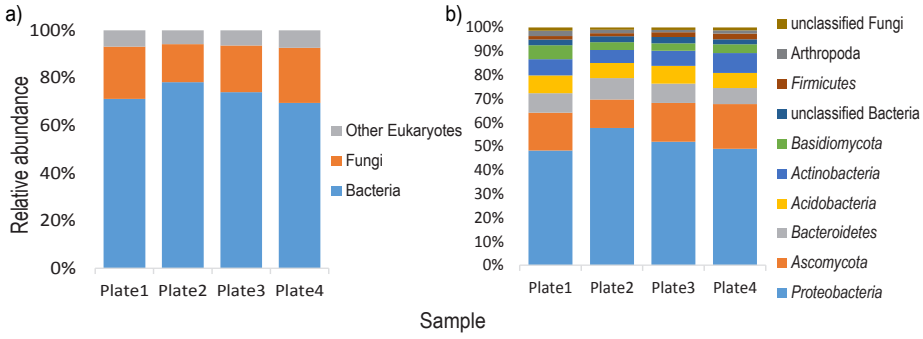

Figure S6: Taxonomic composition and relative abundance of microbial groups at a) kingdom and b) phylum level in samples from the metagenome shotgun of cultivated microrganims based on SSU rRNA gene taxonomic classification. Average from 2 replicates per plate of culture medium.

Figure S7: Venn diagram depicting the number of common and unique glycoside hydrolase (GH) families observed in SIP metagenome and metagenome of cultivate microorganisms` datasets.

a)

	Un.Control	EPS	Heavy
GT41	8.4%	11.1%	11.0%
AA3	5.0%	4.4%	4.8%
GT4	3.4%	3.6%	4.7%
GT2	4.3%	4.1%	4.6%
CE10	3.5%	3.6%	4.2%
GH13	4.4%	3.9%	4.2%
GH109	1.7%	2.9%	3.9%
CE1	2.6%	2.6%	3.3%
CE9	2.8%	2.4%	2.9%
GT83	2.5%	2.5%	2.4%
GH15	1.9%	1.5%	1.7%
AA7	1.3%	1.0%	1.7%
CE7	1.0%	1.2%	1.4%
GH2	2.2%	1.5%	1.3%
GT9	0.8%	1.2%	1.3%
CE14	1.0%	0.8%	1.3%
GH74	2.6%	2.4%	1.2%
GH3	1.9%	1.4%	1.1%
GT35	1.2%	0.9%	1.0%
GH33	0.1%	0.8%	1.0%

b)

	Plate1	Plate2	Plate3	Plate4
GT41	6.5%	6.9%	7.1%	6.9%
AA3	6.1%	6.1%	6.4%	6.8%
GT4	3.9%	4.3%	4.5%	4.2%
CE1	3.8%	3.9%	3.8%	3.7%
GT2	3.6%	4.2%	4.3%	3.8%
CE10	3.6%	3.5%	3.6%	4.0%
GH13	3.2%	3.0%	3.0%	3.0%
CE9	2.8%	2.9%	3.0%	3.0%
GH92	2.3%	2.4%	2.0%	1.9%
GT9	2.1%	2.0%	1.9%	1.8%
GH2	2.1%	2.2%	2.1%	2.0%
GT83	2.0%	1.8%	1.7%	1.7%
GH3	1.8%	1.7%	1.7%	1.7%
GH109	1.6%	1.7%	1.8%	1.6%
GH106	1.6%	1.4%	1.2%	1.2%
GH31	1.5%	1.4%	1.3%	1.3%
GH29	1.4%	1.4%	1.3%	1.1%
GH78	1.3%	1.1%	1.1%	1.1%
AA7	1.3%	1.3%	1.4%	1.6%
GH23	1.3%	1.5%	1.5%	1.3%

c)

	MAG1	MAG2	MAG3	MAG4
GT41	18	10	23	10
CE1	19	6	16	12
GT2	11	8	18	11
GT4	18	4	7	7
CE9	11	4	6	5
AA3	13	4	2	5
CBM50	7	5	7	4
CE10	14	0	9	0
GH23	7	5	5	5
GH18	4	4	3	7
GT51	5	4	4	5
GT83	2	4	6	5
GT9	10	1	3	3
CE4	3	5	1	6
AA3_1	6	3	2	4
AA6	4	4	3	4
GH109	7	3	1	4
GH3	5	3	4	3
CE3	7	1	4	1
GH92	8	0	5	0

d)

	Un.Control	EPS	Heavy
GH13	4.4%	3.9%	4.2%
GH109	1.7%	2.9%	3.9%
GH15	1.9%	1.5%	1.7%
GH2	2.2%	1.5%	1.3%
GH74	2.6%	2.4%	1.2%
GH3	1.9%	1.4%	1.1%
GH23	0.9%	1.1%	0.9%
GH78	0.9%	0.9%	0.8%
GH29	0.8%	1.0%	0.7%
GH92	0.8%	0.7%	0.7%
GH1	0.8%	0.7%	0.6%
GH28	1.1%	1.0%	0.5%
GH31	1.1%	1.0%	0.4%
GH39	0.9%	0.7%	0.4%
GH51	1.0%	0.7%	0.4%
GH106	1.1%	0.8%	0.4%
GH43	0.7%	0.6%	0.4%
GH127	0.8%	0.6%	0.4%
GH27	0.8%	0.9%	0.2%
GH144	0.7%	0.4%	0.2%

e)

	Plate1	Plate2	Plate3	Plate4
GH13	3.2%	3.0%	3.0%	3.0%
GH92	2.3%	2.4%	2.0%	1.9%
GH2	2.1%	2.2%	2.1%	2.0%
GH3	1.8%	1.7%	1.7%	1.7%
GH109	1.6%	1.7%	1.8%	1.6%
GH106	1.6%	1.4%	1.2%	1.2%
GH31	1.5%	1.4%	1.3%	1.3%
GH29	1.4%	1.4%	1.3%	1.1%
GH78	1.3%	1.1%	1.1%	1.1%
GH23	1.3%	1.5%	1.5%	1.3%
GH15	1.0%	0.9%	0.9%	0.9%
GH94	0.9%	1.2%	0.9%	0.9%
GH35	0.8%	0.8%	0.7%	0.7%
GH42	0.8%	0.7%	0.7%	0.6%
GH27	0.7%	0.9%	0.8%	0.7%
GH28	0.7%	0.8%	0.7%	0.7%
GH20	0.7%	0.8%	0.8%	0.7%
GH43	0.7%	0.6%	0.7%	0.8%
GH95	0.7%	0.7%	0.7%	0.6%
GH18	0.7%	0.7%	0.7%	0.7%

f)

	MAG1	MAG2	MAG3	MAG4
GH92	8	0	5	0
GH23	7	5	5	5
GH109	7	3	1	4
GH3	5	3	4	3
GH18	4	4	3	7
GH15	3	0	0	0
GH94	2	3	0	2
GH2	2	1	5	1
GH27	2	0	1	0
GH36	2	0	1	0
GH35	1	3	2	3
GH31	1	1	2	2
GH1	1	3	0	2
GH73	1	1	2	1
GH20	0	2	3	3
GH13_23	0	3	1	3
GH53	0	0	0	3
GH125	0	1	1	2
GH128	0	2	0	2
GH29	0	0	3	0

Figure S8: Distribution of the 20 most abundant CAZyme families in a) SIP metagenome samples (relative abundance, average of 4 replicates); b) Metagenome of cultivated microorganisms (relative abundance, average of 2 replicates); c) Metagenome-Assembled Genomes (MAGs) (number of genes), and most abundant glycosyl hydrolases (GH) in d) SIP metagenome samples (relative abundance, average of 4 replicates), e) metagenome of cultivated microorganisms (relative abundance, average of 2 replicates) and f) Metagenome-Assembled Genomes (MAGs) (number of genes).

Table S1: Coordinates of the sampling sites.

Site number	Latitude	Longitude
1	51°59'14.5"N	5°47'32.7"E
2	51°59'15.9"N	5°47'29.5"E
3	51°59'15.7"N	5°47'27.7"E
4	51°59'14.8"N	5°47'23.2"E

Table S2: Physicochemical properties of topsoil-litter samples.

Component	Unit	Average (Sd)	Component	Unit	Average (Sd)
total N	mg N/kg	16535±3217	**CEC**	%	81±0.0
C/N ratio		20±6	**CEC**	mmol+/kg	214±42
Available N	kg N/ha	252±187	**B**	µg B/kg	488±4.2
pH		3.05±0.1	**Cu**	µg Cu/kg	48±9.9
OM	%	55.8±3.5	**Fe**	µg Fe/kg	3070±466.7
Na	mg Na/kg	34.5±4.9	**Mn**	µg Mn/kg	101920±9362.1
P	mg P/kg	42.65±6.4	**Zn**	µg Zn/kg	8860±693.0
K	mg K/k	218±11.3	**Clay**	%	5.5±0.7
Ca	kg Ca/ha	13±0	**Silt**	%	10±14.1
Mg	mg Mg/kg	175±7.1	**Sand**	%	27.5±12.0

Sd: standard deviation

Table S3: COG functions that significantly segregated across treatments selected by Boruta random forests algorithm based on 1000 permutations in the SIP metagenome treatment comparisons (p-value <0.05).

COG	Function		Category
COG0043*	3-polyprenyl-4-hydroxybenzoate decarboxylase	H	Coenzyme transport and metabolism
COG2985*	transport protein/Uncharacterized membrane protein YbjL	P	Inorganic ion transport and metabolism
COG3492*	Deoxycytidine triphosphate deaminase	S	Function unknown
COG3534	Alpha-N-arabinofuranosidase (EC 3.2.1.55	G	G Carbohydrate transport and metabolism
COG3547*	Transposase	L	Replication, recombination and repair
ENOG410XNWR	ABC transporter substrate-binding protein	E	Amino acid transport and metabolism
ENOG410XQPR*	Sulfite exporter TauE/SafE	S	Function unknown
ENOG410XQRG	Transcriptional regulator	K	Transcription
ENOG410XRW3*	secreted protein	S	Function unknown
ENOG410Y9A5*	Cna B-type protein	S	Function unknown
ENOG410YCZ9*	NA	S	Function unknown
ENOG410ZNCC*	NA	S	Function unknown
ENOG4111SDB*	NA	S	Function unknown
COG0665	Glycine/D-amino acid oxidase	E	Amino acid transport and metabolism
COG0745	OmpR -regulatoR	T	Signal transduction mechanisms
COG1309	Transcriptional regulator	K	Transcription
COG2243	Precorrin-2 methylase	H	Coenzyme transport and metabolism
ENOG410XPHN*	NA	S	Function unknown
ENOG410XQ2Q*	sialic acid-specific 9-O-acetylesterase	S	Function unknown
ENOG410XQ46	Protein of unknown function (DUF1549) (*Planctomycetes*)	S	Function unknown
ENOG410XQWX*	NA	S	Function unknown
ENOG410XQXZ	NA	S	Function unknown
ENOG410XRSU	NA (*Acidobacteria*)	S	Function unknown
ENOG410XZSR	NA (*Acidobacteria*)	S	Function unknown
ENOG410Y0YH	Reductase	I	Lipid transport and metabolism
ENOG410ZWC6	Kelch repeat-containing protein	S	Function unknown
ENOG4111HEH	(ABC) transporter	S	Function unknown
ENOG4111HR8	Tetr family transcriptional regulator	K	Transcription

COG (continued)	Function		Category
ENOG4111R18	integral membrane transport protein	P	Inorganic ion transport and metabolism
ENOG4111TA9	NA	S	Function unknown
ENOG41120VN	NA (*Planctomycetes*)	S	Function unknown
ENOG41126W5	NA	S	Function unknown

*Higher abundance in ^{12}C-EPS-amended control samples. Blue – more abundant in unamended control samples. Orange: more abundant in Labeled samples

Table S4: KEGG orthologs that significantly segregated across treatments selected by Boruta random forests algorithm based on 1000 permutations in the SIP metagenome treatment comparisons (p-value <0.05).

KEGG	Annotation	KEGG Metabolic Pathway
K00467	Lactate 2-monooxygenase	Metabolic pathways/Pyruvate
K01811*	Alpha-D-xyloside xylohydrolase	-
K01999*	Branched-chain amino acid transport system substrate-binding protein	ABC transporters/quorum sensing
K02083	Allantoate deiminase	Metabolic pathways/Microbial metabolism/ Purines
K02274*	Cytochrome c oxidase subunit I	Metabolic pathways/Oxidative phosphorilation
K03074	Preprotein translocase subunit secf	Protein export/Bacterial secretion
K03307*	Solute:Na+ symporter, SSS family	-
K03832*	Periplasmic protein tonb	-
K07263*	Zinc protease	-
K07795	Putative tricarboxylic transport membrane protein	Two-component system
K08964	Methylthioribulose-1-phosphate dehydratase	Metabolic pathways/Amino-acids
K09758*	Aspartate 4-decarboxylase	Metabolic pathways/Amino-acids
K10826	Putative ABC transporter transmembrane protein	-
K18900	Lysr family transcriptional regulator, regulator for bpeef and oprc	-
K00200*	Formylmethanofuran dehydrogenase subunit A	Metabolic pathways/microbial metabolism/ carbon metabolism/methane metabolism
K00249	Acyl-coA dehydrogenase	Antibiotics/fatty acids/secondary metabolites
K00477	Phytanoyl-coA hydroxylase	Peroxysome
K01087	Trehalose 6-phosphate phosphatase	Metabolic pathways/starch sucrose met
K01133	Choline-sulfatase	-
K01258	Tripeptide aminopeptidase	-
K01515	ADP-ribose pyrophosphatase	Metabolic pathways/Purine metabolism
K01992	ABC-2 type transport system permease protein	-
K02551	2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase	Metabolic pathways/Ubiquinone biosynthesis
K02657	Twitching motility two-component system response regulator pilG	Two component system/Biofilm
K03075	Preprotein translocase subunit secG	Bacterial secretion/Protein export
K03167*	DNA topoisomerase VI subunit B	-
K03620	Ni/Fe-hydrogenase 1 B-type cytochrome subunit	Two componente system
K05786	Chloramphenicol-sensitive protein rarD	-
K05970*	Sialate O-acetylesterase	-
K06606*	2-keto-myo-inositol isomerase	Metabolic pathways/microbial metabolism/ inositol metabolism
K08282	Non-specific serine/threonine protein kinase	-
K08602*	Oligoendopeptidase F	-
K09684	Pucr family transcriptional regulator, purine catabolism regulatory protein	-
K10301*	F-box protein 21	-
K10714*	Methylene-tetrahydromethanopterin dehydrogenase	Metabolic pathways/ Microbial metabolism/ Carbon metabolism/ Methane metabolism

KEGG (continued)	Annotation	KEGG Metabolic Pathway
K11708*	Manganese/zinc/iron transport system permease protein	ABC transporters
K15342	CRISP-associated protein Cas1	-
K17329	N,N'-diacetylchitobiose transport system substrate-binding protein	ABC transporters
K19169*	DNA sulfur modification protein dndB	-
K19542	MFS transporter, DHA2 family, tetracycline/ oxytetracycline resistance protein	-

*Higher abundance in EPS amended samples. Blue – more abundant in unamended control samples. Orange: more abundant in Labeled samples

Table S5: CAZyme families that significantly segregated across treatments selected by Boruta random forests algorithm based on 1000 permutations in the SIP metagenome treatment comparisons (p-value <0.05).

CAZY family	Function
AA1*	Multicopper oxidases
CE8*	Pectin methylesterase
GH18*	Chitinase
GH2	β-galactosidase
GH27*	α-galactosidase
GH28*	Pectin polygalacturonase
GH3	β-glucosidase/xylosidase
GH31*	α-glucosidases
GH36	α-galactosidase/N-acetylgalactosaminidase
GH42*	β-galactosidases
GH51	Hemicellulases
GH79	Proteoglycanases
GH115*	Xylan α-glucuronidase
CBM66	Fructan-binding modules
CBM9*	Xylan-binding modules
CE6	Acetyl xylan esterase
GH109	α-N-acetylgalactosaminidase
GH117	α-1,3-L-neoagarooligosaccharide hydrolase
GH17	1,3;1,4-β-D-glucan endohydrolases
GH32	Invertases
GH33*	Sialidases
GH50	β-agarase
GH71	α-1,3-glucanase
GT12*	[N-acetylneuraminyl]-galactosylglucosylceramide N-acetylgalactosaminyltransferase
GT4	Mannosyltransferases/N-acetylglucosaminyltransferases
GT87	Polyprenol-P-Man: α-1,2-mannosyltransferase
PL10*	Pectate lyase

AA-auxiliary activity; CE-carbohydrate esterase; GH-glycoside hydrolase; GT-glycoside transferase; PL-polysaccharide lyases. *Higher abundance in EPS-amendedsamples. Blue – more abundant in unamended control samples. Orange: more abundant in Labeled samples

Table S6: Most abundant CAZyme families (above 1% abundance) and most abundant KEGG orthologs (above 0.2% abundance) in the shotgun metagenome of cultivated microorganisms.

Category	Ave abundance (%)	Associated functions
CAZy family		
GT41	6.83%	UDP-GlcNAc: peptide β-N-acetylglucosaminyltransferase; UDP-Glc: peptide N-β-glucosyltransferase
AA3	6.35%	Glucose-methanol-choline (GMC) family of oxidoreductases,
GT4	4.24%	Mannosyltransferases/ N-acetylglucosaminyltransferases
CE1	3.78%	Large variety of substrates

Category (continued)	Ave abundance (%)	Associated functions
GT2	3.96%	Cellulose synthase, chitin synthase, mannosyltransferase, glucosyltransferase, galactosyltransferase, rhamnosyltransferase
CE10	3.67%	Esterases acting on non-carbohydrate substrates
GH13	3.05%	α-glucosidase
CE9	2.93%	Deacetylation of N-acetylglucosamine-6-phosphate to glucosamine-6-phosphate
GH92	2.15%	α-mannosidase
GT9	1.97%	Lipopolysaccharide N-acetylglucosaminyltransferase/Heptosyltransferase
GH2	2.07%	β-galactosidase
GT83	1.80%	Undecaprenyl phospho-α-L-4-amino-4-deoxy-L-arabinose; dodecaprenyl phospho-β-galacturonic acid:lipopolysaccharide core α-galacturonosyl transferase
GH3	1.71%	β-glucosidases/galactosidases/xylosidases
GH109	1.69%	α-N-acetylgalactosaminidase
GH106	1.34%	α-L-rhamnosidase/rhamnogalacturonan α-L-rhamnohydrolase
GH31	1.40%	α-xylosidase/glucosidase/mannosidase
GH29	1.29%	α-fucosidase
GH78	1.16%	α-L-rhamnosidase/rhamnogalacturonan α-L-rhamnohydrolase
AA7	1.40%	Gluco/chito-oligosaccharide oxidase
GH23	1.39%	Peptidoglycan lyase
CBM50	1.02%	Bind to the N-acetylglucosamine residues in bacterial peptidoglycans and in chitin
GT51	1.08%	Murein polymerase
CE4	1.08%	De-acylation of polysaccharides
KEGG		
K03296	0,50%	Hydrophobic/amphiphilic exporter-1 (mainly G- bacteria), HAE1 family
K02014	0,50%	Iron complex outermembrane recepter protein
K02004	0,46%	Putative ABC transport system permease protein
K00059	0,41%	3-oxoacyl-[acyl-carrier protein] reductase
K03406	0,32%	Methyl-accepting chemotaxis protein
K01995	0,29%	Branched-chain amino acid transport system ATP-binding protein
K00128	0,28%	Aldehyde dehydrogenase (NAD+)
K18139	0,27%	Outer membrane protein, multidrug efflux system
K01996	0,27%	Branched-chain amino acid transport system ATP-binding protein
K18138	0,26%	Multidrug resistance, efflux pump MexAB-OprM
K08191	0,26%	MFS transporter, ACS family, hexuronate transporter
K02030	0,25%	Polar amino acid transport system substrate-binding protein
K02056	0,25%	Simple sugar transport system ATP-binding protein
K00249	0,25%	Acyl-CoA dehydrogenase
K03088	0,24%	RNA polymerase sigma-70 factor, ECF subfamily
K03762	0,24%	MFS transporter, MHS family, proline/betaine transporter
K01897	0,24%	Long-chain acyl-CoA synthetase
K02483	0,23%	Two-component system, OmpR family, response regulator
K01990	0,22%	ABC-2 type transport system ATP-binding protein
K08369	0,21%	MFS transporter, putative metabolite:H+ symporter
K03296	0,50%	Hydrophobic/amphiphilic exporter-1 (mainly G- bacteria), HAE1 family

Ave: average; AA-auxiliary activity; CE-carbohydrate esterase; GH-glycoside hydrolase; GT-glycoside transferase; PL-polysaccharide lyases

Table S7: MAGs coverage in all samples.

Sample	MAG1	MAG2	MAG3	MAG4
OCP1	160.7	18.9	1805.1	13.2
OCP2	126.8	29.4	1285.5	20.6
OCP3	252.2	81.7	1826.3	51.1
OCP4	119.9	58.3	1348.9	45.5
OCP5	191.6	0.0	1715.5	0.0
OCP6	127.2	48.0	1443.8	40.0
OCP7	150.3	0.0	1376.4	0.0
OCP8	178.4	90.3	1149.0	54.5

Table S8: Most abundant KEGG orthologs in MAGs and their associated functions. A selection of the top 10 most abundant KEGG orthologs in each genome is displayed.

KEGG	Associated function	Category
K18139	Outer membrane protein, multidrug efflux system	Transporters
K02029	Polar amino acid transport system permease protein	Transporters
K03762	MFS transporter, MHS family, proline/betaine transporter	Transporters
K08369	MFS transporter, putative metabolite:H+ symporter	Transporters
K10441	ribose transport system ATP-binding protein	Transporters
K00059	3-oxoacyl-[acyl-carrier protein] reductase	Lipid metabolism
K03088	RpoE	Transcription machinery
K02014	Iron complex outermembrane recepter protein	Transporters
K03296	Hydrophobic/amphiphilic exporter-1 (mainly G- bacteria), HAE1 family	Transporters
K01990	ABC-2 type transport system ATP-binding protein	Transporters
K03286	OmpA-OmpF porin, OOP family	Transporters
K02004	Putative ABC transport system permease protein	Transporters
K16089	Outer membrane receptor for ferrienterochelin and colicins	Transporters
K03406	Methyl-accepting chemotaxis protein	Signal transduction
K00799	Glutathione S-transferase	Metabolism of amino acids
K02660	Twitching motility protein PilJ	Secretion system
K02030	Polar amino acid transport system substrate-binding protein	Transporters
K02027	Multiple sugar transport system substrate-binding protein	Transporters
K10440	Tibose transport system permease protein	Transporters
K02483	Two-component system, OmpR family, response regulator	Two-component system
K01996	Branched-chain amino acid transport system ATP-binding protein	Transporters
K02056	Simple sugar transport system ATP-binding protein	Transporters
K01995	Branched-chain amino acid transport system ATP-binding protein	Transporters

Table S9: Sugar transporters in MAG1 annotated with eggNOG database.

KEGG	Gene	Type of transporter
K10111	*malK*	Multiple sugar transport
K10112	*msmX*	Multiple sugar transport
K10227	*smoE*	Sorbitol/mannitol transport system
K10228	*smoF*	Sorbitol/mannitol transport system
K10229	*smoG*	Sorbitol/mannitol transport system
K10439	*rbsB*	Ribose transport system
K10440	*rbsC*	Ribose transport system
K10441	*rbsA*	Ribose transport system
K10537	*araF*	L-arabinose transport system
K10538	*araH*	L-arabinose transport system
K10539	*araG*	L-arabinose transport system
K10543	*xylF*	D-xylose transport system
K10544	*xylH*	D-xylose transport system
K10545	*xylG*	D-xylose transport system
K10552	*frcB*	Fructose transport system
K10553	*frcC*	Fructose transport system
K10554	*frcA*	Fructose transport system
K10559	*rhaS*	Rhamnose transport system
K10560	*rhaP*	Rhamnose transport system
K10561	*rhaQ*	Rhamnose transport system
K10562	*rhaT*	Rhamnose transport system
K17315	*gtsA*	Glucose/mannose transport system
K17316	*gtsB*	Glucose/mannose transport system
K17317	*gtsC*	Glucose/mannose transport system

Table S10: Sugar transporters in MAG2 annotated with eggNOG database.

KEGG (continued)	Gene	Type of transporter
K06726	*rbsD*	D-ribose pyranase
K10108	*malE*	Maltose/maltodextrin transport system
K10109	*malF*	Maltose/maltodextrin transport system
K10110	*malG*	Maltose/maltodextrin transport system
K10111	*malK*	Multiple sugar transport
K10112	*msmX*	Multiple sugar transport
K10117	*msmE*	Raffinose/stachyose/melibiose transport system
K10118	*msmF*	Raffinose/stachyose/melibiose transport system
K10119	*msmG*	Raffinose/stachyose/melibiose transport system
K10191	*lacK*	Lactose/L-arabinose transport system
K10193	*togM*	Oligogalacturonide transport system
K10228	*smoF*	Sorbitol/mannitol transport system
K10229	*smoG*	Sorbitol/mannitol transport system
K10234	*aglG*	Alpha-glucoside transport system
K10235	*aglK*	Alpha-glucoside transport system
K10236	*thuE*	Trehalose/maltose transport system
K10237	*thuF*	Trehalose/maltose transport system
K10238	*thuG*	Trehalose/maltose transport system
K10240	*cebE*	Cellobiose transport system
K10241	*cebF*	Cellobiose transport system
K10242	*cebG*	Cellobiose transport system
K10439	*rbsB*	Ribose transport system
K10440	*rbsC*	Ribose transport system
K10441	*rbsA*	Ribose transport system
K10537	*araF*	L-arabinose transport system
K10538	*araH*	L-arabinose transport system
K10539	*araG*	L-arabinose transport system
K10540	*mglB*	Methyl-galactoside transport system
K10541	*mglC*	Methyl-galactoside transport system
K10542	*mglA*	Methyl-galactoside transport system
K10543	*xylF*	D-xylose transport system
K10544	*xylH*	D-xylose transport system
K10545	*xylG*	D-xylose transport system
K10548	-	Putative multiple sugar transport
K10550	*alsC*	D-allose transport system
K10552	*frcB*	Fructose transport system
K10553	*frcC*	Fructose transport system
K10554	*frcA*	Fructose transport system
K10558	*lsrA*	AI-2 transport system
K10559	*rhaS*	Rhamnose transport system
K10560	*rhaP*	Rhamnose transport system
K10561	*rhaQ*	Rhamnose transport system
K10562	*rhaT*	Rhamnose transport system
K15770	*cycB*	Arabinogalactan oligomer /
K15771	*ganP*	Arabinogalactan oligomer /
K15772	*ganQ*	Arabinogalactan oligomer /
K17204	*eryE*	Erythritol transport system
K17207	*xltA*	Putative xylitol transport
K17208	*ibpA*	Inositol transport system
K17209	*iatP*	Inositol transport system
K17210	*iatA*	Inositol transport system
K17213	-	Inositol transport system substrate-binding
K17214	-	Inositol transport system permease
K17215	-	Inositol transport system ATP-binding
K17241	*aguE*	Alpha-1 4-digalacturonate transport
K17242	*aguF*	Alpha-1 4-digalacturonate transport
K17244	*chiE*	Putative chitobiose transport
K17245	*chiF*	Putative chitobiose transport
K17246	*chiG*	Putative chitobiose transport

KEGG (continued)	Gene	Type of transporter
K17315	*gtsA*	Glucose/mannose transport system
K17316	*gtsB*	Glucose/mannose transport system
K17317	*gtsC*	Glucose/mannose transport system

Table S11: General type transporters in MAG3 annotated with eggNOG database.

KEGG	Type of transporter
K06147	ATP-binding cassette subfamily B bacterial
K02004	Putative ABC transport system permease
K02003	Putative ABC transport system ATP-binding
K01992	ABC-2 type transport system permease
K01990	ABC-2 type transport system ATP-binding
K03286	OmpA-OmpF porin OOP family
K11085	ATP-binding cassette subfamily B bacterial
K16013	ATP-binding cassette subfamily C bacterial

Table S12: Sugar transporters in MAG4 annotated with eggNOG database.

KEGG	Gene	Type of transporter
K06726	*rbsD*	D-ribose pyranase
K10108	*malE*	Maltose/maltodextrin transport system substrate-binding protein
K10109	*malF*	Maltose/maltodextrin transport system permease protein
K10110	*malG*	Maltose/maltodextrin transport system permease protein
K10111	*malK*	Multiple sugar transport system ATP-binding
K10112	*msmX*	Multiple sugar transport system ATP-binding
K10117	*msmE*	Raffinose/stachyose/melibiose transport system substrate-binding protein
K10118	*msmF*	Raffinose/stachyose/melibiose transport system permease protein
K10119	*msmG*	Raffinose/stachyose/melibiose transport system permease protein
K10191	*lacK*	Lactose/L-arabinose transport system ATP-binding protein
K10193	*togM*	Oligogalacturonide transport system permease protein
K10227	*smoE*	Sorbitol/mannitol transport system substrate-binding protein
K10228	*smoF*	Sorbitol/mannitol transport system permease protein
K10229	*smoG*	Sorbitol/mannitol transport system permease protein
K10234	*aglG*	Alpha-glucoside transport system permease protein
K10235	*aglK*	Alpha-glucoside transport system ATP-binding protein
K10236	*thuE*	Trehalose/maltose transport system substrate-binding protein
K10237	*thuF*	Trehalose/maltose transport system permease protein
K10238	*thuG*	Trehalose/maltose transport system permease protein
K10240	*cebE*	Cellobiose transport system substrate-binding protein
K10241	*cebF*	Cellobiose transport system permease protein
K10242	*cebG*	Cellobiose transport system permease protein
K10439	*rbsB*	Ribose transport system substrate-binding protein
K10440	*rbsC*	Ribose transport system permease protein
K10441	*rbsA*	Ribose transport system ATP-binding protein
K10537	*araF*	L-arabinose transport system substrate-binding protein
K10538	*araH*	L-arabinose transport system permease protein
K10539	*araG*	L-arabinose transport system ATP-binding protein
K10540	*mglB*	Methyl-galactoside transport system substrate-binding protein
K10541	*mglC*	Methyl-galactoside transport system permease protein
K10542	*mglA*	Methyl-galactoside transport system ATP-binding protein
K10543	*xylF*	D-xylose transport system substrate-binding protein
K10544	*xylH*	D-xylose transport system permease protein
K10545	*xylG*	D-xylose transport system ATP-binding protein
K10548	-	Putative multiple sugar transport system
K10550	*alsC*	D-allose transport system permease protein
K10552	*frcB*	Fructose transport system substrate-binding protein
K10553	*frcC*	Fructose transport system permease protein

KEGG (continued)	Gene	Type of transporter
K10554	*frcA*	Fructose transport system ATP-binding protein
K10559	*rhaS*	Rhamnose transport system substrate-binding protein
K10560	*rhaP*	Rhamnose transport system permease protein
K10561	*rhaQ*	Rhamnose transport system permease protein
K10562	*rhaT*	Rhamnose transport system ATP-binding protein
K15770	*cycB*	Arabinogalactan oligomer / maltooligosaccharide transport
K15771	*ganP*	Arabinogalactan oligomer / maltooligosaccharide transport
K15772	*ganQ*	Arabinogalactan oligomer / maltooligosaccharide transport
K17208	*ibpA*	Inositol transport system substrate-binding protein
K17213	-	Inositol transport system substrate-binding protein
K17214	-	Inositol transport system permease protein
K17215	-	Inositol transport system ATP-binding protein
K17241	*aguE*	Alpha-1 4-digalacturonate transport system substrate-binding
K17242	*aguF*	Alpha-1 4-digalacturonate transport system permease
K17244	*chiE*	Putative chitobiose transport system substrate-binding
K17245	*chiF*	Putative chitobiose transport system permease
K17246	*chiG*	Putative chitobiose transport system permease
K17313	*treU*	Trehalose transport system permease protein
K17315	*gtsA*	Glucose/mannose transport system substrate-binding protein
K17316	*gtsB*	Glucose/mannose transport system permease protein
K17317	*gtsC*	Glucose/mannose transport system permease protein
K17324	*glpS*	Glycerol transport system ATP-binding protein
K17325	*glpT*	Glycerol transport system ATP-binding protein

Table S13: Families of CAZymes observed in the MAGs, number of ORFs in each genome and associated enzymatic functions.

Family	MAG1	MAG2	MAG3	MAG4	Enzymatic activity
GH1	1	3	0	2	β-glucosidases and β-galactosidases:
GH10	0	1	0	1	Endo-beta-1,3-xylanase, endo-beta-1,4-xylanases
GH102	1	1	0	1	Lytic transglycosidease
GH103	1	1	1	1	Lytic transglycosideases
GH104	0	0	1	0	Lytic transglyosylases
GH105	0	0	1	0	Rhamnogalacturonidases
GH109	7	3	1	4	α-N-acetylgalactosaminidase
GH12	1	0	0	0	Endo-β-1,4-glucanase, endo-β-1,3-1,4-glucanase
GH123	0	0	1	0	N-acetyl-β-galactosaminidases
GH125	0	1	1	2	α-mannosidases
GH127	0	0	0	1	Arabinofuranosidase
GH128	0	2	0	2	β-1,3-glucanases
GH13_10	1	0	0	0	α-hydrolases, transglycosidases and isomerases
GH13_11	1	1	1	1	α-hydrolases, transglycosidases and isomerases
GH13_15	0	1	0	1	α-hydrolases, transglycosidases and isomerases
GH13_16	1	0	0	0	α-hydrolases, transglycosidases and isomerases
GH13_23	0	3	1	3	α-hydrolases, transglycosidases and isomerases
GH13_26	1	0	0	0	α-hydrolases, transglycosidases and isomerases
GH13_3	1	0	0	0	α-hydrolases, transglycosidases and isomerases
GH13_36	0	1	1	1	α-hydrolases, transglycosidases and isomerases
GH13_9	1	1	0	1	α-hydrolases, transglycosidases and isomerases
GH130	0	0	1	0	β-mannoside phosporilases
GH133	1	0	0	0	Amylo-α-1,6-glucosidase
GH135	1	0	0	0	Galactosaminogalactan hydrolase
GH141	1	0	0	0	α-L-fucosidase/xylanase
GH144	0	1	0	1	Endo-β-1,2-glucanase
GH15	3	0	0	0	Glucoamylase
GH16	1	0	0	0	β-1,4 or β-1,3 glycosidic bonds in various glucans and galactans 1,3-β-D-glucan endohydrolases and 1,3;1,4-β-D-glucan
GH17	0	0	2	0	Endohydrolases
GH18	4	4	3	7	Chitinases (EC 3.2.1.14) and endo-β-N-acetylglucosaminidase

Family (continued)	MAG1	MAG2	MAG3	MAG4	Enzymatic activity
GH19	0	0	1	0	Chitinases
GH2	2	1	5	1	β-galactosidases, β-glucuronidases, β-mannosidases, exo-β-glucosaminidases
GH20	0	2	3	3	Exo-acting β-N-acetylglucosaminidases, β-N-acetylgalactosamindase and β-6-SO3-N-Acetylglucosaminidases
GH23	7	5	5	5	Lytic transglycosideases
GH24	0	0	2	0	Lysozyme
GH27	2	0	1	0	α-Galactosidase
GH28	1	1	0	2	Polygalacturonases.
GH29	0	0	3	0	α-fucosidases
GH3	5	3	4	3	Exo-acting β-D-glucosidases, α-L-arabinofuranosidases, β-D-xylopyranosidases
GH30	0	0	1	0	β-glucosylceramidase, β-1,6-glucanase, and β-xylosidase
GH31	1	1	2	2	α-glucosidases
GH33	0	1	1	1	Sialidases and trans-sialidases
GH35	1	3	2	3	β-galactosidases
GH36	2	0	1	0	α-N-acetylgalactosaminidase
GH37	1	0	1	0	Trehalase
GH42	1	0	1	1	β-galactosidases, α-L-arabinosidase, β-D-fucosidase
GH49	1	0	0	0	dextranase, isopullulanase
GH5	1	0	0	0	Endo- and exoglucanases, endo- and exomannanases
GH5_13	1	0	0	0	Endo- and exoglucanases, endo- and exomannanases
GH5_19	0	1	0	1	Endo- and exoglucanases, endo- and exomannanases
GH5_43	1	0	0	0	Endo- and exoglucanases, endo- and exomannanases
GH5_48	0	1	0	1	Endo- and exoglucanases, endo- and exomannanases
GH50	0	0	1	0	β-agarase
GH51	0	0	0	1	L-arabinfuranosidases
GH53	0	0	0	3	β-1,4-galactanase
GH54	2	0	0	0	α-L-arabinofuranosidase and β-xylosidase
GH55	1	1	1	1	β-1,3-glucanases, including both exo- and endo
GH63	1	1	1	1	Exo-acting α-glucosidases
GH64	1	1	0	2	β-1,3-glucanase
GH65	0	0	2	0	Maltose phosphorylase, trehalose phosphorylase, kojibiose phosphorylase
GH73	1	1	2	1	β-N-acetylglucosaminidases.
GH76	0	1	1	0	Endo-acting α-mannanases
GH77	1	0	0	0	4-α-glucanotransferase
GH78	1	0	1	0	α-L-rhamnosidases
GH79	1	0	1	0	β-glucuronidase , β-4-O-methyl-glucuronidase , baicalin
GH87	1	0	0	0	β-glucuronidase , heparanase and hyaluronidase
GH9	0	0	1	1	α-1,3-glucanase
GH92	8	0	5	0	Cellulases
GH94	2	3	0	2	Exo-acting α-mannosidases, phosphorylase, cellodextrin phosphorylase, chitobiose phosphorylase
GH95	1	0	0	0	1,2-α-L-fucosidases, 1,2-α-L-galactosidases

Chapter 7

General Discussion

Acidobacteria is a phylum widely distributed in several soil types but their functional roles in ecosystem processes are still largely elusive. At present, cultured representatives of the 15 recently defined classes are available only for *Acidobacteriia*, *Blastocatellia*, *Holophagae*, *Vicinamibacteria* and *Thermoanaerobaculia* (Dedysh & Yilmaz, 2018). Therefore, most of the information and studies on the characterization of *Acidobacteria* are derived from subdivision 1 (class *Acidobacteriia*), which contains the majority of cultivated species to date. The overall goal of this thesis was to investigate and understand the metabolism of two strains belonging to the genus *Granulicella* (class *Acidobacteriia*), and the functions and environmental fate of extracellular polymeric substances (WH15EPS). The *Granulicella* genus is known for the copious production of EPS (Pankratov & Dedysh, 2010), and the EPS of our strains were chemically and physically characterized previously, showing interesting emulsification properties (Kielak *et al.*, 2017). The main objectives of my thesis were to i) study the responses *Granulicella* sp. strain WH15 to cellobiose as carbon source; ii) investigate the impact of manganese (Mn) and other trace elements, such as Fe, Zn, Cu, Co, Ni, B and Mo, on the growth and metabolism of *Granulicella* sp. strains WH15 and 5B5; iii) explore the assimilation of EPS of *Granulicella* sp. strainWH15 by other soil bacteria derived from the same environment where WH15 was isolated from; and iv) evaluate the potential of WH15EPS for enrichment of microbes, especially yet unclassified bacteria, capable of EPS degradation in culture medium. Here, the major findings of my thesis are discussed and future perspectives for studying the ecology of *Acidobacteria* are presented.

Culture media optimization and metabolism of *Granulicella*

One of the major hurdles of studying *Acidobacteria* is their isolation and propagation in culture media. Since they are slow growers, it can take weeks to months to develop colonies (Eichorst *et al.*, 2011, de Castro *et al.*, 2013). Recently, changes in traditional culture methods and the use of unconventional culture media composition have increased the number of *Acidobacteria* isolates that can be cultured in the lab. Currently, 62 species have been described compared to only 14 species in 2011 (de Castro, 2011). Media modifications include low concentration of nutrients (Janssen *et al.*, 2002, Stevenson *et al.*, 2004), unusual or complex polysaccharides as carbon sources (Pankratov *et al.*, 2008, Eichorst *et al.*, 2011), longer incubation periods (de Castro *et al.*, 2013), addition of humic acids and quorum-sensing molecules (Stevenson *et al.*, 2004), and the employment of soil solutions and inhibitors for unwanted microorganisms (de Castro *et al.*, 2013, Foesel *et al.*, 2013).

Once the isolates are obtained, the utilization of richer culture media, with higher concentrations of carbon and other nutrients can be applied for better cell proliferation (de Castro *et al.*, 2013, Kielak *et al.*, 2016). Campanharo *et al*. (2016) optimized a culture medium in order to improve growth of the *Acidobacteria* subdivision 1 strains used in my thesis, i.e. *Granulicella* strains WH15 and 5B5. In the solid culture medium PSYL 5 (**P**hosphate, **S**ucrose,

Yeast extract, **L**iquid medium, pH **5**), the strains produced visible colonies after around 3 days of incubation, a much shorter time frame as compared to earlier studies where *Acidobacteria* strains took 14-168 days to develop colonies (Eichorst *et al.*, 2011).

In order to study these strains and their EPS, further optimizations of the culture medium and growth conditions for extractable EPS yield were necessary. Our experiments showed that the agar brand was important for growth and EPS production in the laboratory. Initial experiments showed that, using Sigma Agar A1296-500, strain WH15 did not grow well and strain 5B5 was not able to grow at all. Our results demonstrated that both strains performed the best in BD agar (Figure 1). It is a common knowledge in laboratory procedures that the purity of the reagents is important for the reproducibility of results, since minor differences in medium composition can impact the growth of microorganisms (Atlas, 2010). However, few studies addressed the differences between agar brands, composition and preparation. For instance, Bosmans *et al.* (2016) showed that calcium content of the agar strongly affected antimicrobial activity of Firmicutes strain ST15.15/036 against *Agrobacterium*. Their study also addressed the metal ion composition of the different agar brands, showing that the products I used in this study had varied metal ion concentration, which could be affecting the performance of our strains in solid medium. Therefore, we tested if the addition of the trace element solution SL10 (described in chapter 4) to PSYL5 could enhance the growth of our *Granulicella* strains on media with the different agar brands. Indeed, the addition of SL10 (1X and 10X concentrations) enhanced bacterial growth, especially for Sigma (Figure1d) and Fluka (Figure 1e) brands, showing that the metal ion content of these agar brands was not fulfilling the requirements for the growth of our strains. In order to identify which component(s) was responsible for the growth enhancement, further experiments with the addition of metal ions separately were performed. Nevertheless, the best option was still BD agar (Figure 1c and 1f), since no extra requirements for bacterial development were necessary (Table 1).

Table 1: Summary of *Granulicella* strains growth and EPS production in culture medium containing different agar sources at different concentrations of SL10 solution in two temperatures.

	20°C		**30°C**	
	WH15	**5B5**	**WH15**	**5B5**
Agar brand	**PSYL 5 (growth/ EPS)**	**PSYL5 (growth/EPS)**	**PSYL 5 (growth/ EPS)**	**PSYL5 (growth/EPS)**
Sigma	**+/+**	**-/-**	--	--
Sigma 1X	**+++/+**	**+/+**	**+++/+**	**+/+**
Sigma 10X	**+++/+++**	**++/+**	**+++/+++**	**+++/++**
Fluka	**++/++**	**+/+**	**++/++**	**+/+**
Fluka 1X	**+++/++**	**++/+**	**+++/+**	**+++/+**
Fluka 10X	**+++/++**	**++/+**	**+++/++**	**+++/+**
BD	**+++/++**	**+++/+**	**+++/+++**	**+++/+**
BD 1X	**+++/++**	**+++/+**	**+++/+++**	**+++/+**
BD 10X	**+++/++**	**+++/++**	**+++/+++**	**+++/++**

1X: trace element solution SL10 1ml/l; 10X: trace element solution SL10 10 ml/l.

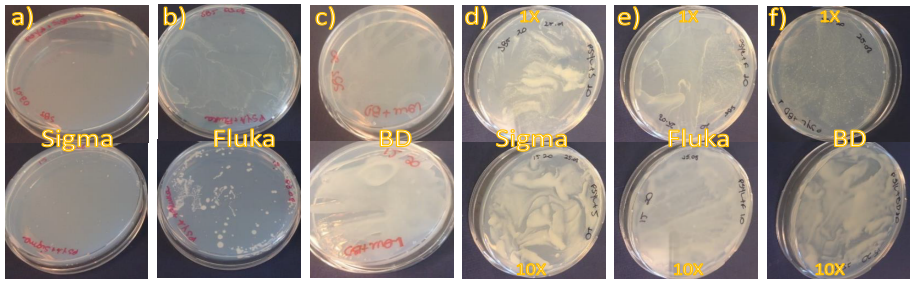

Figure 1: Growth of *Granulicella* strains 5B5 (top) and WH15 (bottom) in PSYL5 culture medium containing different agar brands after incubation for 15 days at 20 ^{0}C. a) Sigma agar, b) Fluka agar c) BD agar. Growth of both strains after the addition of trace element solution and incubation for 15 days at 20 ^{0}C. d) Sigma agar, e) Fluka agar, f) BD agar; 1X: trace element solution SL10 1ml/l; 10X: trace element solution SL10 10 ml/l.

Culture media optimization demonstrated that both strains performed better when growing in culture media with higher carbon concentration. The two strains were originally isolated from wood-decay material using a 10X diluted culture medium TSB (3g of sugar) (Valášková *et al.*, 2009), while the optimized culture medium PSYL5 had a 10 times higher quantity of carbon (30g of sugar). In chapter 3, I investigated the impact of two concentrations of carbon source (0.025 g and 30g of cellobiose) on the metabolism of *Granulicella* sp. WH15 strain using a multi-omics approach. Firstly, we sequenced the complete genome of strain WH15, which has large number of genes dedicated to carbohydrate metabolism, involving the utilization of various carbon sources, hydrolysis of complex polysaccharides and biosynthesis of exopolymers (EPS) corroborating with other published *Granulicella* strains (Pankratov & Dedysh, 2010). The impact of higher concentrations in the metabolism of *Granulicella*, however, has not been studied previously. My proteomic and transcriptomic data analyses indicated that the higher cellobiose concentration, after 3 days of incubation, triggered a stress response in the strain studied. The transcriptomics profile of strain WH15 under high cellobiose concentration revealed the expression of several stress proteins, such as factor σW and toxin-antitoxin (TA) system HigAB, which in other bacterial genera are involved in responses to environmental stress stimuli, such as salt stress, presence of detergents (Zweers *et al.*, 2012), nutritional stress, heat shock and others (Tachdjian & Kelly, 2006, Jorgensen *et al.*, 2008). In addition, genes linked to osmoprotection were upregulated, such as *gfo4* (the glucose-fructose oxidoreductase), and glycerol transporter *glpF*, which could be important to balance the impact of the cellobiose quantity in the culture medium. The proteomic profile of strain WH15 at high cellobiose concentration was consistent with the transcriptomic response, revealing several membrane efflux pumps and transporter proteins, which could be involved in cell wall osmotic protection mechanisms. *Acidobacteria* in general possess a large variety of transporters involved in the acquisition and secretion of a wide range of substrates, which is suggested to be an adaptation advantage especially in low nutrient environments (Kielak *et al.*, 2016). Furthermore, I showed that higher cellobiose concentrations caused the

repression of pathways related to the metabolism of other carbon sources, especially the pentose phosphate (PP) pathway. Those results suggest that, *Granulicella* sp. strain WH15 reacts to higher carbon concentrations by upregulating genes associated with the transport of toxic compounds and by reallocating resources for cell maintenance of basic metabolism instead of growth. However, the mechanisms underlying adaptation towards enhanced growth still have to be addressed in further studies.

Another important question raised by bacterial growth optimization procedures was the impact of the trace elements on the metabolism of *Granulicella* sp. WH15 and 5B5 strains. Both strains were isolated from an environment (wood in late stage of decomposition) containing high amounts of manganese (chapter 5) ~17 times higher than the average observed in normal local soil without litter (5.8-8.0 mg Mn/Kg). Hence, in chapter 4, I examined the effects of trace element solution SL10 on the growth of both WH15 and 5B5 strains. The addition of the trace element solution had a clear impact in the growth yield of our strains, which responded similarly to 1X concentration of SL10 solution. However, the higher concentration of SL10 suppressed the growth of WH15 strain, while still enhancing the growth of 5B5 strain, highlighting differences in growth responses of the two strains. In order to identify the metal ion with the strongest effect on bacterial growth, I performed separate experiments for each of the 8 metal ions present in the SL10 solution. The analysis showed that only manganese (Mn) affected positively the growth of both strains, leading to a further investigation of their proteomic profile upon Mn exposure. Interestingly, the results did not show overexpression of any specific pathway, demonstrating that the presence of Mn did not have a clear impact in the regulation of protein expression. Nonetheless, higher number of upregulated proteins that can use Mn^{2+} as co-factor was detected, as well as several candidate transporters that could be involved in Mn homeostasis of *Granulicella*.

Both our strains showed tolerance to Mn at a level that is comparable to *Serratia* strains isolated from Mn water mines in Brazil (Barboza *et al.*, 2017). Even though *Acidobacteria* were already associated to heavy-metal contaminated sites in culture-independent surveys (Gremion *et al.*, 2003, Guo *et al.*, 2017), no study to date addressed heavy metal resistance of acidobacterial isolates. Nonetheless, the resistance to higher Mn concentrations, in comparison to other metal ions, presumably is linked to the survival of both *Granulicella* strains in decaying wood material. White rot fungi, for instance *Hypholoma fasciculare,* found in the same environment as our strains, can create deposits of Mn oxides in degraded wood (Blanchette, 1984).

Mn is an important co-factor of enzymes involved sugar, lipid and protein metabolism (Jensen & Jensen, 2014). The presence of Mn may have, for instance, enhanced *Granulicella* strains growth by improving metabolic activities linked to cell cycle and division. In addition, Mn can be a protection factor against reactive oxygen species (Jakubovics & Jenkinson, 2001, Jensen & Jensen, 2014) generated by fungal peroxidases and phenol oxidases during degradation of lignin and cellulose (Kielak *et al.*, 2016), that would further support our hypothesis that our

strains have a competitive advantage in this environment, not only tolerating Mn but also using the metal for protection against reactive oxygen species.
Furthermore, the gene for the most common transcriptional Mn homeostasis regulator, *mntR*, could not be detected in the genome, suggesting that manganese homeostasis in these two *Granulicella* strains is under the control of a different regulator. In these analyses, I identified a wide range of transporters that could be involved in metal ion homeostasis for both *Granulicella* strains. It was already observed that strains of *Acidobacteria* have a large proportion of their genomes dedicated to transport (Kielak et al., 2016). Most of the genes are related to drug/metabolite transport; however, a broad range of substrate categories can be observed, suggesting that *Acidobacteria* may have advantages in nutrient uptake especially in oligotrophic, nutrient-limited conditions (Kielak *et al.*, 2016). Efflux is one of the key mechanisms that bacteria have developed to tolerate the presence of metal ions (Porcheron *et al.*, 2013). Along with potential transporters involved in manganese homeostasis, such as metal translocating P-type ATPase copA and metal ABC transporter ATP-binding protein troB, I detected 78 genes in strain WH15 and 23 genes in strain 5B5 that are candidates for the transportation of a diverse range of metal ions, such as Cu, Ag, As, Zn. Even though the strains did not show a high resistance to most of the metals in culture medium, further tests in liquid culture media with absorbance measurements should be performed in order to examine the range of metal resistance of our *Granulicella* strains and determine their minimum inhibitory concentration (MIC). Combining the results of chapter 3 and chapter 4, I showed that our strains are equipped with several types of transporters, which could be involved not only in nutrient uptake but also in detoxification and resistance important for their survival in soil and wood decay environments.
The growth optimization tests that I have carried out in chapters 3 and 4 allowed for improvements in obtaining higher bacterial growth yields, as well as enhanced EPS production, steps that were crucial for further examination of the ecological functions and environmental fate of *Granulicella* EPS. Furthermore, the investigations I performed in chapters 3 and 4 demonstrated that *Granulicella* strains are equipped with several features, namely tolerance to higher concentrations of carbon sources and Mn, presence of several types of glycosyde hydrolases and a wide range of transporters, which might be fundamental characteristics for their development and prevalence in soil environments.

7

Granulicella EPS – ecological role and application

An additional bacterial trait associated with protection against harsh environments is the production of extracellular polymeric substances (EPS). *Granulicella* species are known to produce large amounts of EPS (Pankratov & Dedysh, 2010), biopolymers with protective capabilities against antimicrobial compounds and environmental stresses such as drought (Souli & Giamarellou, 1998, Krembs *et al.*, 2011, Upadhyay *et al.*, 2011). Microbial EPS typically

consist of mucoid, viscous substances that mediate contact and exchange of nutrients among cells and the environment, as well as providing ideal moisture conditions for enzymatic reactions (Wingender *et al.*, 1999). EPS protective capacity was reported for several types of environmental stresses, such as drought, temperature (Bhatnagar *et al.*, 2014), pH, salinity, metal ions (Upadhyay *et al.*, 2011, Upadhyay *et al.*, 2017), as well as antimicrobial compounds (Davenport *et al.*, 2014). The main components of EPS are carbohydrates (Flemming & Wingender, 2010), yet few studies addressed the importance of EPS as carbon source. To this end, I applied Stable Isotope probing (SIP) coupled to rRNA gene sequencing to identify microbes capable of using EPS as a carbon source (Chapter 5). I used WH15EPS, since strain WH15 produces higher amounts of EPS in the laboratory than strain 5B5. The characterization of WH15EPS showed that the biopolymer is composed of xylose, mannose, glucose, galactose, rhamnose, glucuronic and galacturonic acids (Kielak *et al.*, 2017). Hence, metabolization of WH15EPS would demand a variety of glycoside hydrolases. I showed that WH15EPS enriched for a specific microbial community from the topsoil litter and I identified the main EPS metabolizers within 3 timepoints. After 35 days of incubation, the addition of WH15EPS to soil enriched mostly bacteria belonging to the phylum *Planctomycetes*, especially *Singulisphaera*. Interestingly, *Planctomycetes*, similar to *Acidobacteria*, possess few isolated representatives due to difficulties in cultivation. However, genomic studies of the available isolates, as well as cultivation-independent studies identified a wide range of glycosyl hydrolases in *Planctomycetes*, enzymes involved in degrading complex polysaccharides (Ivanova *et al.*, 2017, Ivanova *et al.*, 2017). Unclassified *Planctomycetes* were also enriched in a similar study by Wang *et al.* (2015) in which the authors applied a microbial biopolymer with a different composition to soil. In addition, we also observed the enrichment of *Rhodanobacter,* a bacterial genus associated with the degradation of aromatic compounds (Uhlik *et al.*, 2012), and other poorly characterized groups, such as *Pedospheraceae*, which belongs to the phylum *Verrucomicrobia* (Spring *et al.*, 2016). Among the fungal genera, I detected *Trichoderma, Scleroderma* and *Mortierella* and also several ectomycorrhizal fungi, known producers of extracellular enzymes for organic material decomposition (Valášková *et al.*, 2009, Urbanová *et al.*, 2015).

Co-occurrence analysis demonstrated that *Singulisphaera* had potential positive interactions with other bacterial phyla, suggesting cooperation with other bacteria for the metabolization of WH15EPS, and some negative connections, possibly due to competition with fungal genera. Even though there is no experimental evidence that strain WH15 can use its own EPS as a carbon source, *Granulicella* was detected in the labeled treatment. It should be emphasized that our co-occurrence network analysis should be evaluated cautiously. I propose that more targeted metabolic studies can be more insightful for the generation of ecological inferences, while only pure evaluation of total DNA content of samples can yield misleading interpretations. Furthermore, studies show that the total DNA pool of an environment can contain non-active and dormant microorganisms, masking the more active components of

microbial communities (Kuramae *et al.*, 2013, Harkes *et al.*, 2019, Lupatini *et al.*, 2019).
In addition to the microbial taxonomic classification, SIP allows targeting functions associated with the incorporation of the labeled material. With this in mind, I analyzed metagenomics sequences originated from the SIP experiment of chapter 5 in order to find interesting genes that encoded enzymes involved in the hydrolysis of WH15EPS (Chapter 6). I also employed WH15EPS as an enrichment factor and carbon source in culture medium, to increase the number of unidentified microorganisms. Taxonomic analyses of the SIP metagenomics dataset were consistent with those from chapter 5, showing an increase in the abundance of *Planctomycetes*. In addition, I also observed a high abundance of sequences affiliated to the phyla *Proteobacteria*, *Actinobacteria* and *Acidobacteria*, which are widely known to be involved in carbon degradation processes, such as degradation and assimilation of cellulose, xylan, pectin and chitin (Haichar *et al.*, 2007, Kielak *et al.*, 2016). On the other hand, the culture-based metagenomics showed the predominance of *Proteobacteria*, which are usually easier to isolate from environmental samples in culture medium, but also a high abundance of unclassified bacteria, demonstrating the potential of EPS for the enrichment and isolation of previously uncultivated microbes. Futhermore, the culture-based metagenomics dataset allowed the assembly of 4 metagenome-assembled genomes (MAGs) of unclassified *Proteobacteria*. Besides unknown microbes, I was able to isolate 201 strains of bacteria and fungi, among them 143 *Dyella*, 27 *Burkholderia*, 11 *Pseudomonas*, 1 *Mucilaginibacter, 2 Buttiauxella*, 4 *Rahnella*, 3 *Enterobacter,* 2 *Penicillium,* 1 *Trichoderma* and 1 *Pseudogymnoascus,* which will be tested in further studies for enzymatic activity.
Collectively, these results confirm and extend the study of Verastegui *et al.* (2014), in which the presence of a complex polysaccharide promoted the enrichment of microorganisms capable of breaking down several plant-derived carbohydrates. In further functional analyses, I showed that the microbial community enriched in both cultivation dependent and independent approaches was abundant with carbohydrate-associated enzymes (CAZymes) with potential applications in diverse sectors of industry. Overall, the most abundant CAZyme families were glycosyl transferases, which can be explained by the high proportion (~1.5% of total genes) of GTs in the genomes of most microorganisms in general (Lairson *et al.*, 2008). GTs are not widely explored in industries since their application requires activated donor precursors (Yuan *et al.*, 2018). However, they can be employed for manufacturing of modified natural products with medical applications, such as anticancer compounds (Schmid *et al.*, 2016).
Overall, I showed the presence of 310 CAZyme families, from which 38.4% (119) were GH families. Even though the potential enzymes might belong to slow growing microorganisms in laboratory conditions, such as *Acidobacteria*, *Planctomyctes* and *Verrucomicrobia*, sequences can still be targeted for further heterologous expression and characterization. Within the most abundant GHs, I found families of amylases, acetylgalactosaminidases, glucuronidases, agarases, invertases and endoglucanases. For instance, amylases are widely employed in food industry, where they can be used for production of glucose and maltose

syrups, reduction of viscosity of syrups, production of clarified fruit juices, solubilization of starch for brewing processes and manufacture of baked products (Liu & Kokare, 2017). Other GHs can be applied for the production of oligosaccharides with antioxidant activities for food, pharmaceutical and cosmetic industries (Fu & Kim, 2010), as well as be used in biological control, bioremediation processes (Dahiya *et al.*, 2005), production of biofuels and paper (McKee *et al.*, 2016). Another important application of GHs is the removal or inhibition of biofilm. Amylase and cellulose solutions have already been used to inhibit biofilm formation of *S. aureus* and *P. aeruginosa* decreasing significantly bacterial biomass and increasing the effectiveness of antibiotics treatments (Craigen, 2011, Flemming *et al.*, 2016).

In chapters 5 and 6, I showed that WH15EPS can be a potential carbon source for a diverse range of microorganisms, suggesting direct and indirect associations between *Granulicella* and other soil bacteria, especially *Singulisphaera*. I demonstrated that WH15EPS can be used as an enrichment factor for the isolation of known and unknown microbes with biotechnological potential, as well as for targeting sequences of a wide range of CAZYmes. In conclusion, I showed that targeted metabolic studies can be effective to unravel microbial interactions and reinforce characterization of these groups to better understand their ecological functions in the environment.

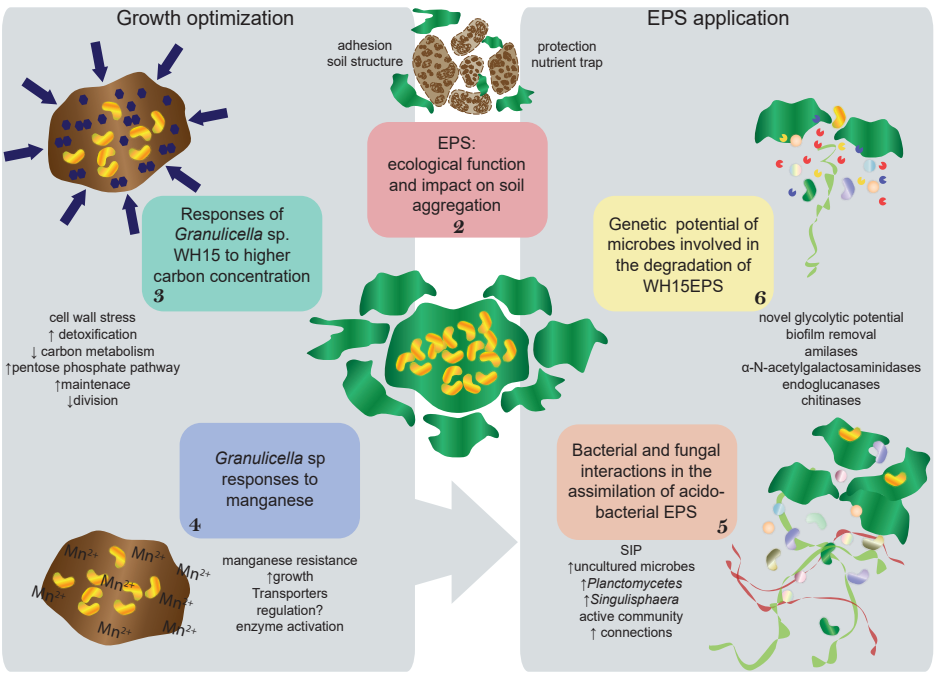

Figure 2:: Schematic overview of the main topics and most important findings described in this thesis.

Concluding remarks and future prospects

Our understanding of *Acidobacteria* has gradually increased throughout the years, but there is still a huge knowledge gap, especially concerning subdivisions with no cultivated representatives. To bridge this gap, I studied two strains of *Granulicella*, *Acidobacteria* belonging to the subgroup 1, class *Acidobacteriia*. This is the first time that the impact of trace elements and carbon source concentration on the metabolism of *Acidobacteria* strains was tested with different 'omics approaches. With my research, I showed that trace elements, especially Mn, are important for the growth of *Granulicellla* strains and can be used to enhance their growth in culture medium. Together with a higher carbon concentration, the addition of trace elements is necessary for improving *Acidobacteria* growth and can be used in future research for isolation and propagation of other *Acidobacteria* strains.

I further demonstrated that our strains harbour a wide range of transporters involved in nutrient transport and detoxification, which can play a role in surviving stressful and harsh conditions. In addition, the transporters can be involved in homeostasis and resistance to the high Mn concentrations released by fungi in wood decomposition environments, the natural habitat from where they were isolated. The variety of transporters found in the genomes and proteomes of our *Granulicella* strains suggest they contribute to stress resistance of several *Acidobacteria*. Further studies, however, are needed to validate these hypotheses.

The research presented in this thesis highlights the importance of targeted metabolic studies to disentangle potential microbial interactions and unknown functions. In this context, the application of modern techniques, such as nanoscale secondary ion mass spectrometry (nanoSIMS) and DNA-SIP coupled to high-throughput sequencing can be useful for unraveling further metabolic mechanisms and interactions among microbes *in situ*. Additional studies can be performed involving isolated strains between which we found connections, such as *Granulicella* and *Singulisphaera*, evaluating the interactions of these genera during the metabolization of labeled carbon and nitrogen sources, such as glucose, cellulose or chitin, for instance. Furthermore DNA-SIP can be applied for the study of other microbial EPS in an ecological context. In general, the cultivation of novel microbes is still necessary, since it allows their functional characterization and application in studies that include pairwise or group interactions, either in culture media or in environmental samples. Furthermore, the isolation of new species belonging to *Acidobacteria* classes that do not have cultured representatives is fundamental for the exploration of their functions, since the variation within the phylum is broad, not allowing generalizations (Kielak *et al.*, 2016).

In this thesis I showed that the use of a complex polymer (WH15EPS) enriched microbial communities in unclassified microbes using culture-dependent and independent approaches; therefore, WH15EPS and other microbial biopolymers with different compositions may be an interesting source of variation for the cultivation of new species. Furthermore, the analysis of new metagenome-assembled genomes may provide novel information that could be applied for their isolation. The isolation of unclassified microbes may rely in the development of new

cultivation techniques, but also in the use of unconventional compounds in the preparation of culture media, as well as different growth conditions or the combination of both.

Using WH15EPS as an enrichment factor, I showed that *Acidobacteria* and other slow-growing phyla have an unexplored genetic and biotechnological potential that can be investigated through heterologous studies, not needing their cultivation in laboratory. My results indicated that microorganisms that incorporated and used WH15EPS as a carbon source have a wide glycolytic potential, possessing genes involved in the degradation of several complex polymers. I found several enzymes with potential applications in industry and the extent of their efficiency can be evaluated in further studies. Gene sequences belonging to enzymes of interest in metagenome data can be recovered from assembled metagenome data and used for expression. In addition, labeled DNA can be used for library construction and functional screening (Verastegui *et al.*, 2014), which also can be performed for the microorganisms I isolated. Overall, my findings in this thesis demonstrate that *Granulicella* are resourceful bacteria, which produce EPS with diverse ecological functions and a wide range of applications. Furthermore, the use of EPS allowed the enrichment and cultivation of unclassified *Proteobacteria*, demonstrating that even the biotechnological potential of well studied phyla is far from being fully explored.

In conclusion, the data presented in this thesis establish a solid fundamental basis for more mechanistic studies of *Acidobacteria* and other uncultivated microbes. Here I list questions to be answered in future studies on this high abundant and ubiquitous unexplored group of bacteria, *Acidobacteria* inhibiting soil ecosystems. More studies are needed to address the role of Mn in protection against oxidative stress and gene regulation of Mn homeostasis in *Granulicella* and other *Acidobacteria* strains. In this thesis we observed that *Planctomycetes* are capable to incorporate WH15EPS. Thus, we raise questions related to the potential interactions of *Acidobacteria* with other soil microbes. Does *Granulicella* produce EPS in presence of *Singulisphaera*, in laboratory and field conditions? How is the interaction between *Granulicella* and *Singulisphaera* during the degradation of cellulose or hemicellulose? Do they compete or collaborate? Is the dynamics of interaction dependent on nutrient source or concentration? Those are interesting questions that, once addressed will expand further our knowledge on *Acidobacteria*, providing more insights into their ecological roles and the survival strategies of this yet underexplored bacterial phylum, the interactions of these bacteria with other soil dwelling microbes, and possible biotechnological applications of their genes and their biopolymers.

References

A

Abbas AS & Edwards C (1990) Effects of metals on *Streptomyces coelicolor* growth and actinorhodin production. *Appl Environ Microbiol* **56**: 675-680.

Alami Y, Achouak W, Marol C & Heulin T (2000) Rhizosphere soil aggregation and plant growth promotion of sunflowers by an exopolysaccharide-producing *Rhizobium* sp strain isolated from sunflower roots. *Appl Environ Microbiol* **66**: 3393-3398.

Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF & Quince C (2014) Binning metagenomic contigs by coverage and composition. *Nat Methods* **11**: 1144-1146.

Amellal N, Burtin G, Bartoli F & Heulin T (1998) Colonization of wheat roots by an exopolysaccharide-producing pantoea agglomerans strain and its effect on rhizosphere soil aggregation. *Appl Environ Microbiol* **64**: 3740-3747.

Amellal N, Bartoli F, Villemin G, Talouizte A & Heulin T (1999) Effects of inoculation of EPS-producing Pantoea agglomerans on wheat rhizosphere aggregation. *Plant Soil* **211**: 93-101.

Amézketa E (1999) Soil Aggregate Stability: A Review. *J Sustain Agr* **14**: 83-151.

Angers DA & Mehuys GR (1989) Effects of Cropping on Carbohydrate Content and Water-Stable Aggregation of a Clay Soil. *Can J Soil Sci* **69**: 373-380.

Anwar MA, Kralj S, Pique AV, Leemhuis H, van der Maarel MJEC & Dijkhuizen L (2010) Inulin and levan synthesis by probiotic *Lactobacillus gasseri* strains: characterization of three novel fructansucrase enzymes and their fructan products. *Microbiology* **156**: 1264-1274.

Ashraf M, Hasnain S, Berge O & Mahmood T (2004) Inoculating wheat seedlings with exopolysaccharide-producing bacteria restricts sodium uptake and stimulates plant growth under salt stress. *Biol Fertil Soils* **40**.

Aspiras RB, Allen ON, Harris RF & Chesters G (1971) Aggregate stabilization by filamentous microorganisms. *Soil Sci* **112**: 3.

Atlas RM (2010) *Handbook of microbiological media*. CRC Press c2004., Boca Raton, Florida.

Ayeldeen M, Negm A, El-Sawwaf M & Kitazume M (2017) Enhancing mechanical behaviors of collapsible soil using two biopolymers. *J Rock Mech Geotech Eng* **9**: 329-339.

Azeredo J & Oliveira R (2000) The role of exopolymers in the attachment of *Sphingomonas paucimobilis*. *Biofouling* **16**: 59-67.

Aziz RK, Bartels D, Best AA, *et al.* (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9**.

B

Bae J, Oh E & Jeon B (2014) Enhanced transmission of antibiotictesistance in *Campylobacter jejuni* biofilms by natural transformation. *Antimicrob Agents Chemother* **58**: 7573-7575.

Baldock JA (2002) Interactions of organic materials and microorganisms with minerals in the stabilization of soil structure. *Interactions between soil particles and microorganisms,* Vol. 8 (Huang PM, Bollag J-M & Senesi N, eds.), John Wiley & Sons, Ltd, England.

Banerjee D, Jana M & Mahapatra S (2009) Production of exopolysaccharide by endophytic *Stemphylium sp*. *Micol Apl Int* **21**: 6.

Bao Y, Dolfing J, Wang B, Chen R, Huang M, Li Z, Lin X & Feng Y (2019) Bacterial communities involved directly or indirectly in the anaerobic degradation of cellulose. *Biol Fertil Soils* **55**: 201-211.

Barboza NR, Morais MMCA, Queiroz PS, Amorim SS, Guerra-Sá R & Leão VA (2017) High manganese tolerance and biooxidation ability of Serratia marcescens isolated from manganese mine water in Minas Gerais, Brazil. *Front Microbiol* **8**.

Barns SM, Takala SL & Kuske CR (1999) Wide distribution and diversity of members of the bacterial kingdom Acidobacterium in the environment. *Appl Environ Microbiol* **65**: 1731-1737.

Beare MH, Hendrix PF, Cabrera ML & Coleman DC (1994) Aggregate-protected and unprotected organic matter pools in conventional- and no-tillage soils. *Soil Sci Soc Am J* **58**: 787.

Becker A, Katzen F, Pühler A & Ielpi L (1998) Xanthan gum biosynthesis and application: a biochemical/genetic perspective. *Appl Microbiol Biotechnol* **50**: 145-152.

Belova SE, Ravin NV, Pankratov TA, Rakitin AL, Ivanova AA, Beletsky AV, Mardanov AV, Sinninghe Damsté JS & Dedysh SN (2018) Hydrolytic Capabilities as a Key to Environmental Success: Chitinolytic and Cellulolytic *Acidobacteria* From Acidic Sub-arctic Soils and Boreal Peatlands. *Front Microbiol* **9**.

Bengtsson-Palme J, Ryberg M, Hartmann M, *et al.* (2013) Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods Ecol Evol* **4**.

Bergmann GT, Bates ST, Eilers KG, Lauber CL, Caporaso JG, Walters WA, Knight R & Fierer N (2011) The under-recognized dominance of Verrucomicrobia in soil bacterial communities. *Soil Biol Biochem* **43**: 1450-1455.

Berini F, Casciello C, Marcone GL & Marinelli F (2017) Metagenomics: novel enzymes from non-culturable microbes. *FEMS Microbiol Lett* **364**.

Berne C, Ducret A, Brun YV & Hardy GG (2015) Adhesins involved in attachment to abiotic surfaces by gram-negative bacteria. *Microbiol Spectr* **3**.

Bezzate S, Aymerich S, Chambert R, Czarnes S, Berge O & Heulin T (2000) Disruption of the *Paenibacillus polymyxa* levansucrase gene impairs its ability to aggregate soil in the wheat rhizosphere. *Environ Microbiol* **2**: 333-342.

Bhatnagar M, Pareek S, Bhatnagar A & Ganguly J (2014) Rheology and characterization of a low viscosity emulsifying exopolymer from desert borne *Nostoc calcicola*. *Indian J Biotechnol* **13**: 241-246.

Bianchetti CM, Elsen NL, Fox BG & Phillips GN (2011) Structure of cellobiose phosphorylase from *Clostridium thermocellum* in complex with phosphate. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **67**: 1345-1349.

Blanchette RA (1984) Manganese accumulation in wood decayed by white rot fungi. *Phytopathology* **74**.

Blaud A, Lerch TZ, Chevallier T, Nunan N, Chenu C & Brauman A (2012) Dynamics of bacterial communities in relation to soil aggregate formation during the decomposition of 13C-labelled rice straw. *Appl Soil Ecol* **53**: 1-9.

Bleriot C, Effantin G, Lagarde F, Mandrand-Berthelot MA & Rodrigue A (2011) RcnB Is a periplasmic protein essential for maintaining intracellular Ni and Co concentrations in *Escherichia coli*. *J Bacteriol* **193**: 3785-3793.

Bödeker ITM, Nygren CMR, Taylor AFS, Olson Å & Lindahl BD (2009) ClassII peroxidase-encoding genes are present in a phylogenetically wide range of ectomycorrhizal fungi. *ISME J* **3**: 1387-1395.

Boer Wd, Folman LB, Summerbell RC & Boddy L (2005) Living in a fungal world: impact of fungi on soil bacterial niche development. *FEMS Microbiol Rev* **29**: 795-811.

Bolger AM, Lohse M & Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120.

Bolla K, Gopinath BV, Shaheen SZ & Charya MAS (2010) Optimization of carbon and nitrogen sources of submerged culture process for the production of mycelial biomass and exopolysaccharides by *Trametes versicolor*. *Int J Biotechnol Mol Biol Res* **1**: 7.

Boonchai R, Kaewsuk J & Seo G (2014) Effect of nutrient starvation on nutrient uptake and extracellular polymeric substance for microalgae cultivation and separation. *Desalin Water Treat* **55**: 360-367.

Bosmans L, De Bruijn I, De Mot R, Rediers H & Lievens B (2016) Agar composition affects in vitro screening of biocontrol activity of antagonistic microorganisms. *J Microbiol Methods* **127**: 7-9.

Braak CJFt & Smilauer P (2012) *Canoco reference manual and user's guide: software for ordination, version 5.0*. Microcomputer Power, Ithaca USA.

Bronick CJ & Lal R (2005) Soil structure and management: a review. *Geoderma* **124**: 3-22.

Buchfink B, Xie C & Huson DH (2014) Fast and sensitive protein alignment using DIAMOND. *Nat Methods* **12**: 59-60.

Burke DJ, Smemo KA & Hewins CR (2014) Ectomycorrhizal fungi isolated from old-growth northern hardwood forest display variability in extracellular enzyme activity in the presence of plant litter. *Soil Biol Biochem* **68**: 219-222.

Bushnell B (2015) BBMap.

C

Caesar-TonThat T-C & Cochran VL (2000) Soil aggregate stabilization by a saprophytic lignin-decomposing basidiomycete fungus I. Microbiological aspects. *Biol Fertil Soils* **32**: 6.

Caesar-TonThat TC, Caesar AJ, Gaskin JF, Sainju UM & Busscher WJ (2007) Taxonomic diversity of predominant culturable bacteria associated with microaggregates from two different agroecosystems and their ability to

aggregate soil in vitro. *Appl Soil Ecol* **36**: 10-21.

Caesar-TonThat TC, Stevens WB, Sainju UM, Caesar AJ, West M & Gaskin JF (2014) Soil-Aggregating Bacterial Community as Affected by Irrigation, Tillage, and Cropping System in the Northern Great Plains. *Soil Sci* **179**: 11-20.

Campanharo JC, Kielak AM, Castellane TCL, Kuramae EE & Lemos EGdM (2016) Optimized medium culture for *Acidobacteria* subdivision 1 strains. *FEMS Microbiol Lett* **363**: fnw245.

Caravaca F, Hernández T, García C & Roldán A (2002) Improvement of rhizosphere aggregate stability of afforested semiarid plant species subjected to mycorrhizal inoculation and compost addition. *Geoderma* **108**: 133-144.

Cardman Z, Arnosti C, Durbin A, Ziervogel K, Cox C, Steen AD, Teske A & Spormann AM (2014) *Verrucomicrobia* are candidates for polysaccharide-degrading bacterioplankton in an arctic fjord of Svalbard. *Appl Environ Microbiol* **80**: 3749-3756.

Carini P, Marsden PJ, Leff JW, Morgan EE, Strickland MS & Fierer N (2016) Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nat Microbiol* **2**.

Carzaniga T, Antoniani D, Dehò G, Briani F & Landini P (2012) The RNA processing enzyme polynucleotide phosphorylase negatively controls biofilm formation by repressing poly-N-acetylglucosamine (PNAG) production in *Escherichia coli* C. *BMC Microbiol* **12**: 270.

Chamizo S, Cantón Y, Miralles I & Domingo F (2012) Biological soil crust development affects physicochemical characteristics of soil surface in semiarid ecosystems. *Soil Biol Biochem* **49**: 96-105.

Chandrangsu P, Rensing C & Helmann JD (2017) Metal homeostasis and resistance in bacteria. *Nat Rev Microbiol* **15**: 338-350.

Chang I & Cho G-C (2012) Strengthening of Korean residual soil with β-1,3/1,6-glucan biopolymer. *Constr Build Mater* **30**: 30-35.

Chang I, Im J, Prasidhi AK & Cho G-C (2015) Effects of Xanthan gum biopolymer on soil strengthening. *Constr Build Mater* **74**: 65-72.

Chang I, Im J, Lee S-W & Cho G-C (2017) Strength durability of gellan gum biopolymer-treated Korean sand with cyclic wetting and drying. *Constr Build Mater* **143**: 210-221.

Chen Y, Pu G, Lian B, Pei X, Huang G, Wang Q & Lv Y (2018) Interactions between Two Fungi Strains during Litter Decomposition through a Microcosm Experiment: Different Degradative Enzyme Activities. *Adv Enzyme Res* **06**: 1-9.

Cheng HP & Walker GC (1998) Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. *J Bacteriol* **180**: 5183-5191.

Chenu C (1995) Extracellular polysaccharides : an interface between microorganisms and soil constituents. *Environmental impact of soil component interactions Natural and Anthropogenic organics,*(Huang PM, Berthelin J, Bollag JM, McGill WB & Page AL, eds.), CRC Lewis Publishers, Boca Raton, USA.

Chenu C & Roberson EB (1996) Diffusion of glucose in microbial extracellular polysaccharide as affected by water potential. *Soil Biol Biochem* **28**: 877-884.

Chenu C & Cosentino D (2011) Microbial regulation of soil structural dynamics. *The Architecture and Biology of Soils: Life in Inner Space,*(Ritz K & Young I, eds.), 37-70. CABI, Wallingford, Oxfordshire, UK.

Chin C-S, Alexander DH, Marks P, *et al.* (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* **10**: 563-569.

Christensen-Dalsgaard M, Jørgensen MG & Gerdes K (2010) Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Mol Microbiol* **75**: 333-348.

Christensen SK, Mikkelsen M, Pedersen K & Gerdes K (2001) RelE, a global inhibitor of translation, is activated during nutritional stress. *Proc Natl Acad Sci U S A* **98**: 14328-14333.

Coates JD, Ellis DJ, Gaw CV & Lovley DR (1999) Geothrix fermentans gen. nov., sp. nov., a novel Fe(III)-reducing bacterium from a hydrocarbon-contaminated aquifer. *Int J Syst Bacteriol* **49 Pt 4**: 1615-1622.

Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR & Tiedje JM (2014) Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* **42**: D633-D642.

Costa OYA, Raaijmakers JM & Kuramae EE (2018) Microbial extracellular polymeric substances: ecological function

and impact on soil aggregation. *Front Microbiol* **9**: 1636.

Coughlan LM, Cotter PD, Hill C & Alvarez-Ordóñez A (2015) Biotechnological applications of functional metagenomics in the food and pharmaceutical industries. *Front Microbiol* **6**.

Craigen B (2011) The Use of Commercially Available Alpha-Amylase Compounds to Inhibit and Remove Staphylococcus aureus Biofilms. *Open Microbiol J* **5**: 21-31.

Crowe MA, Power JF, Morgan XC*, et al.* (2013) *Pyrinomonas methylaliphatogenes* gen. nov., sp. nov., a novel group 4 thermophilic member of the phylum *Acidobacteria* from geothermal soils. *Int J Syst Evol Microbiol* **64**: 220-227.

Crowley JD, Traynor DA & Weatherburn DC (2000) Enzymes and proteins containing manganese: an overview. *Met Ions Biol Syst,* Vol. 37 209-278.

Cuthbertson L, Mainprize IL, Naismith JH & Whitfield C (2009) Pivotal roles of the outer membrane polysaccharide dxport and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. *Microbiol Mol Biol Rev* **73**: 155-177.

Cvetkovic A, Menon AL, Thorgersen MP*, et al.* (2010) Microbial metalloproteomes are largely uncharacterized. *Nature* **466**: 779-782.

Czarnes S, Hallett PD, Bengough AG & Young IM (2000) Root- and microbial-derived mucilages affect soil structure and water transport. *Eur J Soil Sci* **51**: 435-443.

D

Dahiya N, Tewari R, Tiwari RP & Hoondal GS (2005) Chitinase Production in Solid-State Fermentation by Enterobacter sp. NRG4 Using Statistical Experimental Design. *Curr Microbiol* **51**: 222-228.

Davenport EK, Call DR & Beyenal H (2014) Differential protection fromtobramycin by extracellular polymeric substances from *Acinetobacter baumannii* and *Staphylococcus aureus* biofilms. *Antimicrob Agents Chemother* **58**: 4755-4761.

de Castro VHL (2011) Identificação, isolamento e caracterização de bactérias de solo de cerrado pertencentes ao filo *Acidobacteria*. Dissertation Thesis, Universidade Católica de Brasília, Brasília.

de Castro VHL, Schroeder LF, Quirino BF, Kruger RH & Barreto CC (2013) *Acidobacteria* from oligotrophic soil from the Cerrado can grow in a wide range of carbon source concentrations. *Can J Microbiol* **59**: 746-753.

De Vuyst, Vanderveken, Van de V & Degeest (1998) Production by and isolation of exopolysaccharides from *Streptococcus thermophilus* grown in a milk medium and evidence for their growth-associated biosynthesis. *J Appl Microbiol* **84**: 1059-1068.

Dedysh S, Kulichevskaya I, Huber K & Overmann J (2017) Defining the taxonomic status of described subdivision 3 Acidobacteria: proposal of Bryobacteraceae fam. nov. *Int J Syst Evol Microbiol* **67**: 498-501.

Dedysh SN & Yilmaz P (2018) Refining the taxonomic structure of the phylum *Acidobacteria*. *Int J Syst Evol Microbiol* **68**: 3796-3806.

Dedysh SN, Kulichevskaya IS, Serkebaeva YM, Mityaeva MA, Sorokin VV, Suzina NE, Rijpstra WIC & Damste JSS (2012) Bryocella elongata gen. nov., sp nov., a member of subdivision 1 of the Acidobacteria isolated from a methanotrophic enrichment culture, and emended description of Edaphobacter aggregans Koch et al. 2008. *Int J Syst Evol Microbiol* **62**: 654-664.

Dimitrov MR, Veraart AJ, de Hollander M, Smidt H, van Veen JA & Kuramae EE (2017) Successive DNA extractions improve characterization of soil microbial communities. *PeerJ* **5**: e2915.

Dinel H, Lévesque PEM, Jambu P & Righi D (1992) Microbial activity and long-chain aliphatics in the formation of stable soil aggregates. *Soil Sci Soc Am J* **56**: 1455.

Dominguez-Ferreras A, Perez-Arnedo R, Becker A, Olivares J, Soto MJ & Sanjuan J (2006) Transcriptome profiling reveals the importance of plasmid pSymB for osmoadaptation of *Sinorhizobium meliloti*. *J Bacteriol* **188**: 7617-7625.

Doncheva NT, Assenov Y, Domingues FS & Albrecht M (2012) Topological analysis and interactive visualization of biological networks and protein structures. *Nat Protoc* **7**: 670-685.

Dray S & Dufour A-B (2007) The ade4 Package: implementing the duality diagram for ecologists. *J Stat Softw* **22**.

Druzhinina IS, Seidl-Seiboth V, Herrera-Estrella A, Horwitz BA, Kenerley CM, Monte E, Mukherjee PK, Zeilinger S, Grigoriev IV & Kubicek CP (2011) *Trichoderma*: the genomics of opportunistic success. *Nat Rev Microbiol* **9**: 749-

759.

DuBois M, Gilles KA, Hamilton JK, Rebers PA & Smith F (1956) Colorimetric method for determination of sugars and related substances. *Anal Chem* **28**: 350-356.

Dumitriu S (2005) *Polysaccharides: structural diversity and functional versatility*. Marcel Dekker, New York.

E

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460-2461.

Edgar RC, Haas BJ, Clemente JC, Quince C & Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194-2200.

Eichorst SA, Breznak JA & Schmidt TM (2007) Isolation and characterization of soil bacteria that define Terriglobus gen. nov., in the phylum Acidobacteria. *Appl Environ Microbiol* **73**: 2708-2717.

Eichorst SA, Kuske CR & Schmidt TM (2011) Influence of plant polymers on the distribution and cultivation of bacteria in the phylum *Acidobacteria*. *Appl Environ Microbiol* **77**: 586-596.

Eichorst SA, Trojan D, Roux S, Herbold C, Rattei T & Woebken D (2018) Genomic insights into the Acidobacteria reveal strategies for their success in terrestrial environments. *Environ Microbiol* **20**: 1041-1063.

Ekkers DM, Cretoiu MS, Kielak AM & van Elsas JD (2011) The great screen anomaly—a new frontier in product discovery through functional metagenomics. *Appl Microbiol Biotechnol* **93**: 1005-1020.

El-Naggar AH, Omar HH, Osman MEH & Ismail. GA (2008) Heavy metal binding capacity of Exopolysaccharides produced by *Anabaena variabilis* and *Nostoc muscorum*. *Egypt J Exp Biol* **4**: 6.

Elisashvili VI, Kachlishvili ET & Wasser SP (2009) Carbon and nitrogen source effects on basidiomycetes exopolysaccharide production. *Appl Biochem Micro* **45**: 531-535.

Entcheva-Dimitrov P & Spormann AM (2004) Dynamics and Control of Biofilms of the Oligotrophic Bacterium Caulobacter crescentus. *J Bacteriol* **186**: 8254-8266.

Everett JA & Rumbaugh KP (2015) Biofilms, quorum sensing and crosstalk in medically important microbes. *Molecular Medical Microbiology*,(Tang Y-W, Sussman M, Liu D, Poxton I & Schwartzman J, eds.), 235-247. Academic Press, Massachussets.

Ezeilo UR, Zakaria II, Huyop F & Wahab RA (2017) Enzymatic breakdown of lignocellulosic biomass: the role of glycosyl hydrolases and lytic polysaccharide monooxygenases. *Biotechnol Biotechnol Equip* 1-16.

F

Falagán C, Foesel B & Johnson B (2017) *Acidicapsa ferrireducens* sp. nov., *Acidicapsa acidiphila* sp. nov., and *Granulicella acidiphila* sp. nov.: novel acidobacteria isolated from metal-rich acidic waters. *Extremophiles* **21**: 459-469.

Farber BF, Kaplan MH & Clogston AG (1990) *Staphylococcus epidermidis* extracted slime inhibits the antimicrobial action of glycopeptide antibiotics. *Int J Infect Dis* **161**: 37-40.

Faria M, Bordin N, Kizina J, Harder J, Devos D & Lage OM (2018) *Planctomycetes* attached to algal surfaces: Insight into their genomes. *Genomics* **110**: 231-238.

Farrés J, Caminal G & López-Santín J (1997) Influence of phosphate on rhamnose-containing exopolysaccharide rheology and production by *Klebsiella* I-714. *Appl Microbiol Biotechnol* **48**: 522-527.

Federici L, Du D, Walas F, Matsumura H, Fernandez-Recio J, McKeegan KS, Borges-Walmsley MI, Luisi BF & Walmsley AR (2005) The Crystal Structure of the Outer Membrane Protein VceC from the Bacterial PathogenVibrio choleraeat 1.8 Å Resolution. *J Biol Chem* **280**: 15307-15314.

Fierer N, Morse JL, Berthrong ST, Bernhardt ES & Jackson RB (2007) Environmental Controls on the Landscape-Scale Biogeography of Stream Bacterial Communities. *Ecology* **88**: 2162-2173.

Figueras MJ, Beaz-Hidalgo R, Hossain MJ & Liles MR (2014) Taxonomic affiliation of new genomes should be verified using average nucleotide identity and multilocus phylogenetic analysis. *Genome Announc* **2**.

Fitzpatrick JJ, Ahrens M & Smith S (2001) Effect of manganese on *Lactobacillus casei* fermentation to produce lactic acid from whey permeate. *Process Biochem* **36**: 671-675.

Fleming D, Chahin L & Rumbaugh K (2016) Glycoside Hydrolases Degrade Polymicrobial Bacterial Biofilms in Wounds. *Antimicrob Agents Chemother*.

Flemming H-C & Wingender J (2010) The biofilm matrix. *Nat Rev Microbiol* 623-633.

Flemming H-C, Neu TR & Wozniak DJ (2007) The EPS matrix: the "house of biofilm cells". *J Bacteriol* **189**: 7945-7947.

Flemming H-C, Wingender J, Mayer C, Korstgens V & Borchard W (2000) Cohesiveness in biofilm matrix polymers. *Symposium of the Society for General Microbiology* **59th**: 87-105.

Flemming H-C, Wingender J, Szewzyk U, Steinberg P, Rice SA & Kjelleberg S (2016) Biofilms: an emergent form of bacterial life. *Nat Rev Microbiol* **14**: 563-575.

Foesel BU, Rohde M & Overmann J (2013) *Blastocatella fastidiosa* gen. nov., sp. nov., isolated from semiarid savanna soil–The first described species of *Acidobacteria* subdivision 4. *Syst Appl Microbiol* **36**: 82-89.

Foesel BU, Wanner G, Mayer S, Rohde M, Overmann J & Luckner M (2016) Occallatibacter riparius gen. nov., sp. nov. and Occallatibacter savannae sp. nov., acidobacteria isolated from Namibian soils, and emended description of the family Acidobacteriaceae. *Int J Syst Evol Microbiol* **66**: 219-229.

Foesel BU, Nägele V, Naether A, *et al.* (2014) Determinants of Acidobacteria activity inferred from the relative abundances of 16S rRNA transcripts in German grassland and forest soils. *Environ Microbiol* **16**: 658-675.

Forster SM (1979) Microbial aggregation of sand in an embryo dune system. *Soil Biol Biochem* **11**: 537-543.

Forster SM & Nicolson TH (1981) Aggregation of sand from a maritime embryo sand dune by microorganisms and higher plants. *Soil Biol Biochem* **13**: 199-203.

Friedman J & Alm EJ (2012) Inferring Correlation Networks from Genomic Survey Data. *PLoS Comput Biol* **8**.

Fu XT & Kim SM (2010) Agarase: Review of Major Sources, Categories, Purification Method, Enzyme Characteristics and Applications. *Mar Drugs* **8**: 200-218.

Fuchs EL, Brutinel ED, Klem ER, Fehr AR, Yahr TL & Wolfgang MC (2010) In vitro and in vivo vharacterization of the *Pseudomonas aeruginosa* Cyclic AMP (cAMP) Phosphodiesterase CpdA, required for cAMP homeostasis and virulence factor regulation. *J Bacteriol* **192**: 2779-2790.

Fujita Y (2014) Carbon Catabolite Control of the Metabolic Network in Bacillus subtilis. *Biosci Biotechnol Biochem* **73**: 245-259.

Fukunaga Y, Kurahashi M, Yanagi K, Yokota A & Harayama S (2008) Acanthopleuribacter pedis gen. nov., sp. nov., a marine bacterium isolated from a chiton, and description of Acanthopleuribacteraceae fam. nov., Acanthopleuribacterales ord. nov., Holophagaceae fam. nov., Holophagales ord. nov. and Holophagae classis nov. in the phylum 'Acidobacteria'. *Int J Syst Evol Microbiol* **58**: 2597-2601.

G

Galperin MY, Makarova KS, Wolf YI & Koonin EV (2015) Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* **43**: D261-D269.

Gamar-Nourani L, Blondeau K & Simonet JM (1998) Influence of culture conditions on exopolysaccharide production by *Lactobacillus rhamnosus* strain C83. *J Appl Microbiol* **85**: 664-672.

Gandhi HP, Ray RM & Patel RM (1997) Exopolymer production by *Bacillus* species. *Carbohydr Polym* **34**: 323-327.

García-Fraile P, Benada O, Cajthaml T, Baldrian P, Lladó S & Löffler FE (2016) Terracidiphilus gabretensis gen. nov., sp. nov., an Abundant and Active Forest Soil Acidobacterium Important in Organic Matter Transformation. *Appl Environ Microbiol* **82**: 560-569.

Gasperi-Mago RR & Troeh FR (1979) Microbial Effects on Soil Erodibility1. *Soil Sci Soc Am J* **43**.

Geoghegan MJ & Brian RC (1948) Aggregate formation in soil. 1. Influence of some bacterial polysaccharides on the binding of soil particles. *Biochem J* **43**: 5-13.

Gillard B, Chatzievangelou D, Thomsen L & Ullrich MS (2019) Heavy-metal-resistant microorganisms in deep-sea sediments disturbed by mining activity: an application toward the development of experimental in vitro systems. *Front Mar Sci* **6**.

Ginige MP, Hugenholtz P, Daims H, Wagner M, Keller J & Blackall LL (2004) Use of stable-isotope probing, full-cycle rRNA analysis, and fluorescence in situ hybridization-microautoradiography to study a methanol-fed denitrifying microbial community. *Appl Environ Microbiol* **70**: 588-596.

Givry S & Duchiron F (2008) Optimization of culture medium and growth conditions for production of L-arabinose isomerase and D-xylose isomerase by *Lactobacillus bifermentans*. *Microbiology* **77**: 281-287.

Godinho AL & Bhosle S (2009) Sand aggregation by exopolysaccharide-producing *Microbacterium arborescens*–– AGSB. *Curr Microbiol* **58**: 616-621.

González-García Y, Heredia A, Meza-Contreras JC, Escalante FME, Camacho-Ruiz RM & Córdova J (2015) Biosynthesis of extracellular polymeric substances by the marine bacterium *Saccharophagus degradans* under different nutritional conditions. *Int J Polym Sci* **2015**: 1-7.

Gorla P, Plocinska R, Sarva K, Satsangi AT, Pandeeti E, Donnelly R, Dziadek J, Rajagopalan M & Madiraju MV (2018) MtrA response regulator controls cell division and cell wall metabolism and affects susceptibility of mycobacteria to the first line antituberculosis drugs. *Front Microbiol* **9**.

Grainge I & Sherratt DJ (1999) Xer Site-specific Recombination. *J Biol Chem* **274**: 6763-6769.

Gremion F, Chatzinotas A & Harms H (2003) Comparative 16S rDNA and 16S rRNA sequence analysis indicates that Actinobacteria might be a dominant part of the metabolically active bacteria in heavy metal-contaminated bulk and rhizosphere soil. *Environ Microbiol* **5**: 896-907.

Grobe S, Wingender J & Flemming H-C (2001) Capability of mucoid *Pseudomonas aeruginosa* to survive in chlorinated water. *Int J Hyg Envir Health* **204**: 139-142.

Grube M, Cernava T, Soh J*, et al.* (2014) Exploring functional contexts of symbiotic sustain within lichen-associated bacteria by comparative omics. *ISME J* **9**: 412-424.

Guedon E, Moore CM, Que Q, Wang T, Ye RW & Helmann JD (2003) The global transcriptional response of *Bacillus subtilis* to manganese involves the MntR, Fur, TnrA and σB regulons. *Mol Microbiol* **49**: 1477-1491.

Guldimann C, Boor KJ, Wiedmann M, Guariglia-Oropeza V & Schaffner DW (2016) Resilience in the face of uncertainty: sgma factor B fine-tunes gene expression to support homeostasis in gram-positive bacteria. *Appl Environ Microbiol* **82**: 4456-4469.

Guo H, Nasir M, Lv J, Dai Y & Gao J (2017) Understanding the variation of microbial community in heavy metals contaminated soil using high throughput sequencing. *Ecotox Environ Safe* **144**: 300-306.

Guo Y, Zhao H, Zuo X, Drake S & Zhao X (2007) Biological soil crust development and its topsoil properties in the process of dune stabilization, Inner Mongolia, China. *Environ Geol* **54**: 653-662.

Gvakharia BO, Permina EA, Gelfand MS, Bottomley PJ, Sayavedra-Soto LA & Arp DJ (2007) Global Transcriptional Response of *Nitrosomonas europaea* to Chloroform and Chloromethane. *Appl Environ Microbiol* **73**: 3440-3445.

H

Haichar FeZ, Achouak W, Christen R, Heulin T, Marol C, Marais M-F, Mougel C, Ranjard L, Balesdent J & Berge O (2007) Identification of cellulolytic bacteria in soil by stable isotope probing. *Environ Microbiol* **9**: 625-634.

Hantke K (2001) Iron and metal regulation in bacteria. *Curr Opin Microbiol* **4**: 5.

Harkes P, Suleiman A, van den Elsen S, de Haan J, Holterman M, Kuramae E & Helder J (2019) Mapping of long-term impact of conventional and organic soil management on resident and active fractions of rhizosphere communities of barley (Hordeum vulgare). *bioRxiv* 546192.

Hausmann B, Pelikan C, Herbold CW*, et al.* (2018) Peatland Acidobacteria with a dissimilatory sulfur metabolism. *ISME J* **12**: 1729-1742.

Hausner M & Wuertz S (1999) High rates of conjugation in bacterial biofilms as determined by quantitative in situ analysis. *Appl Environ Microbiol* **65**: 3710-3713.

Hayes CS & Low DA (2009) Signals of growth regulation in bacteria. *Curr Opin Microbiol* **12**: 667-673.

Helmann JD (2016) Bacillus subtilis extracytoplasmic function (ECF) sigma factors and defense of the cell envelope. *Curr Opin Microbiol* **30**: 122-132.

Hendrickx L, Hausner M & Wuertz S (2003) Natural genetic transformation in monoculture *Acinetobacter* sp. strain BD413 biofilms. *Appl Environ Microbiol* **69**: 1721-1727.

Herve M, Boniface A, Gobec S, Blanot D & Mengin-Lecreulx D (2007) Biochemical characterization and physiological properties of *Escherichia coli* UDP-N-Acetylmuramate:L-Alanyl- -D-Glutamyl-meso- Diaminopimelate Ligase. *J Bacteriol* **189**: 3987-3995.

Hess M, Sczyrba A, Egan R*, et al.* (2011) Metagenomic Discovery of Biomass-Degrading Genes and Genomes from Cow Rumen. *Science* **331**: 463-467.

Ho A, Lonardo DPD & Bodelier PLE (2017) Revisiting life strategy concepts in environmental microbial ecology. *FEMS*

Microbiol Ecol.

Hollander MD (2017) Nioo-Knaw/Hydra: 1.3.3.

Houssin C, Eynard N, Shechter E & Ghazi A (1991) Effect of osmotic pressure on membrane energy-linked functions in Escherichia coli. *Biochim Biophys Acta Bioenerg* **1056**: 76-84.

Hubbard C, McNamara JT, Azumaya C, Patel MS & Zimmer J (2012) The hyaluronan synthase catalyzes the synthesis and membrane translocation of hyaluronan. *J Mol Biol* **418**: 21-31.

Huber KJ & Overmann J (2018) Vicinamibacteraceae fam. nov., the first described family within the subdivision 6 Acidobacteria. *Int J Syst Evol Microbiol* **68**: 2331-2334.

Huber KJ, Wüst PK, Rhode M, Overmann J & Foesel BU (2014) Aridibacter famidurans and Aridibacter kavangonensis, 2 novel species of Acidobacteria subdivision 4 isolated from semiarid savanna soil. *Int J Syst Evol Microbiol* ijs. 0.060236-060230.

Huber KJ, Geppert AM, Wanner G, Fösel BU, Wüst PK & Overmann J (2016) The first representative of the globally widespread subdivision 6 Acidobacteria, Vicinamibacter silvestris gen. nov., sp. nov., isolated from subtropical savannah soil. *Int J Syst Evol Microbiol* **66**: 2971-2979.

Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C & Bork P (2017) Fast Genome-Wide functional annotation through Orthology Assignment by eggNOG-Mapper. *Mol Biol Evol* **34**: 2115-2122.

Huerta-Cepas J, Szklarczyk D, Forslund K, *et al.* (2016) eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* **44**: D286-D293.

Hugenholtz P, Goebel BM & Pace NR (1998) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* **180**: 4765-4774.

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 D Earth Sci* **50**: 128-134.

Hukić M, Seljmo D, Ramovic A, Ibrišimović MA, Dogan S, Hukic J & Bojic EF (2018) The Effect of Lysozyme on Reducing Biofilms by Staphylococcus aureus, Pseudomonas aeruginosa, and Gardnerella vaginalis: An In Vitro Examination. *Microb Drug Resist* **24**: 353-358.

Hunter P (2008) The mob response. The importance of biofilm research for combating chronic diseases and tackling contamination. *EMBO Rep* **9**: 314-317.

Hwang HJ, Kim SW, Xu CP, Choi JW & Yun JW (2004) Morphological and rheological properties of the three different species of basidiomycetes *Phellinus* in submerged cultures. *J Appl Microbiol* **96**: 1296-1305.

Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW & Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**.

I

Ivanova AA, Wegner C-E, Kim Y, Liesack W & Dedysh SN (2016) Identification of microbial populations driving biopolymer degradation in acidic peatlands by metatranscriptomic analysis. *Mol Ecol* **25**: 4818-4835.

Ivanova AA, Wegner C-E, Kim Y, Liesack W & Dedysh SN (2017) Metatranscriptomics reveals the hydrolytic potential of peat-inhabiting *Planctomycetes*. *Antonie Van Leeuwenhoek* **111**: 801-809.

Ivanova AA, Naumoff DG, Miroshnikov KK, Liesack W & Dedysh SN (2017) Comparative genomics of four *Isosphaeraceae Planctomycetes*: a common pool of plasmids and glycoside hydrolase genes shared by *Paludisphaera borealis* PX4T, *Isosphaera pallida* IS1BT, *Singulisphaera acidiphila* DSM 18658T, and strain SH-PL62. *Front Microbiol* **8**: 801–809.

Ivanova AO & Dedysh SN (2012) Abundance, diversity, and depth distribution of *Planctomycetes* in Acidic Northern Wetlands. *Front Microbiol* **3**: 9.

Izano EA, Amarante MA, Kher WB & Kaplan JB (2007) Differentialroles of poly-n-acetylglucosamine surface polysaccharide and extracellular DNA in *Staphylococcus aureus* and *Staphylococcus epidermidis* biofilms. *Appl Environ Microbiol* **74**: 470-476.

Izumi H, Nunoura T, Miyazaki M, Mino S, Toki T, Takai K, Sako Y, Sawabe T & Nakagawa S (2012) Thermotomaculum hydrothermale gen. nov., sp. nov., a novel heterotrophic thermophile within the phylum Acidobacteria from a deep-sea hydrothermal vent chimney in the Southern Okinawa Trough. *Extremophiles* **16**: 245-253.

J

Jain R, Raghukumar S, Tharanathan R & Bhosle NB (2005) Extracellular polysaccharide production by Thraustochytrid protists. *Mar Biotechnol* **7**: 184-192.

Jakubovics NS & Jenkinson HF (2001) Out of the iron age: new insights into the critical role of manganese homeostasis in bacteria. *Microbiology* **147**: 1709-1718.

Janczarek M (2011) Environmental signals and regulatory pathways that influence exopolysaccharide production in *Rhizobia*. *Int J Mol Sci* **12**: 7898-7933.

Jankovic I & Bruckner R (2002) Carbon catabolite repression by the catabolite control protein CcpA in *Staphylococcus xylosus*. *J Mol Microbiol Biotechnol* **4**: 309-314.

Janssen PH (2006) Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl Environ Microbiol* **72**: 1719-1728.

Janssen PH, Yates PS, Grinton BE, Taylor PM & Sait M (2002) Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions *Acidobacteria*, *Actinobacteria*, *Proteobacteria*, and *Verrucomicrobia*. *Appl Environ Microbiol* **68**: 2391-2396.

Jeffries P (1999) Scleroderma. *Ectomycorrhizal Fungi Key Genera in Profile,*187-200.

Jensen AN & Jensen LT (2014) CHAPTER 1. Manganese transport, trafficking and function in invertebrates. 1-33.

Johnson JM, Ludwig A, Furch A, Mithöfer A, Scholz SS, Reichelt M & Oelmüller R (2018) The beneficial root-colonizing fungus *Mortierella* hyalina promotes the aerial growth of Arabidopsis and activates calcium-dependent responses which restrict Alternaria brassicae-induced disease development in roots. *Mol Plant Microbe Interact* 351-363.

Jorgensen MG, Pandey DP, Jaskolska M & Gerdes K (2008) HicA of *Escherichia coli* defines a novel family of translation-independent mRNA interferases in *Bacteria* and *Archaea*. *J Bacteriol* **191**: 1191-1199.

Joshi PA, Singh N & Shekhawat DB (2015) Effect of metal ions on growth and biosurfactant production by halophilic bacteria *Adv Appl Sci Res* **6**: 4.

Joubert LM, Wolfaardt GM & Botha A (2006) Microbial exopolymers link predator and prey in a model yeast biofilm system. *Microb Ecol* **52**: 187-197.

K

Kaci Y, Heyraud A, Barakat M & Heulin T (2005) Isolation and identification of an EPS-producing *Rhizobium* strain from arid soil (Algeria): characterization of its EPS and the effect of inoculation on wheat rhizosphere soil structure. *Res Microbiol* **156**: 522-531.

Kalogiannis S, Iakovidou G, Liakopoulou-Kyriakides M, Kyriakidis DA & Skaracis GN (2003) Optimization of xanthan gum production by *Xanthomonas campestris* grown in molasses. *Process Biochem* **39**: 249-256.

Kambourova M, Mandeva R, Dimova D, Poli A, Nicolaus B & Tommonaro G (2009) Production and characterization of a microbial glucan, synthesized by *Geobacillus tepidamans* V264 isolated from Bulgarian hot spring. *Carbohydr Polym* **77**: 338-343.

Kanehisa M (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* **28**: 27-30.

Kang DD, Froula J, Egan R & Wang Z (2015) MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**.

Kawaharada Y, Kelly S, Nielsen MW*, et al.* (2015) Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature* **523**: 308-312.

Kehr J-C & Dittmann E (2015) Biosynthesis and function of extracellular glycans in Cyanobacteria. *Life* **5**: 164-180.

Kehres DG & Maguire ME (2003) Emerging themes in manganese transport, biochemistry and pathogenesis in bacteria. *FEMS Microbiol Rev* **27**: 263-290.

Keiski C-L, Harwich M, Jain S*, et al.* (2010) AlgK Is a TPR-containing protein and the periplasmic component of a novel exopolysaccharide secretin. *Structure* **18**: 265-273.

Khan RH, Du L, Pang H, Wang Z, Lu J, Wei Y & Huang R (2013) Characterization of an Invertase with pH Tolerance and Truncation of Its N-Terminal to Shift Optimum Activity toward Neutral pH. *PLoS One* **8**.

Khani M, Bahrami A, Chegeni A, Ghafari MD & Mansouran Zadeh A (2016) Optimization of carbon and nitrogen sources for extracellular polymeric substances production by *Chryseobacterium indologenes* MUT.2. *Iran J Biotechnol* **14**: 13-18.

Kielak AM, Cipriano MAP & Kuramae EE (2016) Acidobacteria strains from subdivision 1 act as plant growth-promoting bacteria. *Arch Microbiol*.

Kielak AM, Barreto CC, Kowalchuk GA, van Veen JA & Kuramae EE (2016) The ecology of acidobacteria: moving beyond genes and genomes. *Front Microbiol* **7**: 16.

Kielak AM, Scheublin TR, Mendes LW, van Veen JA & Kuramae EE (2016) Bacterial community succession in pine-wood decomposition. *Front Microbiol* **7**.

Kielak AM, Castellane TCL, Campanharo JC, Colnago LA, Costa OYA, Corradi da Silva ML, van Veen JA, Lemos EGM & Kuramae EE (2017) Characterization of novel *Acidobacteria* exopolysaccharides with potential industrial and ecological applications. *Sci Rep* **7**: 41193.

Kim D & Robyt JF (1994) Production and selection of mutants of *Leuconostoc mesenteroides* constitutive for glucansucrases. *Enzyme Microb Technol* **16**: 659-664.

Kim HS, Nagore D & Nikaido H (2009) Multidrug Efflux Pump MdtBC of *Escherichia coli* Is active only as a B2C heterotrimer. *J Bacteriol* **192**: 1377-1386.

Király Z, El-Zahaby HM & Klement Z (1997) Role of extracellular polysaccharide (EPS) slime of plant pathogenic bacteria in protecting cells to reactive oxygen species. *J Phytopathol* **145**: 59-68.

Kishimoto N, Kosako Y & Tano T (1991) *Acidobacterium capsulatum* gen. nov., sp. nov.: an acidophilic chemoorganotrophic bacterium containing menaquinone from acidic mineral environment. *Curr Microbiol* **22**: 1-7.

Koch IH, Gich F, Dunfield PF & Overmann J (2008) Edaphobacter modestus gen. nov., sp. nov., and Edaphobacter aggregans sp. nov., acidobacteria isolated from alpine and forest soils. *Int J Syst Evol Microbiol* **58**: 1114-1122.

Koczan JM, McGrath MJ, Zhao Y & Sundin GW (2009) Contribution of *Erwinia amylovora* exopolysaccharides Amylovoran and Levan to biofilm formation: implications in pathogenicity. *Phytopathology* **99**: 1237-1244.

Kohler J, Caravaca F, Carrasco L & Roldán A (2006) Contribution of Pseudomonas mendocina and Glomus intraradices to aggregate stabilization and promotion of biological fertility in rhizosphere soil of lettuce plants under field conditions. *Soil Use Manage* **22**: 298-304.

Kolde R (2019) pheatmap: Pretty Heatmaps. R package version 1.0.12.

Kõljalg U, Nilsson RH, Abarenkov K, *et al.* (2013) Towards a unified paradigm for sequence-based identification of fungi. *Mol Ecol* **22**: 5271-5277.

Konan HK, Yapi D, Bi CYY, Koné TFM, Kouadio PEJN & Patrice K (2016) Biochemical characterization of two acid phosphatases purified from breadfruit (Artocarpus communis) seeds. *J Adv Biol Biotechnol* **43**: 1102-1113.

Koster J & Rahmann S (2012) Snakemake--a scalable bioinformatics workflow engine. *Bioinformatics* **28**: 2520-2522.

Kravchenko AN, Negassa WC, Guber AK, Hildebrandt B, Marsh TL & Rivers ML (2014) Intra-aggregate Pore Structure Influences Phylogenetic Composition of Bacterial Community in Macroaggregates. *Soil Sci Soc Am J* **78**: 1924-1939.

Krembs C, Eicken H & Deming JW (2011) Exopolymer alteration of physical properties of sea ice and implications for ice habitability and biogeochemistry in a warmer Arctic. *Proc Natl Acad Sci U S A* **108**: 3653-3658.

Krembs C, Eicken H, Junge K & Deming JW (2002) High concentrations of exopolymeric substances in Arctic winter sea ice: implications for the polar ocean carbon cycle and cryoprotection of diatoms. *Deep-Sea Res Pt I* **49**: 2163-2181.

Küçük Ç & Kivanç M (2009) Extracellular polysaccharide production by *Rhizobium ciceri* from Turkey. *Ann Microbiol* **59**: 141-144.

Kulichevskaya IS, Suzina NE, Liesack W & Dedysh SN (2010) Bryobacter aggregatus gen. nov., sp. nov., a peat-inhabiting, aerobic chemo-organotroph from subdivision 3 of the Acidobacteria. *Int J Syst Evol Microbiol* **60**: 301-306.

Kulichevskaya IS, Suzina NE, Rijpstra WIC, Damste JSS & Dedysh SN (2014) Paludibaculum fermentans gen. nov., sp. nov., a facultative anaerobe capable of dissimilatory iron reduction from subdivision 3 of the Acidobacteria. *Int J Syst Evol Microbiol* **64**: 2857-2864.

Kulichevskaya IS, Ivanova AO, Baulina OI, Bodelier PLE, Damste JSS & Dedysh SN (2008) *Singulisphaera acidiphila* gen. nov., sp. nov., a non-filamentous, *Isosphaera*-like planctomycete from acidic northern wetlands. *Int J Syst*

Evol Microbiol **58**: 1186-1193.

Kulichevskaya IS, Kostina LA, Valaskova V, Rijpstra WI, Damste JS, de Boer W & Dedysh SN (2012) *Acidicapsa borealis* gen. nov., sp. nov. and *Acidicapsa ligni* sp. nov., subdivision 1 *Acidobacteria* from Sphagnum peat and decaying wood. *Int J Syst Evol Microbiol* **62**: 1512-1520.

Kullik I & Giachino P (1997) The alternative sigma factor σ B in *Staphylococcus aureus* : regulation of the sigB operon in response to growth phase and heat shock. *Arch Microbiol* **167**: 151-159.

Kumar AS, Mody K & Jha B (2007) Bacterial exopolysaccharides – a perception. *J Basic Microbiol* **47**: 103-117.

Kumar GK & Ram MR (2014) Effect of carbon and nitrogen sources on exopolysacharide production by rhizobial isolates from root nodules of *Vigna trilobata*. *Afr J Microbiol Res* **8**: 2255-2260.

Kumar R, Rawat KS, Singh J, Singh A & Rai A (2013) Soil aggregation dynamics and carbon sequestration. *J Nat Appl Sci* **5**: 250-267.

Kuramae EE, Yergeau E, Wong LC, Pijl AS, Veen JA & Kowalchuk GA (2012) Soil characteristics more strongly influence soil bacterial communities than land-use type. *FEMS Microbiol Ecol* **79**: 12-24.

Kuramae EE, Hillekens RHE, de Hollander M, van der Heijden MGA, van den Berg M, van Straalen NM & Kowalchuk GA (2013) Tracking Fungal Community Responses to Maize Plants by DNA- and RNA-Based Pyrosequencing. *PLoS One* **8**: e69973.

Kursa MB & Rudnicki WR (2010) Feature Selection with theBorutaPackage. *J Stat Softw* **36**.

Kuske CR, Barns SM & Busch JD (1997) Diverse uncultivated bacterial groups from soils of the arid southwestern United States that are present in many geographic regions. *Appl Environ Microbiol* **63**: 3614-3621.

Kuznetsova A, Brockhoff PB & Christensen RHB (2017) lmerTest package: tests in Linear Mixed Effects Models. *J Stat Softw* **82**.

L

Labes A, Karlsson EN, Fridjonsson OH, Turner P, Hreggvidson GO, Kristjansson JK, Holst O & Schonheit P (2008) Novel Members of Glycoside Hydrolase Family 13 Derived from Environmental DNA. *Appl Environ Microbiol* **74**: 1914-1921.

Lacombe-Harvey M-È, Brzezinski R & Beaulieu C (2018) Chitinolytic functions in actinobacteria: ecology, enzymes, and evolution. *Appl Microbiol Biotechnol* **102**: 7219-7230.

Lairson LL, Henrissat B, Davies GJ & Withers SG (2008) Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu Rev Biochem* **77**: 521-555.

Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357-359.

Lau TC, Wu XA, Chua H, Qian P & Wong PK (2005) Effect of exopolysaccharides on the adsorption of metal ions by *Pseudomonas* Sp CU-1. *Water Sci Technol* **52**: 6.

Lee TK, Lee J, Sul WJ, Iwai S, Chai B, Tiedje JM & Park J (2011) Novel Biphenyl-Oxidizing Bacteria and Dioxygenase Genes from a Korean Tidal Mudflat. *Appl Environ Microbiol* **77**: 3888-3891.

Legendre P & Gallagher ED (2001) Ecologically meaningful transformations for ordination of species data. *Oecol* **129**: 271-280.

Legin E, Ladrat C, Godfroy A, Barbier G & Duchiron F (1997) Thermostable amylolytic enzymes of thermophilic microorganisms from deep-sea hydrothermal vents. *Comptes Rendus Acad Sci* **320**: 893-898.

Lehman AP & Long SR (2013) Exopolysaccharides from *Sinorhizobium meliloti* can protect against H_2O_2-Dependent Damage. *J Bacteriol* **195**: 5362-5369.

Leutner BF, Reineking B, Müller J, Bachmann M, Beierkuhnlein C, Dech S & Wegmann M (2012) Modelling Forest α-Diversity and Floristic Composition — On the Added Value of LiDAR plus Hyperspectral Remote Sensing. *Remote Sens* **4**: 2818-2845.

Lew LC, Liong MT & Gan CY (2013) Growth optimization of *Lactobacillus rhamnosus* FTDC 8313 and the production of putative dermal bioactives in the presence of manganese and magnesium ions. *J Appl Microbiol* **114**: 526-535.

Li B & Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**.

Li D, Liu C-M, Luo R, Sadakane K & Lam T-W (2015) MEGAHIT: an ultra-fast single-node solution for large and complex

metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**: 1674-1676.

Li Y & Zhou H (2009) tRNAs as regulators in gene expression. *Sci China C Life Sci* **52**: 245-252.

Li Z, Gao Y, Nakanishi H, Gao X & Cai L (2013) Biosynthesis of rare hexoses using microorganisms and related enzymes. *Beilstein J Org Chem* **9**: 2434-2445.

Liang D-M, Liu J-H, Wu H, Wang B-B, Zhu H-J & Qiao J-J (2015) Glycosyltransferases: mechanisms and applications in natural product development. *Chem Soc Rev* **44**: 8350-8374.

Liao B, Yan X, Zhang J, Chen M, Li Y, Huang J, Lei M, He H & Wang J (2019) Microbial community composition in alpine lake sediments from the Hengduan Mountains. *Microbiologyopen* **8**.

Liesack W, Bak F, Kreft JU & Stackebrandt E (1994) Holophaga Foetida Gen-Nov, Sp-Nov, a New, Homoacetogenic Bacterium Degrading Methoxylated Aromatic-Compounds. *Arch Microbiol* **162**: 85-90.

Liu H & Fang HH (2002) Extraction of extracellular polymeric substances (EPS) of sludges. *J Biotechnol* **95**: 249-256.

Liu QP, Sulzenbacher G, Yuan H*, et al.* (2007) Bacterial glycosidases for the production of universal red blood cells. *Nature Biotechnol* **25**: 454-464.

Liu X & Kokare C (2017) Microbial Enzymes of Use in Industry. *Biotechnology of Microbial Enzymes,*267-298.

Lladó S, Benada O, Cajthaml T, Baldrian P & García-Fraile P (2016) Silvibacterium bohemicum gen. nov. sp. nov., an acidobacterium isolated from coniferous soil in the Bohemian Forest National Park. *Syst Appl Microbiol* **39**: 14-19.

Lloret J, Wulff BB, Rubio JM, Downie JA, Bonilla I & Rivilla R (1998) Exopolysaccharide II production is regulated by salt in the halotolerant strain *Rhizobium meliloti* EFB1. *Appl Environ Microbiol* **64**: 1024-1028.

Locatelli FM, Goo KS & Ulanova D (2016) Effects of trace metal ions on secondary metabolism and the morphological development of streptomycetes. *Metallomics* **8**: 469-480.

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM & Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**: D490-D495.

López-López O, Knapik K, Cerdán M-E & González-Siso M-I (2015) Metagenomics of an Alkaline Hot Spring in Galicia (Spain): Microbial Diversity Analysis and Screening for Novel Lipolytic Enzymes. *Front Microbiol* **6**.

Losey NA, Stevenson BS, Busse HJ, Sinninghe Damste JS, Rijpstra WI, Rudd S & Lawson PA (2013) Thermoanaerobaculum aquaticum gen. nov., sp. nov., the first cultivated member of Acidobacteria subdivision 23, isolated from a hot spring. *Int J Syst Evol Microbiol* **63**: 4149-4157.

Ludwig W, Bauer SH, Bauer M, Held I, Kirchhof G, Schulze R, Huber I, Spring S, Hartmann A & Schleifer KH (1997) Detection and in situ identification of representatives of a widely distributed new bacterial phylum. *FEMS Microbiol Lett* **153**: 181-190.

Lupatini M, Suleiman AKA, Jacques RJS, Lemos LN, Pylro VS, Van Veen JA, Kuramae EE & Roesch LFW (2019) Moisture Is More Important than Temperature for Assembly of Both Potentially Active and Whole Prokaryotic Communities in Subtropical Grassland. *Microb Ecol*.

Lynch JM (1981) Promotion and inhibition of soil aggregate stabilization by micro-organisms. *J Gen Microbiol* **126**: 4.

M

Mader U, Homuth G, Scharf C, Buttner K, Bode R & Hecker M (2002) Transcriptome and proteome analysis of *Bacillus subtilis* gene expression modulated by amino acid availability. *J Bacteriol* **184**: 4288-4295.

Madsen EL (2006) The use of stable isotope probing techniques in bioreactor and field studies on bioremediation. *Curr Opin Biotechnol* **17**: 92-97.

Makowski D (2018) The psycho package: an efficient and publishing-oriented workflow for psychological science. *J Open Source Softw* **3**.

Malam Issa O, Trichet J, Défarge C, Couté A & Valentin C (1999) Morphology and microstructure of microbiotic soil crusts on a tiger bush sequence (Niger, Sahel). *Catena* **37**: 175-196.

Malam Issa O, Défarge C, Le Bissonnais Y, Marin B, Duval O, Bruand A, D'Acqui LP, Nordenberg S & Annerman M (2006) Effects of the inoculation of cyanobacteria on the microstructure and the structural stability of a tropical soil. *Plant Soil* **290**: 209-219.

Mallela K, Talens-Perales D, Górska A, Huson DH, Polaina J & Marín-Navarro J (2016) Analysis of Domain Architecture and Phylogenetics of Family 2 Glycoside Hydrolases (GH2). *PLoS One* **11**.

Manca MC, Lama L, Improta R, Esposito E, Gambacorta A & Nicolaus B (1996) Chemical composition of two exopolysaccharides from *Bacillus thermoantarcticus*. *Appl Environ Microbiol* **62**: 3265-3269.

Mannisto MK, Rawat S, Starovoytov V & Haggblom MM (2012) *Granulicella arctica* sp. nov., *Granulicella mallensis* sp. nov., *Granulicella tundricola* sp. nov. and *Granulicella sapmiensis* sp. nov., novel acidobacteria from tundra soil. *Int J Syst Evol Microbiol* **62**: 2097-2106.

Martin JP (1971) Decomposition and binding action of polysaccharides in soil. *Soil Biol Biochem* **3**: 33-41.

Martin JP & Richards SJ (1963) Decomposition and binding action of a polysaccharide from *Chromobacterium violacium* in soil. *J Bacteriol* **85**: 1288-1294.

Martinez-Garcia M, Brazel DM, Swan BK, *et al.* (2012) Capturing single cell genomes of active polysaccharide degraders: an unexpected contribution of *Verrucomicrobia*. *PLoS One* **7**: e35314.

Martinez D, Berka RM, Henrissat B, *et al.* (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. Hypocrea jecorina). *Nature Biotechnol* **26**: 553-560.

Marx JG, Carpenter SD & Deming JW (2009) Production of cryoprotectant extracellular polysaccharide substances (EPS) by the marine psychrophilic bacterium *Colwellia psychrerythraea* strain 34H under extreme conditions. *Can J Microbiol* **55**: 63-72.

Masuko T, Minami A, Iwasaki N, Majima T, Nishimura S-I & Lee YC (2005) Carbohydrate analysis by a phenol–sulfuric acid method in microplate format. *Anal Biochem* **339**: 69-72.

Matsutani M, Ito K, Azuma Y, Ogino H, Shirai M, Yakushi T & Matsushita K (2015) Adaptive mutation related to cellulose producibility in *Komagataeibacter medellinensis* (*Gluconacetobacter xylinus*) NBRC 3288. *Appl Microbiol Biotechnol* **99**: 7229-7240.

McCalla TM, Haskins FA & Curley RD (1958) Soil aggregation by microorganisms following soil fumigation. *Agr Hortic*.

McDonald D, Clemente JC, Kuczynski J, *et al.* (2012) The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. *GigaScience* **1**.

McKee LS, Sunner H, Anasontzis GE, Toriz G, Gatenholm P, Bulone V, Vilaplana F & Olsson L (2016) A GH115 α-glucuronidase from *Schizophyllum commune* contributes to the synergistic enzymatic deconstruction of softwood glucuronoarabinoxylan. *Biotechnol Biofuels* **9**.

McMurdie PJ & Holmes S (2013) phyloseq: An r package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**.

Mehta NC, Streuli H, Müller M & Deuel H (1960) Role of polysaccharides in soil aggregation. *J Sci Food Agric* **11**: 40-47.

Mendiburu Fd (2017) Statistical procedures for agricultural research.

Menzel P, Ng KL & Krogh A (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* **7**.

Merchant SS & Helmann JD (2012) Elemental economy. *Advances in Microbial Physiology,* Vol. 60 91-210.

Minh Tran T, MacIntyre A, Khokhani D, Hawes M & Allen C (2016) Extracellular DNases of *Ralstonia solanacearum* modulate biofilms and facilitate bacterial wilt virulence. *Environ Microbiol* **18**: 4103-4117.

Mishra A & Jha B (2013) Microbial Exopolysaccharides. *The Prokaryotes,*(Rosemberg E, ed.) 179-192. Springer-Verlag New York, New York.

Mitchell AL, Scheremetjew M, Denise H, *et al.* (2017) EBI Metagenomics in 2017: enriching the analysis of microbial communities, from sequence reads to assemblies. *Nucleic Acids Res* **46**: D726-D735.

Mohan A, Flora B & Girdhar M (2018) Inulinase: An Important Microbial Enzyme in Food Industry. *Microbial Bioprospecting for Sustainable Development,*237-248.

Moncrieffe MC, Fernandez M-J, Spiteller D, Matsumura H, Gay NJ, Luisi BF & Leadlay PF (2012) Structure of the Glycosyltransferase EryCIII in Complex with its Activating P450 Homologue EryCII. *J Mol Biol* **415**: 92-101.

Moore EK, Villanueva L, Hopmans EC, Rijpstra WIC, Mets A, Dedysh SN, Sinninghe Damsté JS & Schloss PD (2015) Abundant trimethylornithine lipids and specific gene sequences are indicative of *Planctomycete* importance at the oxic/anoxic interface in sphagnum-dominated Northern Wetlands. *Appl Environ Microbiol* **81**: 6333-6344.

More TT, Yadav JSS, Yan S, Tyagi RD & Surampalli RY (2014) Extracellular polymeric substances of bacteria and their potential environmental applications. *J Environ Manage* **144**: 1-25.

Morgenstern A, Paetz C, Behrend A & Spiteller D (2015) Divalent transition-metal-ion stress induces prodigiosin biosynthesis in *Streptomyces coelicolor* m145: formation of coeligiosins. *Chem: Eur J* **21**: 6027-6032.

Morillo Pérez JA, García-Ribera R, Quesada T, Aguilera M, Ramos-Cormenzana A & Monteoliva-Sánchez M (2008) Biosorption of heavy metals by the exopolysaccharide produced by *Paenibacillus jamilae*. *World J Microbiol Biotechnol* **24**: 2699-2704.

Mugnai G, Rossi F, Felde VJMNL, Colesie C, Büdel B, Peth S, Kaplan A & De Philippis R (2017) Development of the polysaccharidic matrix in biocrusts induced by a cyanobacterium inoculated in sand microcosms. *Biol Fertil Soils* **54**: 27-40.

Mulichak AM, Lu W, Losey HC, Walsh CT & Garavito RM (2004) Crystal structure of vancosaminyltransferase GtfD from the vancomycin biosynthetic pathway: interactions with acceptor and nucleotide ligands. *Biochemistry* **43**: 5170-5180.

Mulichak AM, Losey HC, Lu W, Wawrzak Z, Walsh CT & Garavito RM (2003) Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. *Proc Natl Acad Sci U S A* **100**: 9238-9243.

Mummey DL & Stahl PD (2004) Analysis of Soil Whole- and Inner-Microaggregate Bacterial Communities. *Microb Ecol* **48**: 41-50.

Murray LE, Rowley N, Dawes IW, Johnston GC & Singer RA (1998) A yeast glutamine tRNA signals nitrogen status for regulation of dimorphic growth and sporulation. *Proc Natl Acad Sci U S A* **95**: 8619-8624.

N

Nagler M, Podmirseg SM, Griffith GW, Insam H & Ascher-Jenull J (2018) The use of extracellular DNA as a proxy for specific microbial activity. *Appl Microbiol Biotechnol* **102**: 2885-2898.

Nahar S, Mizan MFR, Ha AJ-w & Ha S-D (2018) Advances and Future Prospects of Enzyme-Based Biofilm Prevention Approaches in the Food Industry. *Compr Rev Food Sci Food Saf* **17**: 1484-1502.

Nalin R, Simonet P, Vogel TM & Normand P (1999) *Rhodanobacter lindaniclasticus* gen. nov., sp. nov., a lindane-degrading bacterium. *Int J Syst Bacteriol* **49**: 19-23.

Nampoothiri KM, Singhania RR, Sabarinath C & Pandey A (2003) Fermentative production of gellan using *Sphingomonas paucimobilis*. *Process Biochem* **38**: 1513-1519.

Navarrete AA, Barreto CC, Arnaldo M & Tsai SM (2013) Molecular detection on culture medium of *Acidobacteria* from Amazon soils. *Microbiol Discov* **1**.

Navarrete AA, Kuramae EE, de Hollander M, Pijl AS, van Veen JA & Tsai SM (2013) Acidobacterial community responses to agricultural management of soybean in Amazon forest soils. *FEMS Microbiol Ecol* **83**: 607-621.

NCBI Resource Coordinators (2016) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **44**: D7-D19.

Negrea A, Bjur E, Puiac S, Ygberg SE, Aslund F & Rhen M (2009) Thioredoxin 1 participates in the activity of the *Salmonella enterica* aerovar *Typhimurium* pathogenicity island 2 type III secretion system. *J Bacteriol* **191**: 6918-6927.

Neufeld JD, Vohra J, Dumont MG, Lueders T, Manefield M, Friedrich MW & Murrell JC (2007) DNA stable-isotope probing. *Nat Protoc* **2**: 860-866.

Nguyen NH, Song Z, Bates ST, Branco S, Tedersoo L, Menke J, Schilling JS & Kennedy PG (2016) FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol* **20**: 241-248.

Nichols CM, Bowman JP & Guezennec J (2005) Effects of incubation temperature on growth and production of exopolysaccharides by an antarctic sea ice bacterium grown in batch culture. *Appl Environ Microbiol* **71**: 3519-3523.

Nicolaus B, Moriello VS, Lama L, Poli A & Gambacorta A (2004) Polysaccharides from extremophilic microorganisms. *Orig Life Evol Biosph* **34**: 159-169.

Nicolaus B, Panico A, Manca MC, Lama L, Gambacorta A, Maugeri T, Gugliandolo C & Caccamo D (2000) A thermophilic *Bacillus* isolated from an eolian shallow hydrothermal vent able to produce exopolysaccharides. *Syst Appl Microbiol* **23**: 426-432.

Nidetzky B, Fürlinger M, Gollhofer D, Scopes RK, Haltrich D & Kulbe KD (1997) Improved operational stability of cell-

free glucose-fructose oxidoreductase from *Zymomonas mobilis* for the efficient synthesis of sorbitol and gluconic acid in a continuous ultrafiltration membrane reactor. *Biotechnol Bioeng* **53**: 623-629.

Nielsen PH & Jahn A (1999) Extraction of EPS. *Microbial Extracellular Polymeric Substances,*(Wingender J, Neu TR & Flemming H-C, eds.), 49-72. Springer Berlin Heidelberg, Berlin.

Nikaido H (1996) Multidrug efflux pumps of gram-negative bacteria. *J Bacteriol* **178**: 5853-5859.

O

Oh HN, Park D, Seong HJ, Kim D & Sul WJ (2019) Antarctic tundra soil metagenome as useful natural resources of cold-active lignocelluolytic enzymes. *J Microbiol* **57**: 865-873.

Okada K, Minehira M, Zhu X, Suzuki K, Nakagawa T, Matsuda H & Kawamukai M (1997) The *ispB* gene encoding octaprenyl diphosphate synthase is essential for growth of *Escherichia coli*. *J Bacteriol* **179**: 3058-3060.

Okamura K, Kawai A, Yamada T & Hiraishi A (2011) Acidipila rosea gen. nov., sp. nov., an acidophilic chemoorganotrophic bacterium belonging to the phylum Acidobacteria. *FEMS Microbiol Lett* **317**: 138-142.

Oksanen J, Blanchet FG, Friendly M*, et al.* (2018) vegan: community ecology package. R package version 2.4-6.

Okshevsky M & Meyer RL (2013) The role of extracellular DNA in the establishment, maintenance and perpetuation of bacterial biofilms. *Crit Rev Microbiol* **41**: 341-352.

Olm MR, Brown CT, Brooks B & Banfield JF (2017) dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* **11**: 2864-2868.

Oshkin IY, Kulichevskaya IS, Rijpstra WIC, Sinninghe Damsté JS, Rakitin AL, Ravin NV & Dedysh SN (2019) Granulicella sibirica sp. nov., a psychrotolerant acidobacterium isolated from an organic soil layer in forested tundra, West Siberia. *Int J Syst Evol Microbiol* **69**: 1195-1201.

Oussenko IA, Abe T, Ujiie H, Muto A & Bechhofer DH (2005) Participation of 3'-to-5' exoribonucleases in the turnover of *Bacillus subtilis* mRNA. *J Bacteriol* **187**: 2758-2767.

P

Padilla-Benavides T, Long JE, Raimunda D, Sassetti CM & Argüello JM (2013) A Novel P1B-type Mn2+-transporting ATPase Is required for secreted protein metallation in mycobacteria. *J Biol Chem* **288**: 11334-11347.

Padmanabhan P, Padmanabhan S, DeRito C, Gray A, Gannon D, Snape JR, Tsai CS, Park W, Jeon C & Madsen EL (2003) Respiration of 13C-labeled substrates added to soil in the field and subsequent 16S rRNA gene analysis of 13c-labeled soil DNA. *Appl Environ Microbiol* **69**: 1614-1622.

Pal C, Bengtsson-Palme J, Rensing C, Kristiansson E & Larsson DGJ (2014) BacMet: antibacterial biocide and metal resistance genes database. *Nucleic Acids Res* **42**: D737-D743.

Pan Y, Cassman N, de Hollander M, Mendes LW, Korevaar H, Geerts RHEM, van Veen JA & Kuramae EE (2014) Impact of long-term N, P, K, and NPK fertilization on the composition and potential functions of the bacterial community in grassland soil. *FEMS Microbiol Ecol* **90**: 195-205.

Pan Y, Cassman N, de Hollander M, Mendes LW, Korevaar H, Geerts R, van Veen JA & Kuramae EE (2014) Impact of long-term N, P, K, and NPK fertilization on the composition and potential functions of the bacterial community in grassland soil. *FEMS Microbiol Ecol* **90**: 195-205.

Pankratov TA & Dedysh SN (2010) *Granulicella paludicola* gen. nov., sp nov., *Granulicella pectinivorans* sp nov., *Granulicella aggregans* sp nov and *Granulicella rosea* sp nov., acidophilic, polymer-degrading acidobacteria from Sphagnum peat bogs. *Int J Syst Evol Microbiol* **60**: 2951-2959.

Pankratov TA, Serkebaeva YM, Kulichevskaya IS, Liesack W & Dedysh SN (2008) Substrate-induced growth and isolation of *Acidobacteria* from acidic Sphagnum peat. *ISME J* **2**: 551-560.

Pankratov TA, Kirsanova LA, Kaparullina EN, Kevbrin VV & Dedysh SN (2012) Telmatobacter bradus gen. nov., sp nov., a cellulolytic facultative anaerobe from subdivision 1 of the Acidobacteria, and emended description of Acidobacterium capsulatum Kishimoto et al. 1991. *Int J Syst Evol Microbiol* **62**: 430-437.

Papp-Wallace KM & Maguire ME (2006) Manganese transport and the role of manganese in virulence. *Annu Rev Microbiol* **60**: 187-209.

Parikh A & Madamwar D (2006) Partial characterization of extracellular polysaccharides from cyanobacteria. *Bioresour Technol* **97**: 1822-1827.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P & Tyson GW (2015) CheckM: assessing the quality of microbial

genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043-1055.

Pascual J, Wüst PK, Geppert A, Foesel BU, Huber KJ & Overmann J (2015) Novel isolates double the number of chemotrophic species and allow the first description of higher taxa in Acidobacteria subdivision 4. *Syst Appl Microbiol* **38**: 534-544.

Patel JJ & Gerson T (1974) Formation and utilisation of carbon reserves by *Rhizobium*. *Arch Microbiol* **101**: 211-220.

Patel K & Golemi-Kotra D (2016) Signaling mechanism by the *Staphylococcus aureus* two-component system LytSR: role of acetyl phosphate in bypassing the cell membrane electrical potential sensor LytS. *F1000Res* **4**.

Paterson GK, Cone DB, Northen H, Peters SE & Maskell DJ (2009) Deletion of the gene encoding the glycolytic enzyme triosephosphate isomerase (Tpi) alters morphology of *Salmonella enterica* serovar *typhimurium* and decreases fitness in mice. *FEMS Microbiol Lett* **294**: 45-51.

Paul D, Pandey G, Meier C, van der Meer JR & Jain RK (2006) Bacterial community structure of a pesticide-contaminated site and assessment of changes induced in community structure during bioremediation. *FEMS Microbiol Ecol* **57**: 116-127.

Pavlova K & Grigorova D (1999) Production and properties of exopolysaccharide by *Rhodotorula acheniorum* MC. *Food Res Int* **32**: 473-477.

Pereira de Castro A, Sartori da Silva MRS, Quirino BF, da Cunha Bustamante MM & Krüger RH (2016) Microbial diversity in cerrado biome (neotropical savanna) soils. *PLoS One* **11**.

Perez-Riverol Y, Csordas A, Bai J*, et al.* (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**: D442-D450.

Petry S, Furlan S, Crepeau MJ, Cerning J & Desmazeaud M (2000) Factors affecting exocellular polysaccharide production by *Lactobacillus delbrueckii subsp. bulgaricus* grown in a chemically defined medium. *Appl Environ Microbiol* **66**: 3427-3431.

Pinnell LJ, Dunford E, Ronan P, Hausner M & Neufeld JD (2014) Recovering glycoside hydrolase genes from active tundra cellulolytic bacteria. *Can J Microbiol* **60**: 469-476.

Poli A, Moriello VS, Esposito E, Lama L, Gambacorta A & Nicolaus B (2004) Exopolysaccharide production by a new *Halomonas* strain CRSS isolated from saline lake Cape Russell in Antarctica growing on complex and defined media. *Biotechnol Lett* **26**: 1635-1638.

Porcheron G, Garenaux A, Proulx J, Sabri M & Dozois CM (2013) Iron, copper, zinc, and manganese transport and regulation in pathogenic Enterobacteria: correlations between strains, site of infection and the relative importance of the different metal transport systems for virulence. *Front Cell Infect Microbiol* **3**: 90.

Pruesse E, Peplies J & Glöckner FO (2012) SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**: 1823-1829.

Puri S, Hohle TH & O'Brian MR (2010) Control of bacterial iron homeostasis by manganese. *Proc Natl Acad Sci U S A* **107**: 10691-10695.

Q

Queiroz PS, Ruas FAD, Barboza NR, de Castro Borges W & Guerra-Sá R (2018) Alterations in the proteomic composition of *Serratia marcescens* in response to manganese (II). *BMC Biotechnol* **18**.

Qurashi AW & Sabri AN (2012) Bacterial exopolysaccharide and biofilm formation stimulate chickpea growth and soil aggregation under salt stress. *Braz J Microbiol* **43**: 1183-1191.

R

R Core Team (2015) *R: A Language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Rashid MI, Mujawar LH, Shahzad T, Almeelbi T, Ismail IMI & Oves M (2016) Bacteria and fungi can contribute to nutrients bioavailability and aggregate formation in degraded soils. *Microbiol Res* **183**: 26-41.

Rawat SR, Männistö MK, Starovoytov V, Goodwin L, Nolan M, Hauser L, Land M, Davenport KW, Woyke T & Häggblom MM (2013) Complete genome sequence of *Granulicella tundricola* type strain MP5ACTX9T, an *Acidobacteria* from tundra soil. *Stand Genomic Sci* **9**: 449-461.

Rawat SR, Männistö MK, Starovoytov V, Goodwin L, Nolan M, Hauser LJ, Land M, Davenport KW, Woyke T & Häggblom MM (2013) Complete genome sequence of *Granulicella mallensis* type strain MP5ACTX8T, an acidobacterium

from tundra soil. *Stand Genomic Sci* **9**: 71-82.

Rehm BHA (2010) Bacterial polymers: biosynthesis, modifications and applications. *Nat Rev Microbiol* **8**: 578-592.

Remminghorst U & Rehm BHA (2006) Bacterial alginates: from biosynthesis to applications. *Biotechnol Lett* **28**: 1701-1712.

Reuber TL & Walker GC (1993) Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. *Cell* **74**: 269-280.

Richardson AR, Somerville GA & Sonenshein AL (2015) Regulating the Intersection of Metabolism and Pathogenesis in Gram-positive Bacteria. *Microbiol Spectr* **3**.

Roberson EB & Firestone MK (1992) Relationship between desiccation and exopolysaccharide production in a soil *Pseudomonas sp*. *Appl Environ Microbiol* **58**: 1284-1291.

Roberson EB, Chenu C & Firestone MK (1993) Microstructural changes in bacterial exopolysaccharides during desiccation. *Soil Biol Biochem* **25**: 1299-1301.

Roberson EB, Shennan C, Firestone MK & Sarig S (1995) Nutritional Management of Microbial Polysaccharide Production and Aggregation in an Agricultural Soil. *Soil Sci Soc Am J* **59**.

Robinson MD, McCarthy DJ & Smyth GK (2009) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139-140.

Roca C, Alves VD, Freitas F & Reis MAM (2015) Exopolysaccharides enriched in rare sugars: bacterial sources, production, and applications. *Front Microbiol* **6**.

Rogers SL & Burns RG (1994) Changes in aggregate stability, nutrient status, indigenous microbial populations, and seedling emergence, following inoculation of soil withNostoc muscorum. *Biol Fertil Soils* **18**: 209-215.

Rognes T, Mahé F & xflouris (2015) vsearch: VSEARCH version 1.0.16.

Romaní AM, Fischer H, Mille-Lindblom C & Tranvik LJ (2006) Interactions of bacteria and fungi on decomposing litter: differential extracellular enzyme activities. *Ecology* **87**: 2559-2569.

Rosenzweig R, Shavit U & Furman A (2012) Water Retention Curves of Biofilm-Affected Soils using Xanthan as an Analogue. *Soil Sci Soc Am J* **76**.

Rosner JL & Martin RG (2009) An Excretory function for the *Escherichia coli* outer membrane pore TolC: upregulation of marA and soxS transcription and rob activity due to metabolites accumulated in tolC mutants. *J Bacteriol* **191**: 5283-5292.

Rossi F & De Philippis R (2015) Role of Cyanobacterial Exopolysaccharides in Phototrophic Biofilms and in Complex Microbial Mats. *Life* **5**: 1218-1238.

Rossi F, Mugnai G & De Philippis R (2017) Complex role of the polymeric matrix in biological soil crusts. *Plant Soil*.

Ruhal R, Kataria R & Choudhury B (2013) Trends in bacterial trehalose metabolism and significant nodes of metabolic pathway in the direction of trehalose accumulation. *Microb Biotechnol* **6**: 493-502.

Ryder C, Byrd M & Wozniak DJ (2007) Role of polysaccharides in *Pseudomonas aeruginosa* biofilm development. *Curr Opin Microbiol* **10**: 644-648.

S

Sá-Correia I, Fialho AM, Videira P, Moreira LM, Marques AR & Albano H (2002) Gellan gum biosynthesis in *Sphingomonas paucimobilis* ATCC 31461: genes, enzymes and exopolysaccharide production engineering. *J Ind Microbiol Biotechnol* **29**: 170-176.

Sacheti P, Patil R, Dube A, *et al.* (2014) Proteomics of arsenic stress in the gram-positive organism *Exiguobacterium* sp. PS NCIM 5463. *Appl Microbiol Biotechnol* **98**: 6761-6773.

Sadeghi SH, Kheirfam H, Homaee M, Darki BZ & Vafakhah M (2017) Improving runoff behavior resulting from direct inoculation of soil micro-organisms. *Soil Tillage Res* **171**: 35-41.

Saeed AI, Sharov V, White J, *et al.* (2003) TM4: A free,open-source system for microarray data management and analysis. *Biotechniques* **34**: 374-378.

Sakurai K, Arai H, Ishii M & Igarashi Y (2010) Transcriptome response to different carbon sources in *Acetobacter aceti*. *Microbiology* **157**: 899-910.

Sanchez-Peinado MD, Gonzalez-Lopez J, Martinez-Toledo MV, Pozo C & Rodelas B (2010) Influence of linear

alkylbenzene sulfonate (LAS) on the structure of Alphaproteobacteria, Actinobacteria, and Acidobacteria communities in a soil microcosm. *Environ Sci Pollut Res* **17**: 779-790.

Schabereiter-Gurtner C, Saiz-Jimenez C, Pinar G, Lubitz W & Rolleke S (2002) Altamira cave Paleolithic paintings harbor partly unknown bacterial communities. *FEMS Microbiol Lett*.

Schellenberger S, Kolb S & Drake HL (2009) Metabolic responses of novel cellulolytic and saccharolytic agricultural soil Bacteria to oxygen. *Environ Microbiol* **12**: 845-861.

Schlemper TR, van Veen JA & Kuramae EE (2017) Co-variation of bacterial and fungal communities in different sorghum cultivars and growth stages is soil dependent. *Microb Ecol* **76**: 205-214.

Schlesner H (1994) The development of media suitable for the microorganisms morphologically resembling *Planctomyces* spp., *Pirellula* spp., and other *Planctomycetales* from various aquatic habitats using dilute media. *Syst Appl Microbiol* **17**: 135-145.

Schmid J, Sieber V & Rehm B (2015) Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. *Front Microbiol* **6**.

Schmid J, Heider D, Wendel NJ, Sperl N & Sieber V (2016) Bacterial Glycosyltransferases: Challenges and Opportunities of a Highly Diverse Enzyme Class Toward Tailoring Natural Products. *Front Microbiol* **7**.

Schneider T, Schmid E, de Castro JV, Cardinale M, Eberl L, Grube M, Berg G & Riedel K (2011) Structure and function of the symbiosis partners of the lung lichen (*Lobaria pulmonaria* L. Hoffm.) analyzed by metaproteomics. *Proteomics* **11**: 2752-2756.

Schneider T, Keiblinger KM, Schmid E, Sterflinger-Gleixner K, Ellersdorfer G, Roschitzki B, Richter A, Eberl L, Zechmeister-Boltenstern S & Riedel K (2012) Who is who in litter decomposition? Metaproteomics reveals major microbial players and their biogeochemical functions. *ISME J* **6**: 1749-1762.

Schön A (1999) Ribonuclease P: the diversity of a ubiquitous RNA processing enzyme. *FEMS Microbiol Rev* **23**: 391-406.

Schwab C, Walter J, Tannock GW, Vogel RF & Gänzle MG (2007) Sucrose utilization and impact of sucrose on glycosyltransferase expression in *Lactobacillus reuteri*. *Syst Appl Microbiol* **30**: 433-443.

Scieglinska D & Krawczyk Z (2014) Expression, function, and regulation of the testis-enriched heat shock HSPA2 gene in rodents and humans. *Cell Stress Chaperon* **20**: 221-235.

Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**: 2068-2069.

Seemann T (2018) BAsic Rapid Ribosomal RNA Predictor 0.9.

Sessitsch A, Weilharter A, Gerzabek MH, Kirchmann H & Kandeler E (2001) Microbial Population Structures in Soil Particle Size Fractions of a Long-Term Fertilizer Field Experiment. *Appl Environ Microbiol* **67**: 4215-4224.

Shannon P (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**: 2498-2504.

Sheng GP, Yu HQ & Yue Z (2006) Factors influencing the production of extracellular polymeric substances by *Rhodopseudomonas acidophila*. *Int Biodeter Biodegr* **58**: 89-93.

Shih I-L, Yu J-Y, Hsieh C & Wu J-Y (2009) Production and characterization of curdlan by *Agrobacterium* sp. *Biochem Eng J* **43**: 33-40.

Shu C-H & Lung M-Y (2004) Effect of pH on the production and molecular weight distribution of exopolysaccharide by *Antrodia camphorata* in batch cultures. *Process Biochem* **39**: 931-937.

Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG & Banfield JF (2018) Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol* **3**: 836-843.

Singh V, Maniar K, Bhattacharyya R & Banerjee D (2018) A novel insight in favor of structure–function relationship for 16S rRNA. *Mol Biol Rep* **45**: 1569-1573.

Singleton DR, Powell SN, Sangaiah R, Gold A, Ball LM & Aitken MD (2005) Stable-isotope probing of bacteria capable of degrading salicylate, naphthalene, or phenanthrene in a bioreactor treating contaminated soil. *Appl Environ Microbiol* **71**: 1202-1209.

Sokolov M, Lu L, Tucker W, Gao F, Gegenheimer PA & Richter ML (1999) The 20 C-terminal amino acid residues of the chloroplast ATP synthase γ subunit are not essential for activity. *J Biol Chem* **274**: 13824-13829.

Song M, Jiang L, Zhang D, Luo C, Wang Y, Yu Z, Yin H & Zhang G (2016) Bacteria capable of degrading anthracene, phenanthrene, and fluoranthene as revealed by DNA based stable-isotope probing in a forest soil. *J Hazard Mater*

308: 50-57.

Souli M & Giamarellou H (1998) Effects of slime produced by clinical isolates of coagulase-negative staphylococci on activities of various antimicrobial agents. *Antimicrob Agents Chemother* **42**: 939-941.

Sparling GP & Cheshire MV (1985) Effect of periodate oxidation on the polysaccharide content and microaggregate stability of rhizosphere and non-rhizosphere soils. *Plant Soil* **88**: 113-122.

Spring S, Bunk B, Spröer C, Schumann P, Rohde M, Tindall BJ & Klenk H-P (2016) Characterization of the first cultured representative of *Verrucomicrobia* subdivision 5 indicates the proposal of a novel phylum. *ISME J* **10**: 2801-2816.

Srikanth R, Reddy CHSSS, Siddartha G, Ramaiah MJ & Uppuluri KB (2015) Review on production, characterization and applications of microbial levan. *Carbohydr Polym* **120**: 102-114.

Staley JT & Konopka A (1985) Measurement of in Situ Activities of Nonphotosynthetic Microorganisms in Aquatic and Terrestrial Habitats. *Annu Rev Microbiol* **39**: 321-346.

Stasinopoulos SJ, Fisher PR, Stone BA & Stanisich VA (1999) Detection of two loci involved in (1 -> 3)-beta-glucan (curdlan) biosynthesis by *Agrobacterium* sp. ATCC31749, and comparative sequence analysis of the putative curdlan synthase gene. *Glycobiology* **9**: 31-41.

Staudt AK, Wolfe LG & Shrout JD (2011) Variations in exopolysaccharide production by *Rhizobium tropici*. *Arch Microbiol* **194**: 197-206.

Stevenson BS, Eichorst SA, Wertz JT, Schmidt TM & Breznak JA (2004) New strategies for cultivation and detection of previously uncultured microbes. *Appl Environ Microbiol* **70**: 4748-4755.

Stopnisek N, Zühlke D, Carlier A, Barberán A, Fierer N, Becher D, Riedel K, Eberl L & Weisskopf L (2016) Molecular mechanisms underlying the close association between soil *Burkholderia* and fungi. *ISME J* **10**: 253-264.

Stothard P & Wishart DS (2004) Circular genome visualization and exploration using CGView. *Bioinformatics* **21**: 537-539.

Šturšová M, Žifčáková L, Leigh MB, Burgess R & Baldrian P (2012) Cellulose utilization in forest litter and soil: identification of bacterial and fungal decomposers. *FEMS Microbiol Ecol* **80**: 735-746.

Sutherland IW (2001) Microbial polysaccharides from Gram-negative bacteria. *Int Dairy J* **11**: 663-674.

Sutherland IW (2004) Microbial Exopolysaccharides. *Polysaccharides: Structural Diversity and Functional Versatility,*(Dumitriu S, ed.) CRC Press, Florida.

Swaby RJ (1949) The relationship between micro-organisms and soil aggregation. *J Gen Microbiol* **3**: 236-254.

T

Tachdjian S & Kelly RM (2006) Dynamic metabolic adjustments and genome plasticity are implicated in the heat shock response of the extremely thermoacidophilic archaeon *Sulfolobus solfataricus*. *J Bacteriol* **188**: 4553-4559.

Tamames J & Puente-Sánchez F (2019) SqueezeMeta, A Highly Portable, Fully Automatic Metagenomic Analysis Pipeline. *Front Microbiol* **9**.

Tamura K & Nei M (1993) Estimation of the Number of Nucleotide Substitutions in the Control Region of Mitochondrial-DNA in Humans and Chimpanzees. *Mol Biol Evol* **10**: 512-526.

Tang J, Mo Y, Zhang J & Zhang R (2011) Influence of biological aggregating agents associated with microbial population on soil aggregate stability. *Appl Soil Ecol* **47**: 153-159.

Tank M & Bryant DA (2015) Nutrient requirements and growth physiology of the photoheterotrophic Acidobacterium, Chloracidobacterium thermophilum. *Front Microbiol* **06**.

Tebar AR, Ballesteros A & Soria J (1977) Mn2+ electron spin resonance studies on ATP phosphoribosyltransferase from *E. coli*. *Experientia* **33**: 1292-1294.

Thoma R, Hennig M, Sterner R & Kirschner K (2000) Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima*. *Structure* **8**: 265-276.

Thrash JC & Coates Jd (2014) Phylum XVII. Acidobacteria phyl. nov. *Bergey's Manual of Systematic Bacteriology,*(Krieg NR, Ludwig W, Whitman WB, Hedlund BP, Paster BJ, Staley JT, Ward N & Brown D, eds.).

Tieking M, Ehrmann MA, Vogel RF & Gänzle MG (2004) Molecular and functional characterization of a levansucrase from the sourdough isolate *Lactobacillus sanfranciscensis* TMW 1.392. *Appl Microbiol Biotechnol* **66**: 655-663.

Tisdall JM & Oades JM (1982) Organic matter and water-stable aggregates in soils. *J Soil Sci* **33**: 141-163.

Torrent M, Chalancon G, de Groot NS, Wuster A & Madan Babu M (2018) Cells alter their tRNA abundance to selectively regulate protein synthesis during stress conditions. *Sci Signal* **11**.

Torres CAV, Antunes S, Ricardo AR, Grandfils C, Alves VD, Freitas F & Reis MAM (2012) Study of the interactive effect of temperature and pH on exopolysaccharide production by *Enterobacter* A47 using multivariate statistical analysis. *Bioresour Technol* **119**: 148-156.

Trchounian K, Poladyan A & Trchounian A (2016) Optimizing strategy for *Escherichia coli* growth and hydrogen production during glycerol fermentation in batch culture: effects of some heavy metal ions and their mixtures. *Appl Energy* **177**: 335-340.

Tsuneda S, Aikawa H, Hayashi H, Yuasa A & Hirata A (2003) Extracellular polymeric substances responsible for bacterial adhesion onto solid surface. *FEMS Microbiol Lett* **223**: 287-292.

U

Uhlik O, Wald J, Strejcek M, Musilova L, Ridl J, Hroudova M, Vlcek C, Cardenas E, Mackova M & Macek T (2012) Identification of bacteria utilizing biphenyl, benzoate, and naphthalene in long-term contaminated soil. *PLoS One* **7**: e40653.

Umer MI & Rajab SM (2012) Correlation between aggregate stability and microbiological activity in two Russian soil types. *Eurasian J Soil Sci* **1**: 5.

Upadhyay A, Kochar M, Rajam MV & Srivastava S (2017) Players over the Surface: Unraveling the Role of Exopolysaccharides in Zinc Biosorption by Fluorescent Pseudomonas Strain Psd. *Front Microbiol* **8**.

Upadhyay SK, Singh JS & Singh DP (2011) Exopolysaccharide-producing plant growth-promoting rhizobacteria under salinity condition. *Pedosphere* **21**: 214-222.

Urbanová M, Šnajdr J & Baldrian P (2015) Composition of fungal and bacterial communities in forest litter and soil is largely determined by dominant trees. *Soil Biol Biochem* **84**: 53-64.

V

Valášková V, de Boer W, Klein Gunnewiek PJA, Pospíšek M & Baldrian P (2009) Phylogenetic composition and properties of bacteria coexisting with the fungus *Hypholoma* fasciculare in decaying wood. *ISME J* **3**: 1218-1221.

Valdivia-Anistro JA, Eguiarte-Fruns LE, Delgado-Sapién G, Márquez-Zacarías P, Gasca-Pineda J, Learned J, Elser JJ, Olmedo-Alvarez G & Souza V (2016) Variability of rRNA Operon Copy Number and Growth Rate Dynamics of Bacillus Isolated from an Extremely Oligotrophic Aquatic Ecosystem. *Front Microbiol* **6**.

Valepyn E, Berezina N & Paquot M (2012) Optimization of production and preliminary characterization of new exopolysaccharides from *Gluconacetobacter hansenii* LMG1524. *Adv Microbiol* **02**: 488-496.

van Bueren AL, Saraf A, Martens EC, Dijkhuizen L & Drake HL (2015) Differential metabolism of exopolysaccharides from probiotic *Lactobacilli* by the human gut symbiont *Bacteroides thetaiotaomicron*. *Appl Environ Microbiol* **81**: 3973-3983.

Vardharajula S & Ali SZ (2014) Exopolysaccharide production by drought tolerant *Bacillus* spp. and effect on soil aggregation under drought stress. *J Microbiol Biotechnol Food Sci* **4**: 51-57.

Vardharajula S & Ali SZ (2015) The production of exopolysaccharide by *Pseudomonas putida* GAP-P45 under various abiotic stress conditions and its role in soil aggregation. *Microbiology* **84**: 512-519.

Vardharajula S, Sk. Z A, Grover M, Reddy G & Venkateswarlu B (2009) Alleviation of drought stress effects in sunflower seedlings by the exopolysaccharides producing Pseudomonas putida strain GAP-P45. *Biol Fertil Soils* **46**: 17-26.

Velasco Ayuso S, Giraldo Silva A, Nelson C, Barger NN, Garcia-Pichel F & Stams AJ (2017) Microbial Nursery Production of High-Quality Biological Soil Crust Biomass for Restoration of Degraded Dryland Soils. *Appl Environ Microbiol* **83**.

Verastegui Y, Cheng J, Engel K, *et al.* (2014) Multisubstrate isotope labeling and metagenomic analysis of active soil bacterial communities. *mBio* **5**: e01157-01114.

Vester JK, Glaring MA & Stougaard P (2015) Improved cultivation and metagenomics as new tools for bioprospecting in cold environments. *Extremophiles* **19**: 17-29.

Vicente-García V, Ríos-Leal E, Calderón-Domínguez G, Cañizares-Villanueva RO & Olvera-Ramírez R (2004) Detection, isolation, and characterization of exopolysaccharide produced by a strain of *Phormidium* 94a isolated from an arid zone of Mexico. *Biotechnol Bioeng* **85**: 306-310.

Vieira S, Luckner M, Wanner G & Overmann J (2017) Luteitalea pratensis gen. nov., sp. nov. a new member of subdivision 6 Acidobacteria isolated from temperate grassland soil. *Int J Syst Evol Microbiol* **67**: 1408-1414.

Vilain S, Pretorius JM, Theron J & Brozel VS (2009) DNA as an adhesin: *Bacillus cereus* requires extracellular DNA to form biofilms. *Appl Environ Microbiol* **75**: 2861-2868.

Vivanco L, Rascovan N & Austin AT (2018) Plant, fungal, bacterial, and nitrogen interactions in the litter layer of a native Patagonian forest. *PeerJ* **6**.

Vorhölter F-J, Schneiker S, Goesmann A, *et al.* (2008) The genome of *Xanthomonas campestris* pv. *campestris* B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. *J Biotechnol* **134**: 33-45.

W

Walker BJ, Abeel T, Shea T, *et al.* (2014) Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9**.

Wan Z, Brown PJB, Elliott EN & Brun YV (2013) The adhesive and cohesive properties of a bacterial polysaccharide adhesin are modulated by a deacetylase. *Mol Microbiol* **88**: 486-500.

Wang J, Jenkins C, Webb RI & Fuerst JA (2002) Isolation of *Gemmata*-like and *Isosphaera*-like *Planctomycete* bacteria from soil and freshwater. *Appl Environ Microbiol* **68**: 417-422.

Wang N, Zhang S & He M (2016) Bacterial community profile of contaminated soils in a typical antimony mining site. *Environ Sci Pollut Res* **25**: 141-152.

Wang Q, Garrity GM, Tiedje JM & Cole JR (2007) Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol* **73**: 5261-5267.

Wang X, Sharp CE, Jones GM, Grasby SE, Brady AL, Dunfield PF & Nojiri H (2015) Stable-isotope probing identifies uncultured *Planctomycetes* as primary degraders of a complex heteropolysaccharide in soil. *Appl Environ Microbiol* **81**: 4607-4615.

Wang X, Xu P, Yuan Y, Liu C, Zhang D, Yang Z, Yang C & Ma C (2006) Modeling for gellan gum production by *Sphingomonas paucimobilis* ATCC 31461 in a simplified medium. *Appl Environ Microbiol* **72**: 3367-3374.

Ward NL, Challacombe JF, Janssen PH, *et al.* (2009) Three genomes from the phylum *Acidobacteria* provide insight into the lifestyles of these microorganisms in soils. *Appl Environ Microbiol* **75**: 2046-2056.

Warnecke F, Luginbühl P, Ivanova N, *et al.* (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450**: 560-565.

Warner JB & Lolkema JS (2003) CcpA-Dependent carbon catabolite repression in bacteria. *Microbiol Mol Biol Rev* **67**: 475-490.

Wegner C-E & Liesack W (2017) Unexpected Dominance of Elusive Acidobacteria in Early Industrial Soft Coal Slags. *Front Microbiol* **8**.

Whang KS, Lee JC, Lee HR, Han SI & Chung SH (2014) *Terriglobus tenax* sp. nov., an exopolysaccharide-producing acidobacterium isolated from rhizosphere soil of a medicinal plant. *Int J Syst Evol Microbiol* **64**: 431-437.

Whitchurch CB, Tolker-Nielsen T, Ragas PC & Mattick JS (2002) Extracellular DNA required for bacterial biofilm formation. *Science* **295**: 1487-1487.

White, Bruns T, Lee S & Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics. *PCR - Protocols and Applications - A Laboratory Manual,*315-322. Academic Press.

White III RA, Brown J, Colby S, Overall CC, Lee J-Y, Zucker J, Glaesemann KR, Jansson C & Jansson JK (2017) ATLAS (Automatic Tool for Local Assembly Structures) - a comprehensive infrastructure for assembly, annotation, and genomic binning of metagenomic and metatranscriptomic data. *PeerJ* **1**.

Whitney JC & Howell PL (2013) Synthase-dependent exopolysaccharide secretion in gram-negative Bacteria. *Trends Microbio* **21**: 63-72.

Wiater A, Szczodrak J, Pleszczynska M & Próchniak K (2005) Production and use of mutanase from Trichoderma harzianum for effective degradation of streptococcal mutans. *Braz J Microbiol* **36**.

Wickham H (2016) ggplot2. *Use R!*

Wiegand S, Jogler M & Jogler C (2018) On the maverick *Planctomycetes*. *FEMS Microbiol Rev* 739–760.

Williams A, Wilkinson A, Krehenbrink M, Russo DM, Zorreguieta A & Downie JA (2008) Glucomannan-mediated attachment of *Rhizobium leguminosarum* to pea root hairs is required for competitive nodule infection. *J Bacteriol* **190**: 4706-4715.

Wingender J, Neu TR & Flemming H-C (1999) What arebacterial extracellular polymeric substances? *Microbial Extracellular Polymeric Substances: Characterization, Structure and Function,*(Wingender J, Neu TR & Flemming H-C, eds.), 1-19. Springer Berlin Heidelberg, Berlin.

Wingender J, Jaeger K-E & Flemming H-C (1999) Interaction Between Extracellular Polysaccharides and Enzymes. *Microbial Extracellular Polymeric Substances* 231-251. Springer Berlin Heidelberg, Berlin.

Wingender J, Jaeger K-E & Flemming H-C (1999) Interaction between extracellular polysaccharides and enzymes. *Microbial extracellular polymeric substances,*231-251. Springer.

Wintsche B, Glaser K, Sträuber H, Centler F, Liebetrau J, Harms H & Kleinsteuber S (2016) Trace elements induce predominance among methanogenic activity in anaerobic digestion. *Front Microbiol* **7**.

Wolfaardt GM, Lawrence JR & Korber DR (1999) Function of EPS. *Microbial Extracellular Polymeric Substances* 171-200. Springer Berlin Heidelberg, Berlin.

Wu N, Pan H-X, Qiu D & Zhang Y-M (2014) Feasibility of EPS-producing bacterial inoculation to speed up the sand aggregation in the Gurbantünggüt Desert, Northwestern China. *J Basic Microbiol* **54**: 1378-1386.

Wu Y-W, Tang Y-H, Tringe SG, Simmons BA & Singer SW (2014) MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome* **2**.

Wüst PK, Foesel BU, Geppert A, Huber KJ, Luckner M, Wanner G & Overmann J (2016) Brevitalea aridisoli, B. deliciosa and Arenimicrobium luteum, three novel species of Acidobacteria subdivision 4 (class Blastocatellia) isolated from savanna soil and description of the novel family Pyrinomonadaceae. *Int J Syst Evol Microbiol* **66**: 3355-3366.

Y

Yamada K, Okuno Y, Meng X-Y, Tamaki H, Kamagata Y & Hanada S (2014) Granulicella cerasi sp. nov., an acidophilic bacterium isolated from cherry bark. *Int J Syst Evol Microbiol* **64**: 2781-2785.

Yin Y, Mao X, Yang J, Chen X, Mao F & Xu Y (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* **40**: W445-W451.

Yoon S-H, Ha S-m, Lim J, Kwon S & Chun J (2017) A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* **110**: 1281-1286.

Yu H, Rao X & Zhang K (2017) Nucleoside diphosphate kinase (Ndk): A pleiotropic effector manipulating bacterial virulence and adaptive responses. *Microbiol Res* **205**: 125-134.

Yu X, Zhang C, Yang L, Zhao L, Lin C, Liu Z & Mao Z (2015) CrdR function in a curdlan-producing *Agrobacterium* sp. ATCC31749 strain. *BMC Microbiol* **15**: 25.

Yuan S, Yin S, Liu M & Kong J-Q (2018) Isolation and characterization of a multifunctional flavonoid glycosyltransferase from Ornithogalum caudatum with glycosidase activity. *Sci Rep* **8**.

Yuksekdag ZN & Aslim B (2008) Influence of different carbon sources on exopolysaccharide production by *Lactobacillus delbrueckii* subsp. *bulgaricus* (B3, G12) and *Streptococcus thermophilus* (W22). *Braz Arch Biol Technol* **51**: 581-585.

Z

Zgurskaya HI, Krishnamoorthy G, Ntreh A & Lu S (2011) Mechanism and Function of the Outer Membrane Channel TolC in Multidrug Resistance and Physiology of Enterobacteria. *Front Microbiol* **2**.

Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y & Yin Y (2018) dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* **46**: W95-W101.

Zhang X, Xin X, Zhu A, Yang W, Zhang J, Ding S, Mu L & Shao L (2018) Linking macroaggregation to soil microbial community and organic carbon accumulation under different tillage and residue managements. *Soil Tillage Res* **178**: 99-107.

Zhang Y, Rodionov DA, Gelfand MS & Gladyshev VN (2009) Comparative genomic analyses of nickel, cobalt and vitamin B12 utilization. *BMC genomics* **10**: 78.

Zhang Y, Wen Z, Washburn MP & Florens L (2010) Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal Chem* **82**: 2272-2281.

Zhang Y, Deng W, Xie X & Jiao N (2016) Differential incorporation of carbon substrates among microbial populations identified by field-based, DNA stable-isotope probing in South China sea. *PLoS One* **11**: e0157178.

Zimmermann J, Gonzalez JM, Saiz-Jimenez C & Ludwig W (2005) Detection and phylogenetic relationships of highly diverse uncultured acidobacterial communities in altamira cave using 23S rRNA sequence analyses. *Geomicrobiol J* **22**: 379-388.

Zweers JC, Nicolas P, Wiegert T, van Dijl JM & Denham EL (2012) Definition of the σW Regulon of *Bacillus subtilis* in the Absence of Stress. *PLoS One* **7**.s

Summary

Acidobacteria is a widely distributed phylum but their functional roles in ecosystem processes are still largely elusive. The high abundance and ubiquity of *Acidobacteria* in different environments, especially soils, raises intriguing questions about the physiological traits underlying their marked abundance. This lack of fundamental knowledge is mostly due to difficulties to isolate *Acidobacteria* and their slow growth in vitro. Increased knowledge of the traits of different *Acidobacteria* subdivisions is critical for understanding their persistence in soil as well as their interactions with other soil microorganisms.

The *Granulicella* genus belongs to the class *Acidobacteriia* and is known to produce copious amount of Extracellular Polymeric Substances (EPS) in culture medium. EPS are fundamental for microbial life and provide an ideal substrate for chemical reactions, nutrient entrapment and protection against environmental stresses such as salinity and drought. Microbial EPS can enhance the aggregation of soil particles and benefit plants by maintaining the moisture of the environment and trapping nutrients. The EPS of the two strains *Granulicella* sp. WH15 and 5B5 were chemically and physically characterized previously, showing interesting emulsification properties. Therefore, the major goal of my thesis was to investigate and understand the metabolism of these two *Granulicella* sp. strains, and the functions and environmental fate of EPS of *Granulicella* sp. WH15 (WH15EPS). We optimized growth conditions of *Granulicella* sp. WH15 and 5B5 in the laboratory to obtain higher biomass and enhance EPS production while using multi-omics approaches to investigate their metabolic adaptations. Furthermore, we used high-throughput sequencing to detected other soil microorganisms and their functions associated with the degradation of WH15EPS.

In **chapter 3**, we showed that the transcriptional and proteomic responses of *Granulicella* sp. WH15 grown at different concentrations of cellobiose resulted in higher expression of genes encoding excretory functions and reallocation of resources to maintain basic cell metabolism instead of generating new cell biomass. In **chapter 4**, we observed that the addition of a mix of trace elements to culture medium improved significantly the growth yield of *Granulicella* sp. WH15 and 5B5. After evaluating the effects of each of the trace elements separately, our results showed that manganese (Mn) had a significant positive effect on the growth of both strains. Further proteomic and genomic analyses showed that the strains had different proteomic profiles and several uncharacterized metal ion transporters that could be involved in metal ion homeostasis and could contribute to survival under high manganese concentrations.

Optimization of carbon concentration and manganese in culture medium allowed our strains, especially WH15, to grow faster in laboratory conditions, producing extractable amounts of EPS.

In **chapter 5**, we labeled WH15EPS with ^{13}C and investigated its effect on the assembly and co-occurrence of the active bacterial and fungal communities in topsoil by stable isotope probing (SIP). Our results demonstrated that WH15EPS was primarily assimilated by *Planctomycetes*, *Verrucomicrobia*, *Ascomycota* and *Basidiomycota* and co-inertia analysis

suggested overall relationships between these bacterial and fungal kingdoms. We observed the incorporation of WH15EPS by *Singulisphaera* and its connections to other *Planctomycetes* and *Acidobacteria*, which were not reported before.

In **chapter 6**, we applied WH15EPS as an enrichment factor to target microorganisms and functions associated with EPS degradation through culture-independent and culture-dependent techniques. Our results showed a large diversity of glycoside hydrolase families with biotechnological potential and a high number of unclassified microorganisms that could be targeted for future studies.

In **chapter 7**, I integrated the overall findings of my thesis and discussed the most important observations concerning the impact of carbon sources and trace elements on the physiology of *Granulicella* and, more general, the ecological functions and environmental fate of EPS of *Acidobacteria*. The data presented in this thesis establish a solid fundamental basis for more mechanistic studies of *Acidobacteria* and other uncultivated microbes.

Samenvatting

Acidobacteriën zijn wijdverbreid, maar over hun functionele rol in het ecosysteem is nog veel onbekend. De alomtegenwoordigheid van *Acidobacteria* in verschillende ecosystemen, en met name de bodem, werpt prangende vragen op over de fysiologische eigenschappen die hieraan ten grondslag liggen. Het gebrek aan deze fundamentele kennis is voornamelijk te wijten aan het feit dat *Acidobacteria* moeilijk te isoleren zijn en onder laboratoriumomstandigheden een lage groeisnelheid kennen. Meer inzicht in de fysiologische eigenschappen van verschillende subdivisies van de *Acidobacteria* is van belang om hun voorkomen in bodemecosystemen en interacties met andere bodemorganismen beter te kunnen begrijpen.

Het genus *Granulicella* behoort tot de klasse *Acidobacteriia*. Van dit genus is bekend dat het in vloeibare kweken overvloedig extracellulaire polysachariden (EPS) kan produceren. EPS zijn voor bacteriën van fundamenteel belang; ze vormen een ideaal substraat voor chemische reacties, bevorderen acquisitie van nutriënten en beschermen tegen zout- en droogtestress. EPS geproduceerd door microörganismen kunnen op hun beurt planten een voordeel bieden door de aggregatie van bodemdeeltjes te verbeteren en daarmee het vocht- en nutriëntenvasthoudend vermogen van de bodem te vergroten.

Eerder onderzoek naar de chemische en fysische eigenschappen van EPS van de twee *Granulicella* sp.-stammen WH15 en 5B5, toonde aan dat deze sachariden een emulsificerende werking kunnen hebben. De belangrijkste doelstellingen van mijn proefschrift waren daarom i) het in kaart brengen en doorgronden van het metabolisme van de bovengenoemde *Granulicella* stammen, ii) het ontrafelen van de functies van het door stam WH15 geproduceerde EPS (WH15EPS) en iii) het gedrag van deze EPS in bodemecosystemen.

Voor ons onderzoek optimaliseerden we in het laboratorium allereerst de groeicondities van de *Granulicella* stammen WH15 en 5B5. Zo verkregen we hoge biomassa en stimuleerden EPS-productie. Aanpassingen in het metabolisme van WH15 en 5B5 werden vervolgens in kaart gebracht door gebruik te maken van een zogenaamde 'multi-omics approach'. Next-generation sequencing, ook wel high-throughput sequencing genoemd, werd ingezet om bodemgerelateerde micro-organismen te indentificeren die geassocieerd zijn met de afbraak van het WH15EPS.

In **hoofdstuk 3** toonden we, middels transcriptionele en proteome analysesanscriptionele aan dat wanneer *Granulicella* sp. WH15 gekweekt wordt op toenemende concentraties cellobiose, er een verhoogde expressie is van genen die excretie reguleren en genen die een rol spelen in de basale celstofwisseling. Deze verhoogde expressie ging ten koste van de aanmaak van nieuwe biomassa. In **hoofdstuk 4** observeerden we dat de groeiopbrengst van *Granulicella* sp. WH15 en 5B5 significant hoger was wanneer een mix van sporen-elementen aan het groeimedium toegevoegd werd. Mangaan (Mn) bleek, na evaluatie van de effecten van elk van de individuele sporenelementen, een significant positief effect te hebben op de groei van beide stammen. Proteoom- en genoomanalyses brachten aan het licht dat de stammen beschikken over diverse, nog niet-gekarakteriseerde metaalion-transporters die mogelijk betrokken zijn in metaalion-homeostase en waarschijnlijk ook bijdragen aan een

verhoogde overlevingskans bij hoge mangaanconcentraties. Optimalisatie van koolstof- en mangaanconcentraties in het groeimedium liet de geselecteerde stammen, en vooral WH15, onder laboratoriumomstandigheden sneller groeien en meer extraheerbare hoeveelheden EPS produceren.

In **hoofdstuk 5** hebben we WH15EPS gelabeld met ^{13}C en het effect hiervan bestudeerd op de assemblage van de actieve bacteriële en schimmelgemeenschappen. Onze resultaten toonden aan dat WH15EPS voornamelijk werd geassimileerd door *Planctomycetes*, *Verrucomicrobia*, *Ascomycota* en *Basidiomycota*. Co-inertieanalyse suggereerde algemene relaties tussen deze bacteriële en schimmelsoorten. We hebben tevens de opname van WH15EPS door *Singulisphaera* en de verbindingen met andere *Planctomycetes* en *Acidobacteria* waargenomen, die niet eerder waren beschreven.

In **hoofdstuk 6** hebben we, door middel van cultuuronafhankelijke en cultuurafhankelijke technieken, WH15EPS toegepast als een verrijkingsfactor om microörganismen en functies te identificeren die betrokken zijn bij de afbraak van EPS. Onze resultaten toonden een grote diversiteit aan glycosidehydrolases, enzymen met groot biotechnologisch potentieel, alsook een groot aantal niet-geclassificeerde microörganismen.

In **hoofdstuk 7** heb ik de algemene bevindingen van mijn proefschrift geïntegreerd en de belangrijkste resultaten over de impact van koolstofbronnen en sporenelementen op de fysiologie van *Granulicella* en, meer in het algemeen, de ecologische functies en het lot van EPS van *Acidobacteria* in het milieu. De resultaten beschreven in dit proefschrift vormen een solide basis voor meer mechanistische studies van *Acidobacteria* en andere niet-gecultiveerde microben.

Resumo

Acidobacteria é um filo de ampla distribuição geográfica, porém suas funções em processos do ecossistema ainda não foram completamente desvendadas. A abundância e a distribuição de *Acidobacteria* em diferentes ambientes, especialmente em solos, origina questões sobre as características fisiológicas que promovem essa abundância. A falta desse conhecimento fundamental deve-se primariamente às dificuldades de isolamento de *Acidobacteria* e seu crescimento lento *in vitro*. O conhecimento das características das diferentes subdivisões de *Acidobacteria* é de extrema importância para a compreensão de sua persistência em solos, assim como suas interações com outros microrganismos de solo.

O gênero *Granulicella* pertence à classe *Acidobacteriia* e é conhecido pela produção abundante de Substâncias Poliméricas Extracelulares (SPE) em meio de cultura. Os SPE são fundamentais para a vida microbiana e proporcionam um substrato ideal para reações químicas, captura de nutrientes e proteção contra estresses ambientais, como salinidade e seca. Os SPE microbianos podem melhorar a agregação de partículas de solo e beneficiar plantas pela manutenção da umidade do ambiente e disponibilidade de nutrientes. Os SPE de duas cepas de *Granulicella* sp. WH15 e 5B5 foram caracterizados anteriormente, demonstrando propriedades emulsificantes interessantes. Assim, os principais objetivos desta tese foram investigar e compreender o metabolismo de duas cepas pertencentes ao gênero *Granulicella* e investigar as funções e o destino ambiental do SPE de *Granulicella* sp. WH15 (WH15EPS). As condições de crescimento de *Granulicella* sp WH15 e 5B5 foram otimizadas em laboratório para aumento de biomassa e produção de SPE, paralelamente ao uso de técnicas multi-ômicas para investigar as adaptações metabólicas. Além disso, o sequenciamento de alto desempenho foi utilizado para a detecção de microrganismos e funções envolvidas na degradação do WH15EPS.

No **capítulo 3**, foi demonstrado que as respostas transcricionais e proteômicas de *Granulicella* sp. WH15 cultivada em diferentes concentrações de celobiose resultaram em maior expressão das funções excretoras e na realocação de recursos para manutenção do metabolismo celular básico, ao invés da produção de novo material celular. No **capítulo 4**, foi observado que a adição da solução de elementos traço SL10 melhorou significativamente o crescimento de *Granulicella* sp. WH15 e 5B5 em meio de cultura. Após avaliar o efeito de cada um dos elementos traço separadamente, os resultados demonstraram que o manganês (Mn) teve um efeito positivo no crescimento de ambas as cepas. Análises proteômicas e genômicas adicionais mostraram que as cepas tinham perfis proteômicos diferentes e vários transportadores não caracterizados que podem estar envolvidos na homeostase de íons metálicos e contribuir para a sobrevivência das cepas em altas concentrações de manganês. A otimização das concentrações de carbono e manganês em meio de cultura permitiu que ambas as cepas, especialmente WH15, crescessem mais rapidamente em condições de laboratório, produzindo quantidades satisfatórias de SPE.

No **capítulo 5**, o WH15EPS foi marcado com ^{13}C e seu efeito foi investigado na montagem e co-ocorrência de comunidades bacterianas e fúngicas ativas em solo superficial, usando-se

como abordagem a análise de isótopos estáveis (SIP). Os resultados demonstraram que o WH15EPS foi assimilado principalmente por *Planctomycetes*, *Verrucomicrobia*, *Ascomycota* e *Basidiomycota* e a análise de co-inércia sugeriu possíveis relações entre esses reinos. Foi observada a incorporação do WH15EPS por *Singulisphaera* e suas conexões com outros *Planctomycetes* e *Acidobacteria*, que não foram observadas anteriormente.

No **capítulo 6**, o WH15EPS foi utilizado como um fator de enriquecimento para microrganismos e funções envolvidas na degradação do SPE por meio de técnicas independentes e dependentes da cultivo. Os resultados demonstraram uma grande diversidade de famílias de glicosídeo hidrolases com potencial biotecnológico e um alto número de microrganismos não classificados que podem ser empregados em novos estudos.

No **capítulo 7**, são feitas a integração e a discussão das conclusões gerais da tese e das observações mais importantes sobre o impacto das fontes de carbono e dos elementos traço na fisiologia da *Granulicella* e, de maneira mais geral, das funções ecológicas e do destino ambiental do EPS de *Acidobacteria*. Os resultados apresentados nesta tese estabelecem uma base sólida para estudos mais mecanísticos de *Acidobacteria* e outros microrganismos não cultivados.

About the author
Publications

Ohana Yonara de Assis Costa was born on 3rd June 1990 in Manaus (Amazonas), Brazil. In 2011, she completed her BSc degree in Biomedicine at the Catholic University of Brasilia (UCB), Brazil. During her bachelor, she worked with microbial community diversity in goat rumen and metagenomic library screening. In 2012 she started her MSc studies in Genomic Sciences and Biotechnology at the Catholic University of Brasilia (UCB), Brazil. Her MSc research was the first work to evaluate the microbial diversity in the ethanol production process using high-throughput sequencing in Brazil, research performed under the supervision of Prof. Dr. Betania F. Quirino. In 2014 she obtained her MSc degree and started working as a research assistant in projects involving microbial community analysis in oil palm trees and protein expression for ethanol production. In 2016 she moved to The Netherlands to start her PhD research at the Department of Microbial Ecology of the Netherlands Institute of Ecology (NIOO-KNAW) within the International collaboration between CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brazil) and NIOO-KNAW on the Ecology and Genomics of *Acidobacteria* under supervision of Prof. Dr. Eiko Kuramae and in collaboration with the Institute of Biology of Leiden University under the supervision of Prof. Dr. Jos Raaijmakers. The findings of her PhD research are described in this thesis.

Costa OYA, De Hollander M, Pijl A, Liu BB, Kuramae EE (2020) Cultivation-independent and cultivation-dependent metagenomes reveal genetic and enzymatic potential of microbial community involved in the degradation of a complex microbial polymer. *Microbiome*, 8(76).

Costa OYA, Zerillo MM, Zühlke D, Kielak AM, Pijl A, Riedel K, Kuramae EE (2020) Responses of *Acidobacteria Granulicella* sp. WH15 to high carbon revealed by integrated omics analyses *Microorganisms*, v. 8, p. 244.

Kuramae EE, Derksen S, Schlemper TR, Dimitrov MR, **Costa OYA**, Silveira APD (2020) Sorghum growth promotion by *Paraburkholderia tropica* and *Herbaspirillum frisingense*: putative mechanisms revealed by genomics and metagenomics. *Microorganisms*, v.8, p. 725.

Costa OYA, Pijl A, Kuramae EE Identification of bacterial and fungal co-occurrence networks during assimilation of acidobacterial extracellular polymers in soil. *(Chapter 5, submitted).*

Costa OYA, Oguejiofor C, Zühlke D, Barreto CC, Riedel K, Kuramae EE (2020) Impact of different trace elements on the growth and proteome of two strains of *Granulicella*, class "Acidobacteriia". *Frontiers in Microbiology,* 11:1227.

Costa OYA & Kuramae EE (2019) *Acidobacteria*, Reference Module in Life Sciences. doi:10.1016/b978-0-12-809633-8.20780-2.

Costa OYA, Raaijmakers JM, Kuramae EE (2018) Microbial Extracellular Polymeric Substances: ecological function and impact on soil aggregation. *Frontiers in Microbiology*, v. 9, p. 1636-1650.

Costa OYA, Tupinambá DD, Bergmann JC, Barreto CC, Quirino BF (2018) Fungal diversity in oil palm leaves showing symptoms of Fatal Yellowing disease. *PLoS One*, v. 13, p. e0191884.

Kielak AM, Castellane TCL, Campanharo JC, Colnago LA, **Costa OYA,** Corradi Da Silva, ML, Van Veen JA, Lemos EGM, Kuramae, EE (2017) Characterization of novel *Acidobacteria* exopolysaccharides with potential industrial and ecological applications. *Scientific Reports*, v. 7, p. 41193.

Tupinamba DD, Cantao M, **Costa OYA,** Bergmann J, Kruger R, Kyaw CM, Barreto CC, Quirino BF (2016) Archaeal community changes associated with cultivation of Amazon Forest soil with oil palm. *Archaea*, v. 2016, p. 1-14.

Costa OYA, Souto B. M, Tupinambá, DD, Bergmann JC, Kyaw CM, Kruger RH, Barreto CC, Quirino BF (2014) Microbial diversity in sugarcane ethanol production in a Brazilian distillery using a culture-independent method. *Journal of Industrial Microbiology and Biotechnology*, v. 42, p. 73-84.

Bergmann JC, Tupinambá DD, **Costa OYA,** Almeida JRM, Barreto CC, Quirino BF (2013) Biodiesel production in Brazil and alternative biomass feedstocks. *Renewable & Sustainable Energy Reviews*, v. 21, p. 411-420.

Bergmann JC, **Costa OYA,** Gladden JM, Singer S, Heins R, D'haeseleer P, Simmons BA, Quirino BF (2013) Discovery of Two novel β-glucosidases from an Amazon soil metagenomic library. *FEMS Microbiology Letters*, v. 11.

Tupinambá DD, Paluan SF, **Costa OYA**, Bitencourt AC, Bergmann JC, Quirino BF (2013) Utilização da diversidade microbiana brasileira para a produção de etanol a partir de biomassa lignocelulósica. *Microbiologia in Foco*.

Cunha IS, Barreto CC, **Costa OYA,** Bomfim MA, Castro AP, Kruger RH, Quirino BF (2011) *Bacteria* and *Archaea* community structure in the rumen microbiome of goats (*Capra hircus*) from the semiarid region of Brazil. *Anaerobe*, v. 17, p. 118-124.

Education Statement

PE&RC Training and Education Statement

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of literature (4.5 ECTS)

- Microbial extracellular polymeric substances: ecological function and impact on soil aggregation (2018)

Writing of project proposal (4.5 ECTS)

- Physiological and ecological survival strategies of soil bacteria belonging to the phylum *Acidobacteria* (2016)

Post-graduate courses (4.6 ECTS)

- Stable isotope applications in microbiology and environmental studies; Wimek-HIGRADE (2017)
- The power of RNA-seq; EPS (2018)
- Introduction to R for statistical analysis; PE&RC (2018)
- Data carpentry genomics; WUR (2019)
- Multivariate analysis; PE&RC (2019)

Deficiency, refresh, brush-up courses (1.5 ECTS)

- Basic statistics; PE&RC (2018)

Competence strengthening / skills courses (3.6 ECTS)

- Effective Communication; Leiden University (2017)
- Communication in science; Leiden University (2018)
- Time Management; Leiden University (2018)
- In Design; WUR Library (2018)
- Scientific Artwork; WUR Library (2018)

Scientific integrity / ethics in science activity (0.4 ECTS)

- PhD Workshop: principles of good academic research; NIOO (2016)
- Scientific conduct; Leiden University (2017

PE&RC Annual meetings, seminars and the PE&RC weekend (1.8 ECTS)

- PE&RC Weekend (2018, 2019)
- PE&RC Day (2018, 2019)

Discussion groups / local seminars / other scientific meetings (5.9 ECTS)

- NIOO Seminar/ME meetings (2016-2019)
- Metagenomics Journal Club (2017-2019)

International symposia, workshops and conferences (8.8 ECTS)

- Thunen symposia; poster presentation; Braunschweig (2016)
- ISME Symposia; poster presentation; Leipzig (2018)
- Ecology of soil microorganisms; poster presentation; Helsinki (2018)
- Thunen symposia; oral presentation; Braunschweig (2019)

Supervision of MSc students (3 ECTS)

- Optimization of cell growth and EPS production of 5B5 and WH15 strains belonging to *Granulicella* sp. of phylum *Acidobacteria*
- A survey of the soil microbial communities inhabiting heavy metal contaminated rhizosphere

The research described in this thesis was performed at the Department of Microbial Ecology of the Netherlands Institute of Ecology - (NIOO-KNAW); O.Y.A. Costa was supported by an SWB grant from CNPq [202496/2015-5] (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brasil).

This is NIOO-thesis number 176

Cover and layout design by Ohana Yonara de Assis Costa.
Printed by GVO drukkers & vormgevers B.V. ||www.gvo.nl