Technische Universität München

Lehrstuhl für Bodenkunde

Bacterial polysaccharides as drivers of soil aggregation

Barbara Cania

Vollständiger Abdruck der von der TUM School of Life Sciences der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:

Prof. Dr. Johannes Kollmann

Prüfende/-r der Dissertation:

1. Hon.-Prof. Dr. Michael Schloter

2. Prof. Dr. Monika Egerer

3. Prof. Dr. Gabriele Berg

Die Dissertation wurde am 29.01.2021 bei der Technischen Universität München eingereicht und durch die TUM School of Life Sciences am 02.06.2021 angenommen.

“However difficult life may seem, there is always something you can do and succeed at. It matters that you don't just give up."

Stephen Hawking

Table of contents

Summary

Soil is a precious but endangered resource. One of the greatest threats to it is accelerated erosion, which leads to huge losses of land every year. However, erosion can be decelerated by the presence of stable soil aggregates, the formation of which can be induced by bacteria. Specifically, some of these microorganisms have the ability to synthesize exopolysaccharides (EPSs) and lipopolysaccharides (LPSs) that "glue" soil particles together. Unfortunately, our knowledge on aggregate stabilization by these compounds is based mainly on studies of isolated strains. Therefore, little is known about the communities of bacterial polysaccharide producers and factors influencing their structure and ability to aggregate soil. The current thesis addresses this research gap by applying metagenomics to investigate potential producers of EPSs and LPSs in agro- and initial ecosystems as environments where soil is particularly prone to erosion. The identification of organisms with the potential to produce adhesive polysaccharides was based on a metagenomics approach where we combined hidden Markov model searches with blasts against sequences derived from the Kyoto Encyclopedia of Genes and Genomes database, which allowed us to target genes specific for either the assembly and export or the extracellular synthesis of EPSs and LPSs. We analyzed not only the relative abundance of these genes and organisms harboring them but also their absolute abundance estimated based on microbial biomass values. Finally, we complemented the metagenomic data with measurements of different parameters including aggregate stability and 3D structure of soil.

Our analysis showed that the bacterial potential to aggregate soil via the synthesis of adhesive polysaccharides depends mainly on the absolute abundance of the genes involved in the production of EPSs and LPSs and the taxonomic profile of the communities of bacteria harboring these genes. Both of these potential determinants of aggregate stability are highly variable and can be affected by environmental factors such as pH, nutrient availability and soil texture. We found amongst potential polysaccharide producers affected by these factors major phototrophic groups including *Cyanobacteria* and *Chloroflexi*, essential nitrifiers such as *Nitrospiraceae*, plant-associated microbes like *Bradyrhizobiaceae*, and other members of *Proteobacteria* as well as *Acidobacteria*, *Actinobacteria* and other common soil groups. In the investigated agroecosystems, we identified tillage as an important factor that can affect the absolute abundance and taxonomic affiliation of the genes related to EPS and LPS biosynthesis, but its effects are modulated by site-specific conditions like soil texture. Members of *Actinobacteria* were the only bacteria that were negatively affected by tillage at all of the investigated sites. Our findings indicate that the relative abundance of the EPS and LPS genes could influence soil aggregation as well, but its high stability in sample types such as agricultural soils and biocrusts from initial ecosystems even under differing environmental conditions and anthropogenic treatments makes

this parameter less likely to be important for shaping soil aggregation in these habitats compared with the more variable absolute abundance and taxonomic affiliation of the EPS and LPS genes. However, the fact that bacterial communities of agricultural soils and biocrusts in initial ecosystems maintain stable proportions of the genes involved in EPS and LPS synthesis despite different taxonomic composition shows that the ability to produce these compounds is an important trait for bacteria living in these environments.

Zusammenfassung

Boden ist eine wichtige Ressource, die besonders durch fortschreitende Bodenerosion gefährdet wird. Dadurch gehen jedes Jahr große Mengen an fruchtbarem Boden verloren. Stabile Bodenaggregate können dieser Entwicklung entgegenwirken. Die Bildung solcher stabilen Aggregate wird unter anderem durch Bakterien begünstigt, die durch die Produktion von Exopolysacchariden (EPSs) und Lipopolysacchariden (LPSs) die Bodenpartikel zusammenhalten. Bisher wurde das bakterielle Potential zur Produktion von LPS und EPS hauptsächlich an isolierten Stämmen untersucht, während wenig darüber bekannt ist wie sich die Zusammensetzung der EPS/LPS produzierenden, bakteriellen Gemeinschaft unter Feldbedingungen verändert und wie sich dies auf die Bodenaggregierung auswirkt. Die aktuelle Arbeit adressiert diese Forschungslücke, indem EPSs und LPSs produzierende, bakterielle Gemeinschaften in Agrarökosystemen und initialen Ökosystemen, insbesondere biologischen Bodenkrusten, untersucht wurden. Beide sind dafür bekannt besonders anfällig für Erosion zu sein. Um die relevanten Bakterien und Prozesse zu identifizieren haben wir einen Metagenomansatz gewählt. Dazu haben wir zwei bioinformatische Ansätze kombiniert, nämlich die Suche nach Hidden Markov Modellen mit einem Abgleich basierend auf Sequenzen der Datenbank „Kyoto Encyclopedia of Genes and Genomes“. Der Fokus der Analysen lag auf Genen, die für Proteine kodieren, die an der Assemblierung von EPS/LPS Molekülen beteiligt sind, dem darauffolgenden Export in die Umwelt oder der extrazellulären Synthese. Um auch eine Abschätzung der absoluten Abundanz der beteiligten Bakterien zu realisieren, wurden die erhaltenen relativen Abundanzen aus den Metagenomanalysen mit Daten über die mikrobielle Biomasse in den Proben kombiniert. Außerdem wurden die Metagenomdaten mit Messungen der Aggregatstabilität und der 3D Porenstruktur des Bodens kombiniert.

Die Ergebnisse zeigen, dass das bakterielle Potenzial zur Synthese von adhäsiven Polysacchariden hauptsächlich durch die absolute Abundanz der Gene bestimmt wird, und der taxonomischen Zusammensetzung der Bakteriengemeinschaften, die an der Produktion von EPSs und LPSs beteiligt sind. Die Abundanz und Diversität der EPS und LPS Produzenten bestimmt die Aggregatstabilität im Boden. Beide Faktoren werden maßgeblich durch Umweltfaktoren wie pH-Wert, Nährstoffverfügbarkeit und Bodentextur beeinflusst. Zu den bakteriellen EPS/LPS Produzenten, die sich am stärksten in den verschiedenen Umweltproben unterschieden, zählen phototrophe Bakterien, wie *Cyanobacteria* und *Chloroflexi*, Nitrifizierer wie *Nitrospiraceae*, pflanzenassoziierte Bakterien wie *Bradyrhizobiaceae*, weitere *Proteobacteria*, *Acidobacteria*, und *Actinobacteria* und weitere typische Bodenbakterien. In den untersuchten Agrarsystemen, haben wir die Bodenbearbeitung als wichtigen Faktor identifiziert, der sowohl die Häufigkeit als auch die taxonomische Zuordnung der verschiedenen LPS/EPS Produzenten

beeinflusst. Dabei spielte insbesondere die Intensität der Bodenbearbeitung eine wichtige Rolle, aber auch die Standort-spezifischen Randbedingungen wie die Bodentextur. Nur die absolute Abundanz der *Actinobacteria* wurde an allen Standorten negativ durch die Bodenbearbeitung beeinflusst. Obwohl die absoluten Abundanzen der EPS/LPS produzierenden Bakterien in den verschiedenen bewirtschafteten Agrarsystemen und den initialen Ökosystemen variieren, war der relative Anteil an der Gesamtgemeinschaft in allen Versuchen vergleichbar. Lediglich die taxonomische Zuordnung der verschiedenen Gene, die an der EPS/LPS Produktion beteiligt sind variieren stark mit dem Habitat und den Umweltbedingungen. Dies macht deutlich, dass die Synthese von EPS/LPS eine wichtige Funktion in der bakteriellen Gemeinschaft einnimmt.

List of publications and contributions

This thesis is based on the following publications:

- Cania B., Vestergaard G., Kublik S., Köhne J.M., Fischer T., Albert A., Winkler B., Schloter M., and Schulz S. (2020). Biological soil crusts from different soil substrates harbor distinct bacterial groups with the potential to produce exopolysaccharides and lipopolysaccharides. *Microb Ecol* 72, 326-341.
  (P1, first author, published)
- Cania B., Vestergaard G., Krauss M., Fliessbach A., Schloter M., and Schulz S. (2019). A long-term field experiment demonstrates the influence of tillage on the bacterial potential to produce soil structure-stabilizing agents such as exopolysaccharides and lipopolysaccharides. *Environ Microbiome* 14, 1.
  (P2, first author, published)
- Cania B., Vestergaard G., Suhadolc M., Mihelič R., Krauss M., Fliessbach A., Mäder P., Szumełda A., Schloter M., and Schulz S. (2020). Site-specific conditions change the response of potential producers of soil structure-stabilizing agents such as exopolysaccharides and lipopolysaccharides to tillage intensity. *Front Microbiol* 11, 568.
  (P3, first author, published)

Further relevant contributions that are mentioned but not included in this thesis:

- Vuko M., **Cania B.**, Vogel C., Kublik S., Schloter M., and Schulz S. (2020). Shifts in reclamation management strategies shape the role of exopolysaccharide and lipopolysaccharide-producing bacteria during soil formation. *Microb Biotechnol* 13, 584-598.
  (P4, second author, published)

Publication 1

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

Cania Barbara, Vestergaard Gisle, Kublik Susanne, Köhne John Maximilian, Fischer Thomas, Albert Andreas, Winkler Barbro, Schloter Michael, Schulz Stefanie

Short description:

Biocrusts are well-known for stabilizing soil, and many bacterial strains with the ability to produce soil structure-stabilizing agents such as exo- and lipopolysaccharides were isolated from them before. As the data on the community dynamics of the bacterial producers of these compounds during the development of biocrusts was missing, a microcosm experiment was performed where biocrusts were cultivated on two different soil substrates. Shotgun sequencing was used to compare the potential producers of exopolysaccharides and lipopolysaccharides in the initial soils and the developed biocrusts. The potential to form these compounds increased in the biocrusts compared with the soil substrates, but the community composition of potential polysaccharide producers differentiated once the biocrusts started to develop, which shows that the potential to produce exo- and lipopolysaccharides is a trait maintained by the bacterial communities of biocrusts despite shifts in their composition.

Contributions:

- carried out bioinformatical analysis
- conducted statistical analysis, data visualization and interpretation
- wrote the manuscript

Publication 2

A long-term field experiment demonstrates the influence of tillage on the bacterial potential to produce soil structure-stabilizing agents such as exopolysaccharides and lipopolysaccharides

Cania Barbara, Vestergaard Gisle, Krauss Maike, Fliessbach Andreas, Schloter Michael, Schulz Stefanie

Short description:

As tillage intensity affects the stability of soil aggregates, the bacterial potential to produce exo- and lipopolysaccharides that can influence soil aggregation was investigated in a long-term agricultural field trial. Potential polysaccharide producers were compared in and below the tillage horizon under two tillage intensities (conventional and reduced) using shotgun sequencing. The bacterial potential to form exo- and lipopolysaccharides was equally promoted in the tillage horizon under both tillage intensities, but some potential producers of these compounds differed, which indicates that tillage intensity could affect soil aggregation by selecting different key players of the formation of bacterial polysaccharides.

Contributions:

- contributed to planning and designing the experiment
- performed DNA extraction and library preparation
- carried out bioinformatical analysis
- conducted statistical analysis, data visualization and interpretation
- wrote the manuscript

Publication 3

Site-specific conditions change the response of potential producers of soil structure-stabilizing agents such as exopolysaccharides and lipopolysaccharides to tillage intensity

Cania Barbara, Vestergaard Gisle, Suhadolc Marjetka, Mihelič Rok, Krauss Maike, Fliessbach Andreas, Mäder Paul, Szumełda Anna, Schloter Michael, Schulz Stefanie

Short description:

Previous data on soil aggregation suggests that tillage intensity could differentially affect the bacterial potential to form exo- and lipopolysaccharides in soils with different textures. Thus, this study employed shotgun sequencing to investigate the impact of two tillage intensities (conventional and reduced) on potential polysaccharide producers in three agricultural field trials with different soil textures (clayey, silty and sandy). While some responses of potential polysaccharide producers to tillage intensity depended on their taxonomic affiliation, e.g. *Actinobacteria* showed higher potential to form exo- and lipopolysaccharides under reduced tillage, others were site-specific. This indicates that soil texture indeed plays an important role in defining the effects of tillage intensity on bacterial soil aggregation.

Contributions:

- contributed to planning and designing the experiment
- performed DNA extraction and library preparation
- carried out bioinformatical analysis
- conducted statistical analysis, data visualization and interpretation
- wrote the manuscript

Publication 4

Shifts in reclamation management strategies shape the role of exopolysaccharides and lipopolysaccharide-producing bacteria during soil formation

Vuko Miljenka, Cania Barbara, Vogel Cordula, Kublik Susanne, Schloter Michael, Schulz Stefanie

Short description:

As the contribution of bacteria to soil aggregation through the formation of exo- and lipopolysaccharides should be especially crucial in newly established ecosystems without proper soil structure, a managed land reclamation site of a post-mining area was investigated. Shotgun sequencing was used to obtain information on the community dynamics of potential polysaccharide producers during the initial and agricultural management phases of the reclamation process. The potential polysaccharide producers shifted from oligotrophic taxa in the initial phase to bacteria typical for agricultural soils in the agricultural management phase, and their highest abundance was registered directly after the start of the agricultural management phase. This shows that the communities of potential polysaccharide producers are highly adaptable and triggered especially by agricultural management practices.

Contributions:

- carried out bioinformatical analysis
- contributed to statistical analysis, data visualization and interpretation
- critically revised the manuscript

1 Introduction

1.1 Soil – a precious but endangered resource

Soil is one of our most precious resources. It directly or indirectly provides us with food, clothing, shelter and medications, which are basic necessities. Without soil, there would be no human life. However, soil supports not only our species but is one of the largest reservoirs of our planet's biodiversity. In fact, it hosts around one quarter of all living species on Earth (Bach and Wall, 2018). In just a few grams of soil, there are more individual organisms than there are people on our planet. The most abundant amongst soil biota are microbes, which play pivotal roles for ecosystem functioning, such as driving nutrient and organic matter cycling (Falkowski et al., 2008; Wagg et al., 2014). Terrestrial ecosystems host also the biggest portion of primary producers. Specifically, the biomass of producers on land is 450 times larger than in the oceans (Bar-On et al., 2018). Therefore, the health of our planet's environment and all its inhabitants strongly depends on the status of world soils.

However, the demands of the growing human population are putting soil sustainability at risk. Recent evidence shows that already one third of global soils are moderately to highly degraded (FAO, 2011; FAO and ITPS, 2015). Soil degradation means a long-term reduction in the capacity of the soil to meet social and ecological needs (Lal, 2001). Especially threatened by soil erosion is crop production, as due to the increase in population and the decline in soil quality, the area of arable land available per person already decreased by more than half during the last 60 years and is still shrinking (Flachowsky et al., 2017). Consequently, even unsuitable land is being brought under agriculture, which then leads to increased production costs. The economic losses caused by soil degradation are currently estimated at 7.25 billion euros of global gross domestic product per year (Sartori et al., 2019). Furthermore, there is little new land that could be turned into farmlands, and soil formation takes hundreds or even thousands of years (Kalev and Toor, 2018). Overall, soil is a very precious and hardly renewable resource that is being used up at a dangerous rate and, therefore, requires more attention and efforts aimed at its protection and restoration.

The primary cause of soil degradation is accelerated erosion (FAO and ITPS, 2015). Soil erosion itself is the removal of soil materials from their original location (Arriaga et al., 2017). This can be induced for example by water on steep slopes or wind in open areas with scarce vegetation (Montanarella, 2016). When soil erosion results from natural processes, it is referred to as geological erosion (Arriaga et al., 2017). Geological erosion is a very slow process, which can be hardly observed during a single human lifetime. The examples of its long-term effects are canyons, stream channels and valleys (Gilley, 2005). However, soil erosion can be greatly accelerated by human activities, especially poor agricultural practices (Finch et al., 2014). Accelerated erosion occurs at an alarming speed, leading to huge soil

losses. Montgomery (2007) estimated that the erosion rates from conventionally ploughed agricultural fields are 1-2 orders of magnitude higher than the rates of soil formation and geological erosion. According to FAO (FAO, 2019), this could cause the loss of up to 3.9 mm of soil per year. Moreover, the erosion rates seem to be increasing every year due to the intensification of agricultural production. In fact, Borrelli et al. (2017) estimated that the global erosion rates increased between 2001 and 2012 by 2.5 %. This was mainly due to deforestation and cropland expansion in the least developed countries of Africa, South America and Southeast Asia, even though the ongoing adaptation of conservation agriculture decreased the erosion primarily in the United States and many European countries. Yet soils have different susceptibility to erosion depending on their parameters (Finch et al., 2014). The most important soil parameters influencing erodibility are presented in Figure 1.

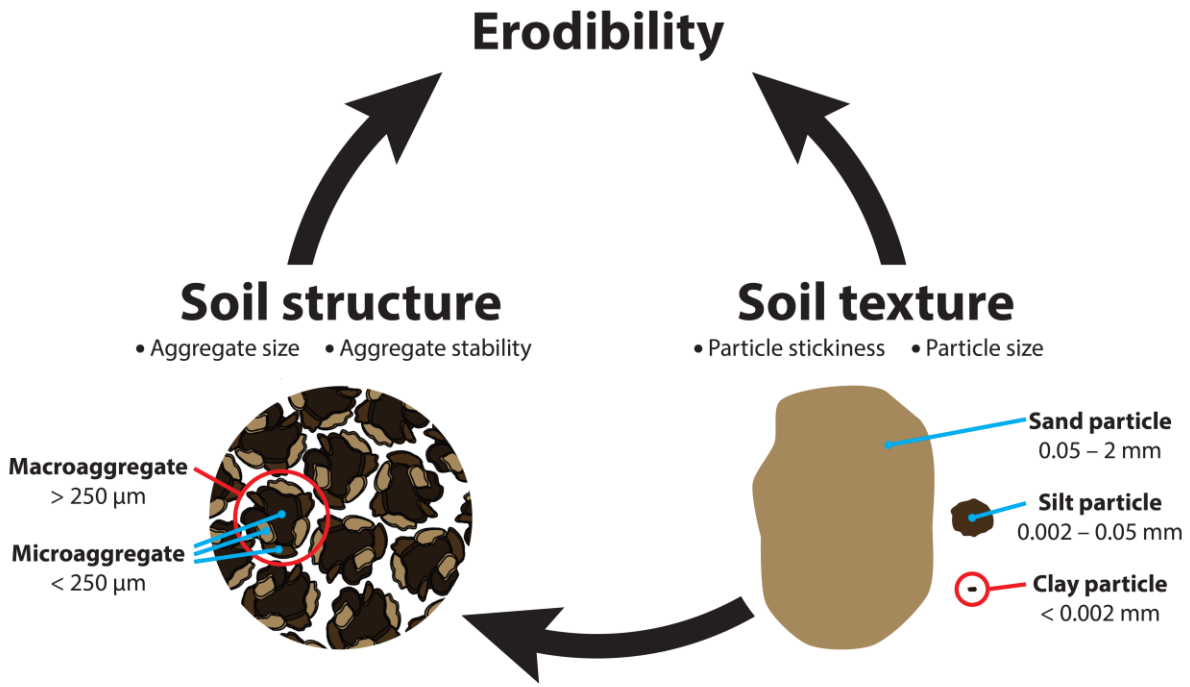


Figure 1. Soil parameters and their most important aspects influencing erodibility. The dependencies between those parameters and erodibility are marked with directional arrows.

1.2 Soil properties influencing erodibility

1.2.1 Soil structure

The major soil parameter that influences the erodibility of a soil is structure (Figure 1). Soil structure refers to the spatial arrangement of individual soil particles. Some soils may have a single-grained structure without any aggregation. One example of that are sand dunes. However, most soils exhibit

a structure in which particles form larger aggregates. Pores and spaces between aggregates and individual particles form a network that affects the flow of water and air through the soil profile, the growth of plant roots and biological activity (Weil and Brady, 2017). Soils with a lot of pores and fissures are said to have a good structure, whereas soils with a poor pore network are compacted and more prone to erosion (Finch et al., 2014). Soil structure is greatly determined by the shape and size of aggregates. Shapes include granular, platy, blocky and columnar forms. Blocky and columnar structures characterize subsoil horizons. Platy forms can be found in both surface and subsoil horizons, and they are mainly a result of soil-forming processes. Granular structures are present in surface horizons of most soils, especially those rich in organic matter. Their exposure to erosive factors is therefore highest, which puts them in focus of this thesis (Weil and Brady, 2017).

Granular aggregates found in soil surface horizons can be divided into macroaggregates (> 250 μm) and microaggregates (< 250 μm) based on their diameter. These aggregates exhibit a hierarchy in which macroaggregates are composed of multiple microaggregates (Totsche et al., 2018). The size of an aggregate greatly determines its susceptibility to erosion, which applies to individual particles as well. In principle, larger particles or aggregates are more resistant to erosion compared with the smaller ones. However, macroaggregates are at the same time more prone than microaggregates to being fragmented by disruptive forces such as water, wind or physical disturbance. Therefore, the ability of granular aggregates to withstand erosion depends not only on their size but also on their stability (Torri et al., 1998). While these two properties of aggregates are important aspects of how soil structure influences erodibility, they are greatly determined by another soil parameter known as soil texture, which can affect soil erodibility on its own as well.

1.2.2 Soil texture

Soil texture (Figure 1) refers to the proportion of different-sized particles in a given soil. Several systems of particle-size classification exist, but the one proposed by the U.S. Department of Agriculture (USDA) is most widely used. According to this system, soil texture is determined by three size fractions: clay (< 0.002 mm), silt (0.002 – 0.05 mm) and sand (0.05 – 2 mm) (Yolcubal et al., 2004). Mineral bodies larger than 2 mm in diameter (e.g. gravels, stones, cobbles and boulders) are not considered as part of soil material (Hillel, 2008). As in the case of soil aggregates, the erodibility of individual soil particles is greatly determined by their size (Torri et al., 1998). However, particles of each size fraction differ also in their propensity to form aggregates, which in turn influences the structure of a soil and its ability to withstand erosion (Weil and Brady, 2017).

Sand particles are the largest, with a relatively small surface area. Because of that, they possess little capacity to hold water or nutrients, and the pores between them tend to be filled mainly with air. Sand itself exhibits insufficient stickiness to form aggregates by themselves, but only fine sand particles can be easily carried away by wind or water. In sharp contrast with sand, stand the smallest particles – clay. As their surface area is relatively large, clay particles have a great capacity to retain water and organic matter. This characteristic makes them highly resistant to wind erosion. Clay particles also have the highest propensity to attract each other and form aggregates. If not aggregated, however, they are highly susceptible to water erosion due to their colloidal nature. An intermediate fraction between sand and clay is silt. While silt particles are better than sand at adsorbing water and other substances, they hardly stick together on their own. The little stickiness that silt may exhibit usually comes from a film of adhering clay. Therefore, due to their low cohesiveness, soils rich in silt and fine sand are most prone to erosion by both water and wind (Finch et al., 2014; Weil and Brady, 2017). Nevertheless, all size fractions can form aggregates and decrease soil erodibility, although the aggregation process and the parameters of the resulting aggregates might differ (Weil and Brady, 2017; Totsche et al., 2018).

1.3 Soil aggregation as a process preventing erosion

Aggregate formation and stabilization make up a bigger process termed soil aggregation (Amezketa, 1999). Several concepts of soil aggregation exist, but generally accepted is the hierarchical model proposed by Tisdall and Oades (Tisdall and Oades, 1982). According to this concept, individual soil particles are bound together and cemented by persistent forces into microaggregates, which in turn are assembled and stabilized as macroaggregates by more transient forces. Microaggregates can also form within macroaggregates (Tisdall and Oades, 1982; Oades and Waters, 1991; Six et al., 2004). The difference between the strength of forces binding soil particles and microaggregates together leads to the generally higher stability of microaggregates over macroaggregates (Totsche et al., 2018). Macroaggregates are also more exposed than microaggregates to disruptive forces, such as agricultural practices, which can decrease their stability or cause their breakdown (Bird et al., 2007; Lehmann et al., 2017a; Totsche et al., 2018). The persistent forces that participate in soil aggregation are mostly a result of physicochemical processes, whereas the transient forces are more related to biological processes (Tisdall and Oades, 1982; Six et al., 2004; Totsche et al., 2018). Therefore, the physicochemical factors are more involved in the formation and stabilization of microaggregates, whereas macroaggregates are formed and cemented mainly by biological activity, although there are exceptions to this rule. Moreover, physicochemical and biological aggregation processes happen simultaneously and interact

with each other, and thus the comprehension of their actual contribution to soil aggregation is still lacking (Lehmann et al., 2017a; Lehmann et al., 2017b; Totsche et al., 2018).

1.3.1 Physicochemical aggregation

Considered as one of the most important physicochemical processes for soil aggregation is the flocculation of soil particles. For this reason, physicochemical aggregation processes might play an especially important role in soils with high clay content, whereas soil aggregation in silty and sandy soils should depend rather on biological factors (Weil and Brady, 2017). As the basal surfaces of clay particles are negatively charged, once they are positioned close enough to each other, cations present between them form bridges bonding the particles strongly together. The positive charges on the edges of clay particles also allow them to bind directly to each other as well as to negatively charged organic molecules. Such organic molecules can also form bonds with clay particles through multivalent cations. Aside from clay, other minerals such as Fe- and Al-(hydr)oxides are important aggregate-forming materials, although they are not as well-studied (Weil and Brady, 2017; Totsche et al., 2018).

Independently from the type of material, the reaction partners need to approach each other physically in order to be able to form a bond. This movement can be aided by different physical processes. Especially important seem to be wetting-drying cycles. As water infiltrates soil and is withdrawn over the course of multiple cycles, it moves small soil particles though capillary forces. The particles as well as aggregates can also be rearranged through shrinking and swelling of soil mass caused by water menisci forces. Similar effect have freezing-thawing cycles. The wetting-drying cycles can be facilitated as well in the root area through the water uptake by plants, which is an example of the interaction of physical and biological processes during soil aggregation (Weil and Brady, 2017; Totsche et al., 2018).

1.3.2 Biological aggregation

The major factors of biological soil aggregation are: i) soil fauna, ii) plant roots and iii) microorganisms (Six et al., 2004). Soil animals such as earthworms and thermites can drive soil aggregation by physically rearranging soil particles or mixing them with their bodily fluids (e.g. mucus, feces and saliva) that act as biding agents (Six et al., 2004; Jouquet et al., 2016). Plant roots mechanically displace soil particles during growth, exert pressure on the surrounding soil, change soil water regime and release organic materials with cementing properties into the soil. Fungi entangle soil particles within their hyphal networks and secrete biopolymers that glue soil particles together (Six et al., 2004; Totsche et al., 2018). Of particular importance is a glycoprotein synthesized by arbuscular mycorrhizal fungi called glomalin,

whose effects on aggregate stability were shown to be much stronger than the physical stabilization by the hyphae alone (Rillig et al., 2002). Bacteria also produce compounds (especially polysaccharides) with aggregating properties, and their filaments likely trap soil particles in a similar manner as fungal hyphae (Belnap and Gardner, 1993; Six et al., 2004; Smith et al., 2004; Lehmann et al., 2017b).

Besides influencing soil aggregation directly, soil fauna, plant roots and microorganisms may affect the aggregating capabilities of each other. For example, animal excretions and plant mucilages alter the activity of soil microorganisms, while symbiotic bacteria and fungi determine the biological processes of plants and animals (Six et al., 2004; Lehmann et al., 2017a; Enagbonma and Babalola, 2019). Soil aggregation is therefore a complex process that depends on a multitude of various mechanisms and interactions between different physicochemical and biological factors. Taking into consideration such complexity, the knowledge gaps regarding the processes involved in soil aggregation are understandable. In order to fill these gasps, further understanding of individual mechanisms driving soil aggregation and how they are affected by different factors is necessary (Six et al., 2004).

Of special importance for preventing soil erosion are the biological agents of aggregation, as those are more vulnerable to human activities compared with the physicochemical factors. According to the recent study by Lehmann et al. (2017b), the most important groups of soil biota influencing soil aggregation are bacteria and fungi. Lehmann et al. found that fungi contribute mainly to the formation of macroaggregates, whereas bacteria strongly affect the formation of both macro- and microaggregates. Therefore, the mechanisms that foster soil aggregation employed by bacteria, especially the formation of polysaccharides with gluing properties, deserve more scientific interest than they have received until now.

1.4 Bacterial polysaccharides as important agents of soil aggregation

Two types of bacterial polysaccharides play an important role in soil aggregation, namely exopolysaccharides (EPSs) and lipopolysaccharides (LPSs). These compounds facilitate the attachment of bacteria to soil particles, which mediates the formation of soil aggregates (Jacques, 1996; Sutherland, 2001a; Six et al., 2004; Totsche et al., 2018). While both EPSs and LPSs have gluing properties, their structures and biosynthetic pathways are different. As it has been shown that even slight structural differences between polysaccharides may result in entirely different physical traits, the structure is an especially important characteristic for EPSs and LPSs, influencing their soil aggregation capabilities (Suresh Kumar et al., 2007).

1.4.1 EPSs

Bacterial EPSs, as their name suggests, are extracellular polymers composed of sugar residues. Additionally, non-carbohydrate substituents such as acetate, pyruvate, succinate and phosphate may be attached (Sengupta et al., 2018). These polymers can form a capsule that is bound, often covalently, to the cell surface, or a slime that is dispersed in the surrounding environment (Sutherland, 1972). EPSs are produced by many different species of both Gram-negative and Gram-positive bacteria (Suresh Kumar et al., 2007). While these compounds are formed by such a wide variety of bacterial taxa, even closely related organisms may produce different types of EPSs (Sutherland and Thomson, 1975; Celik et al., 2008). Furthermore, certain bacterial species are able to synthesize more than one variant of these polymers (Kwon et al., 1994; Matsuyama et al., 2003; Dertli et al., 2013). Moreover, EPS produced by one strain may slightly vary depending on the nutritional and environmental conditions (Suresh Kumar et al., 2007). As a result, EPSs are a very diverse group of compounds. This diversity applies to their sugar composition, sequence of monomeric units as well as molecular size (Suresh Kumar et al., 2007; Cuthbertson et al., 2009; Sengupta et al., 2018). Exemplary chemical structures of different EPSs are presented in Figure 2.

1.4.1.1 Structure

EPSs can be classified based on their sugar composition as homopolysaccharides and heteropolysaccharides. Homopolysaccharides are composed of a single repeating monosaccharide residue, and, according to their linkage bonds and monomeric units, they are grouped into α-D-glucans, β-D-glucans, fructans and polygalactans (Nwodo et al., 2012). Examples of homopolysaccharides are dextran, inulin, levan, curdlan and cellulose (Figure 2a-e) (Schmid et al., 2015). Conversely, heteropolysaccharides contain two or more different sugars (Nwodo et al., 2012). Most commonly found in heteropolysaccharides are D-glucose, D-galactose and L-rhamnose, although L-fucose, L-altrose, L-iduronic acid, N-acetyloglucosamine, N-acetylogalactosamine, glucuronic acid as well as other components are sometimes present (Nwodo et al., 2012; Roca et al., 2015). These units are usually joined by very rigid 1,4- β- and 1,3-β- bonds, or more flexible 1,2-α- and 1,6-α- linkages (Nwodo et al., 2012). Examples of heteropolysaccharides are alginate, hyaluronic acid, sphingans, xanthan, colanic acid and succinoglycan (Figure 2f-k) (Schmid et al., 2015). Furthermore, the repeating units in the backbone of both homo- and heteropolysaccharides can be either linear or branched. The molecular weight of these compounds is also diverse and may vary from 8 to over 5000 kDa (Zeidan et al., 2017). As a result, the number of possible EPS variations is enormous. In the Bacterial Exopolysaccharide Properties and Structures Database (EPS-DB) (Birch et al., 2019), already over a hundred published

polysaccharide structures have been deposited, most (> 80 %) from lactic acid bacteria. Moreover, the database has been established only recently and is still growing. This underlines the difficulties and challenges of the research on bacterial EPSs.

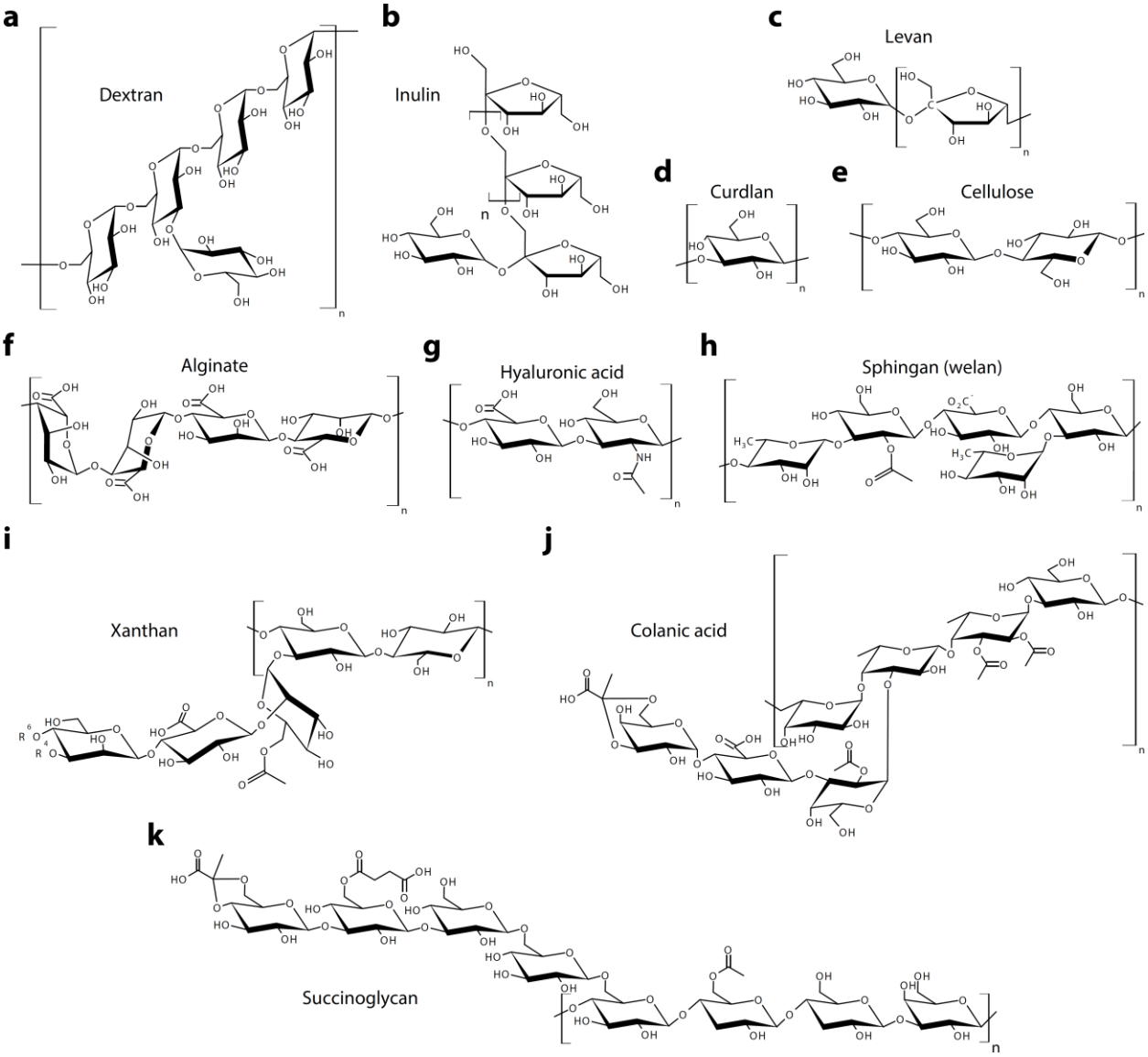


Figure 2. Chemical structures of EPSs mentioned in this thesis: a) dextran, b) inulin, c) levan, d) curdlan, e) cellulose, f) alginate, g) hyaluronic acid, h) sphingan, i) xanthan, j) colanic acid and k) succinoglycan.

1.4.1.2 Biosynthesis

Regardless of the great diversity of bacterial EPSs, these polymers are known to be produced only via four main biosynthetic pathways (Figure 3). Specifically, most EPSs are initially synthesized intracellularly

and exported by one of the following mechanisms: i) the Wzx/Wzy-dependent pathway, ii) the ABC transporter-dependent pathway or iii) the synthase-dependent pathway. Alternatively, they can be synthetized directly outside the cell (Schmid et al., 2015). The first three EPS biosynthetic pathways start in the cytoplasm with the activation of monosaccharides and their conversion into nucleotide sugars, which are then used as building blocks for the polymer strand. For the extracellular EPS biosynthesis, the precursor molecules are obtained by cleaving di- and trisaccharides (Pereira et al., 2013; Schmid et al., 2015). Although not all the following steps in the various EPS biosynthesis systems are yet understood, and especially the decoration with substituents often remains elusive, the main proteins involved in the assembly and export processes are rather well-characterized (Schmid et al., 2015, Schmid, 2018).

The Wzx/Wzy-dependent pathway (Figure 3a) is the most common bacterial mechanism of EPS biosynthesis (Whitfield and Larue, 2008; Whitfield, 2010). It can be found in both Gram-negative and Gram-positive bacteria, although its final steps differ for these two groups (Zeidan et al., 2017). Examples of the polymers synthesized via the Wzx/Wzy-dependent pathway are sphingans, xanthan, colanic acid, succinoglycan (Figure 2h-k) and different capsular polysaccharides (Schmid et al., 2015). In this pathway, the nucleotide sugars are first transferred to a lipid carrier (undecaprenyl diphosphate) located at the cytoplasmic membrane and then assembled into repeating units by several soluble and/or membrane-bound glycosyltransferases (GTs). Therefore, all polymers produced via the Wzx/Wzy-dependent pathway are heteropolysaccharides with highly diverse sugar patterns (i.e. four or five different monomeric units is common). The repeating units assembled on the lipid carrier are subsequently translocated (flipped) across the membrane by the integral protein Wzx and then polymerized by another integral protein – Wzy. In Gram-negative bacteria, the control of chain length and the transport of the polymerized repeating units outside the outer membrane depend on additional protein(s) from the polysaccharide copolymerase (PCP) and the outer membrane polysaccharide export (OPX) families, which form a complex spanning the whole cell envelope (Pereira et al., 2013; Schmid et al., 2015; Zeidan et al., 2017). In Gram-positive bacteria, these protein families are absent, and the chain length of the polymer is instead controlled by a surrogate family of modulation proteins that also act as a scaffold for the assembly machinery (Zeidan et al., 2017).

The ABC transporter-dependent pathway (Figure 3b) is specific for the biosynthesis of capsular polysaccharides in Gram-negative bacteria (Zeidan et al., 2017). As in the Wzx/Wzy-dependent pathway, capsular polysaccharides are assembled at the cytoplasmic membrane by GTs. However, depending on the number of different GTs involved in the assembly process (i.e. one or more), the final polymer may be either a homo- or heteropolysaccharide (Schmid et al., 2015). Another difference is the presence of a linker composed of multiple β-linked 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) residues that anchors

the polymer strand to the lipid carrier in the membrane and is carried by all capsular polysaccharides produced via the ABC transporter-dependent pathway (Schmid et al., 2015; Zeidan et al., 2017; Sande et al., 2019). Moreover, capsular polysaccharides are fully polymerized already at the cytoplasmic face of inner membrane and then exported as complete molecules outside the cell by an efflux pump-like complex. This complex comprises an inner membrane ABC-transporter as well as periplasmic proteins from the PCP and OPX families, and it spans the cell envelope. These PCP and OPX proteins are closely related to those used in the Wzx/Wzy-dependent pathway (Pereira et al., 2013; Schmid et al., 2015).

The synthase-dependent pathway (Figure 3c) is utilized by both Gram-negative and Gram-positive bacteria, mainly for the assembly of homopolysaccharides, although simple heteropolysaccharidic products (i.e. maximum two different monomeric units) also occur (Delbarre-Ladrat et al., 2014; Tytgat and Lebeer, 2014). Examples of polymers produced via this pathway are curdlan, cellulose, alginate, hyaluronic acid (Figure 2d-g) and various capsular polysaccharides (Schmid et al., 2015). The synthase-dependent pathway is clearly distinct from the previously described systems and more diverse (Whitney and Howell, 2013; Tytgat and Lebeer, 2014; Schmid et al., 2015; Low and Howell, 2018). Depending on the polymer, the biosynthesis may be initiated in the presence or absence of a lipid carrier in the cytoplasmic membrane (Whitney and Howell, 2013). Polymer elongation and export may be catalyzed by a single enzyme (e.g. type 3 capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*) or carried out by a multiprotein complex (e.g. alginate biosynthesis in *Pseudomonas* and *Azotobacter*) (Tytgat and Lebeer, 2014; Low and Howell, 2018). Such complex typically consists of an inner membrane-embedded GT that facilitates simultaneous polymer formation and translocation across the membrane. The GT is usually accompanied by an inner membrane receptor, also called copolymerase (but different than the PCP proteins employed by the Wzx/Wzy- and ABC transporter dependent pathways), which post-transcriptionally regulates the polymerization by binding the secondary messenger bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP). In the periplasm, the polymer is protected from degradation by a tetratricopeptide repeat (TPR)-containing scaffold protein. This protein is coupled to an outer membrane β-barrel porin, which in turn exports the polymer outside the cell (Whitney and Howell, 2013; Low and Howell, 2018).

The extracellular biosynthesis (Figure 3d) is employed by Gram-positive bacteria for the production of homopolysaccharides such as dextran, inulin and levan (Figure 2a-c) (Sutherland, 2001b; Ates, 2015; Schmid et al., 2015). Compared with the other EPS biosynthesis pathways, this route is relatively simple. It involves a specific GT (sucrase) that is secreted outside the cell and covalently linked to the cell wall. This enzyme catalyzes the transfer of a monosaccharide from an extracellular oligosaccharide onto a growing polymeric chain. As the synthesis of new glycosidic bonds in the elongated

homopolysaccharides is powered by the energy obtained from cleaving the glycosidic bonds in the sugar donor, this system is essentially independent from the central metabolism (Ates, 2015; Schmid et al., 2015; Zeidan et al., 2017).

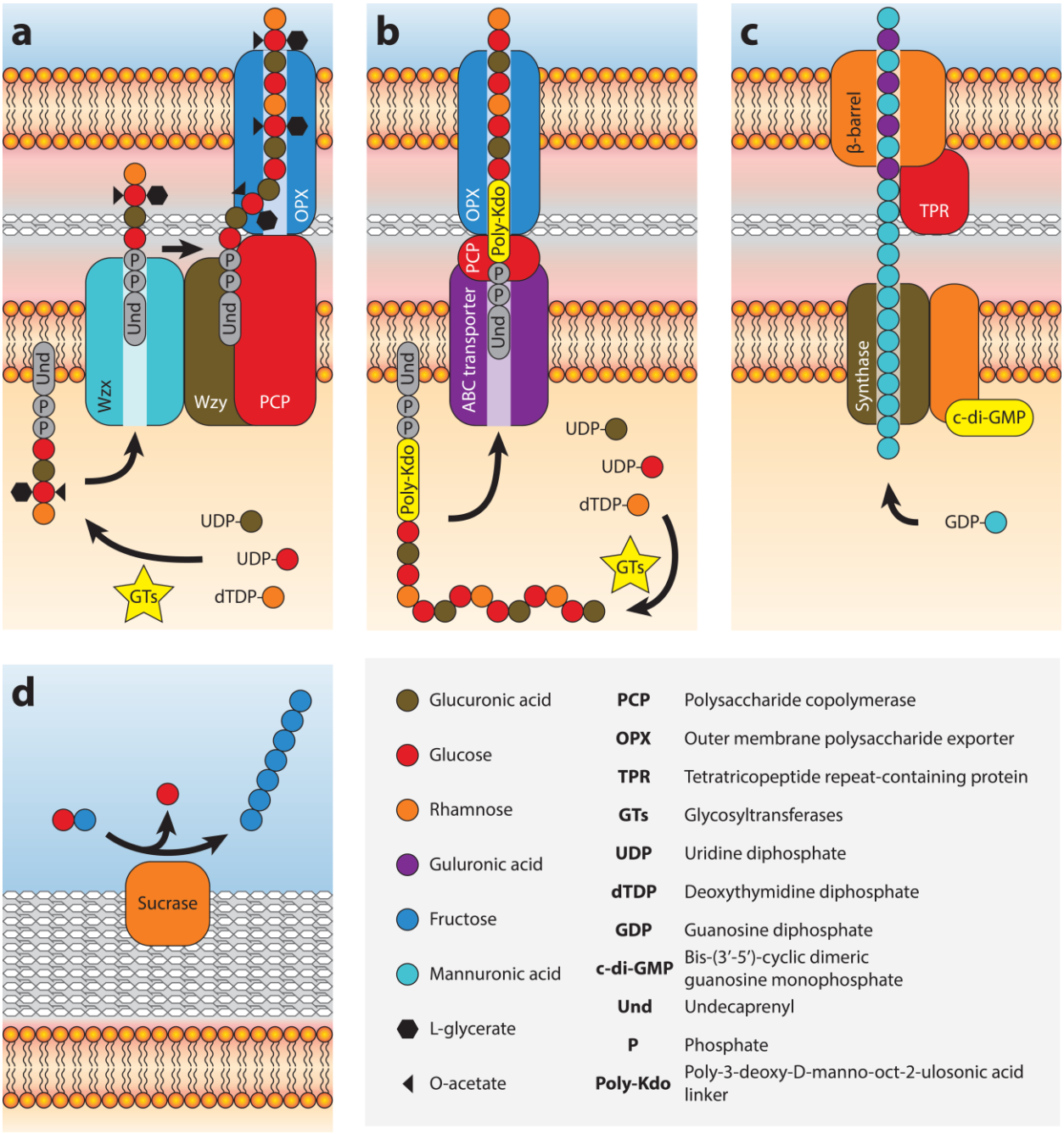


Figure 3. Pathways of the biosynthesis of EPSs found in bacteria: a) Wzx/Wzy-dependent pathway, b) ABC transporter-dependent pathway and c) synthase-dependent pathway represented for Gram-negative bacteria, as well as d) extracellular biosynthesis represented for Gram-positive bacteria (adapted from Schmid, 2018 and Zeidan et al., 2017).

1.4.2 LPSs

LPSs are major components of the outer membrane of most (but not all) Gram-negative bacteria (Sutcliffe, 2010). These compounds typically cover ~75 % of the bacterial cell surface. While LPSs may vary greatly in their structural details even between closely related organisms, their general structure is one of the most conserved within Gram-negative bacteria (Steimle et al., 2016). They are complex glycolipids that normally contain three main regions: i) lipid A, ii) core oligosaccharide and iii) O-antigen polysaccharide (Whitfield and Trent, 2014). However, some bacteria synthesize these compounds only partially. Full-length LPSs are named "smooth" forms, whereas LPSs lacking O-antigen polysaccharide are called "rough" forms (Steimle et al., 2016). In rare cases, LPSs might comprise only lipid A (Wang et al., 2006). The general structure of the different forms of LPS is presented in Figure 4a.

1.4.2.1 Structure

Lipid A is the innermost domain that makes up the outer leaflet of the membrane. Although this molecule is the most conserved part of LPS, it still shows a high degree of structural diversity. Specifically, lipid A consists of a phosphorylated glucosamine disaccharide with attached acyl chains, and it may vary in the number and decoration of the phosphate groups, the number, length and modification of the acyl chains, and occasionally also the chemistry of the sugar backbone. Differences in the architecture of this molecule can be found not only between different species but sometimes even within one species (Steimle et al., 2016).

Linked to lipid A is core oligosaccharide (core OS). This region contains up to 15 monosaccharides, organized in a linear or branched structure, and it can be subdivided into an inner and outer core. Between the two of them, the inner core shows less variability, as it is usually conserved within a family or genus. This part of core OS is proximal to lipid A and normally starts with one or more Kdo residues. Further, L-glycero-D (or L)-manno-heptopyranose and other heptose units are typically present. The positioning of Kdo between sugar and lipid moieties is reminiscent of the Kdo linker used in the ABC transporter-dependent EPS biosynthesis pathway. However, while the Kdo linker is composed of β-linked Kdo residues, the Kdo residues in the inner core of LPS are joined by α- bonds. In addition, the inner core is often decorated with substituents such as phosphate or uronic acids. The less conserved is the outer core, also known as hexose region, as it is usually made of hexose units (Silipo and Molinaro, 2010; Whitfield and Trent, 2014; Sande et al., 2019).

Core OS provides an attachment site for O-antigen polysaccharide, also referred to as O-side chain or just O-antigen, which is the outermost region of LPS as well as its biggest fragment. This molecule is built

of repeating oligosaccharide subunits containing three to five sugars, and it can be made of as many as 40 repeating subunits. At least 20 different sugars can be a part of O-antigen, and some of them (e.g. abequose, colitose, paratose and tyvelose) are rarely found anywhere else in nature. In consequence, the composition of this large domain is the most variable amongst LPS components (Lerouge and Vanderleyden, 2002). As an example, different *Escherichia coli* strains are able to synthesize O-antigen with at least 188 different structures (DebRoy et al., 2016).

1.4.2.2 Biosynthesis

Although Gram-negative bacteria can greatly modify the structure of their LPSs, the basic biosynthesis pathway and export mechanism of these compounds are well-conserved. The overview of LPS biosynthesis pathway is depicted in Figure 4b. The LPS parts are assembled in the following order: lipid A, core OS and O-antigen. The biosynthesis of lipid A takes place first in the cytoplasm and then the cytoplasmic surface of inner membrane. This so-called Lpx pathway is mediated by a number of soluble cytoplasmic enzymes and peripheral membrane proteins also known as Lpx enzymes. In *E. coli*, the assembly of core OS ensues directly on lipid A. This process (Waa pathway) is catalyzed by several GTs associated with inner membrane (Waa proteins). Once the synthesis of lipid A-core OS is completed, the nascent molecule is flipped across the inner membrane by the ABC transporter MsbA (Voss and Trent, 2018). The formation of O-antigen starts independently at the cytoplasmic face of inner membrane and follows one of the three possible pathways similar to the EPS biosynthesis pathways. Regardless of the pathway, all O-antigens are built on a lipid carrier (undecaprenyl diphosphate) and ultimately exported to the periplasm. In the Wzx/Wzy-dependent pathway, the precursor subunits of O-antigen are assembled by specific GTs at the cytoplasmic membrane and translocated by the O-antigen flippase Wzx across the inner membrane, where they are polymerized into O-antigen by the O-antigen polymerase Wzy. ABC transporter-dependent O-antigens are synthesized by specific GTs entirely in the cytoplasm and transported across the membrane by the ABC transporter system composed of the O-antigen ABC transporter permease Wzm and the O-antigen ABC transporter ATP-binding protein Wzt (Greenfield and Whitfield, 2012; DebRoy et al., 2016). The synthase-dependent pathway requires the participation of a synthase that is believed to simultaneously polymerase and translocate the growing O-antigen across the membrane, but details of this process are not well-comprehended, as only the O-antigen of *Salmonella enterica* serovar Borreze is known to be formed using this pathway (Greenfield and Whitfield, 2012; Bohl and Aihara, 2018). On the periplasmic, O-antigen and lipid A-core OS are ligated by WaaL. The so obtained fully formed LPS molecule is extracted from the outer leaflet of inner membrane, transported across the periplasm, and finally inserted into the outer leaflet of outer membrane, all done

by a protein bridge comprising seven different proteins ($LptB_2FGCADE$) (Dong et al., 2017; Owens et al., 2019). Although some protein families involved in the biosynthesis of EPSs and LPSs are related, they can be distinguished based on the presence of characteristic domains (Pereira et al., 2013).

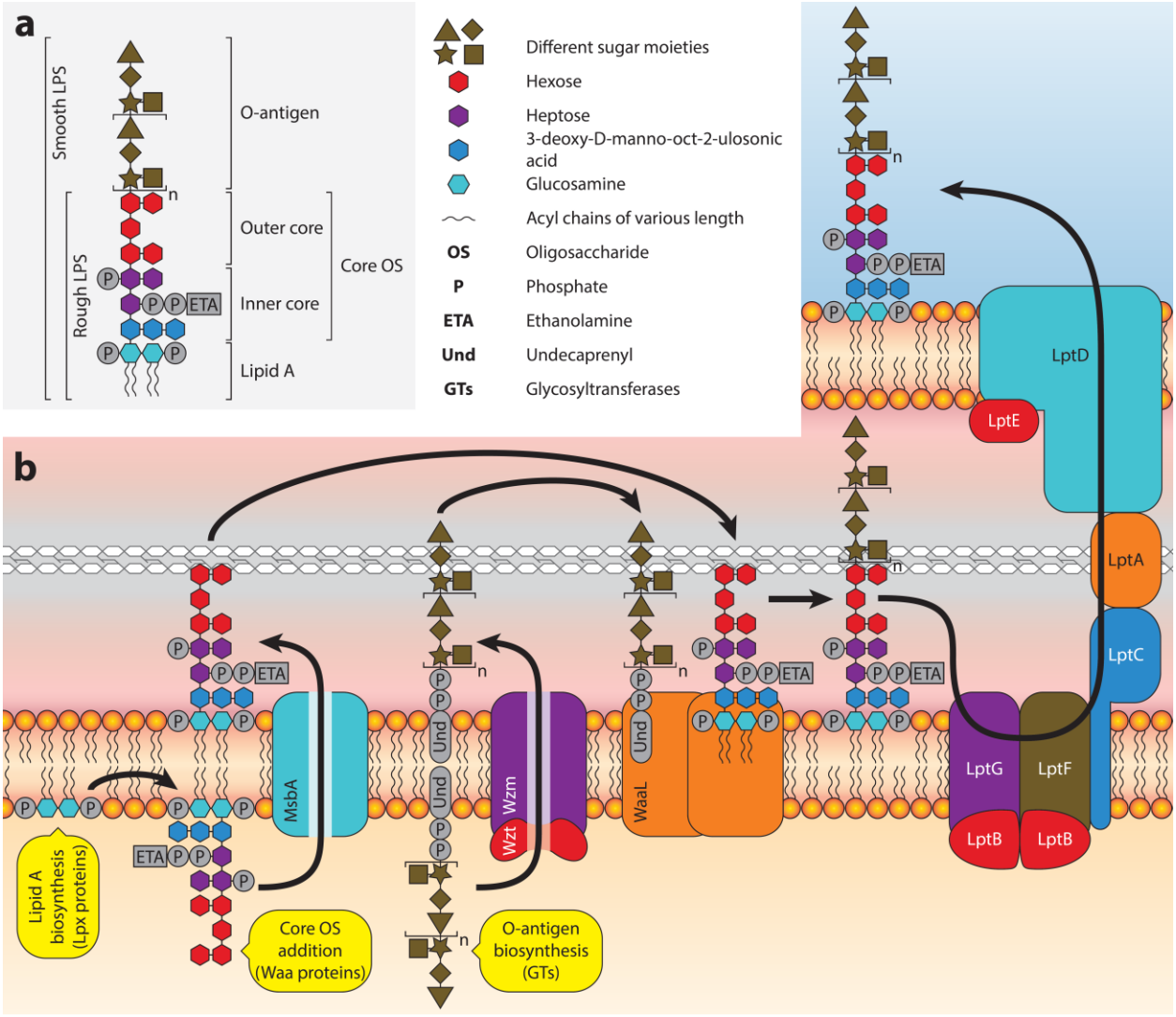


Figure 4. Overview of LPS structure (a) and biosynthesis pathway (b) (adapted from Owens et al., 2019, Pérez-Burgos et al., 2019, Steimle et al., 2016 and Whitfield and Trent, 2014).

1.5 The role of bacterial polysaccharides in specific environments

Although EPSs and LPSs play a crucial role in maintaining proper soil structure and preventing soil erosion, bacteria produce them primarily to accommodate their own physiological needs. In fact, EPSs greatly increase the ecological fitness of bacteria, and LPSs are even essential for the viability of some of them (Suresh Kumar et al., 2007; Zhang et al., 2013). Most functions attributed to both types of

compounds are of protective nature. Cells surrounded by a barrier made of EPSs are more resistant against predation, phagocytosis and environmental stresses (e.g. desiccation, UV radiation, extreme temperatures and elevated salt concentration) (Suresh Kumar et al., 2007; Kehr and Dittmann, 2015). Similar protection, although not as extensive, is provided by LPSs (Garmiri et al., 2008; Duncan et al., 2018). Bacterial polysaccharides also have the ability to bind and control the penetration of useful and harmful substances into the cell. This plays a role in protecting the cell from antibiotics or heavy metals as well as retaining trace elements under the conditions of their limited availability (Langley and Beveridge, 1999; Papo and Shai, 2005; Suresh Kumar et al., 2007). EPSs are especially effective as carbon reserves but also carbon sinks that help to regulate the carbon/nitrogen balance (Otero and Vincenzini, 2004; Costa et al., 2018). Furthermore, both types of polysaccharides are important for biofilm development, establishing symbiotic relationship with plants and infecting animal hosts (Kierek and Watnick, 2003; Suresh Kumar et al., 2007; Lindhout et al., 2009; Quelas et al., 2010; Matsuura, 2013; Kehr and Dittmann, 2015). Finally, the improvement of soil structure through the aggregating capabilities of bacterial polysaccharides is beneficial not only for the soil resistance to erosion but also for creating favorable hydrological niches for the bacteria themselves (Benard et al., 2019). The exact physiological as well as environmental roles of EPSs and LPSs are determined by the habitat in which the bacterium lives. The significance of adhesive polysaccharides in different environments will be explained in more detail on the example of biological soil crusts and agricultural soils.

1.5.1 Biological soil crusts

Bacterial polysaccharides hold especially high importance in the context of biological soil crusts, also known as biocrusts. The reason is that biocrusts are essential for the functioning of many terrestrial ecosystems (Sancho et al., 2014). They are assemblages of bacteria, archaea, fungi, algae, mosses and lichens, which form a coherent layer of living material intermingled with soil particles within the first millimeters of topsoil (Belnap and Lange, 2003). Organisms of biocrusts are embedded in a matrix of extracellular polymeric substances, amongst which dominate polysaccharides. The protective properties of these polymers make biocrusts highly stress-tolerant and allow them to live under extreme environmental conditions (Rossi et al., 2018). In fact, biocrusts can be found almost everywhere, from the Polar Regions (Williams et al., 2017) to the African Namib desert (Büdel et al., 2009). However, they are particularly prevalent in habitats with scarce vegetation, as light plays an important role in their development (Sancho et al., 2014). They are also the first colonizers of new ecosystems and after disturbances (Veste, 2005). In fact, biocrusts are considered to be “ecosystem engineers” because they enhance carbon and nitrogen pools, increase soil temperature and stability, and improve seedling

germination (Belnap and Lange, 2003; Williams et al., 2017). They can regulate moisture content in soil by increasing or reducing water penetrability as well (Rossi et al., 2018). Thanks to these abilities, biocrusts have high potential for improving soil resistance to erosion and are important drivers of soil development (Weber et al., 2016). They owe these properties largely to polysaccharide production.

At the beginning, biocrusts are formed mainly by cyanobacteria and other photosynthetic as well heterotrophic bacteria. Therefore, initial biocrusts have the form of biofilms (Mazor et al., 1996). During their development, capsular polysaccharides and LPSs play an especially important role in the attachment of free-living cells to a surface, whereas EPSs are essential for the consolidation of immobilized cells into a mature biofilm (Vogeleer et al., 2014). Already at this stage, biocrusts promote soil aggregation thanks to the adhesive properties of bacterial polysaccharides (Costa et al., 2018). In the later stages, microorganisms forming biocrusts are largely replaced by eukaryotes (Lan et al., 2012). As mentioned before, these organisms also possess mechanisms that drive aggregate formation. However, EPSs and LPSs produced by bacteria are still necessary to start biocrust development, initiate the aggregation of soil particles and prepare the conditions for the establishment of eukaryotic organisms (Mazor et al., 1996). Therefore, the polysaccharide-producing bacterial members of biocrusts should be considered as a seed and first hotspot for the overall processes of soil formation and stabilization. This also makes them especially valuable in initial ecosystems, which are characterized by poor soil structure, low nutrient content, few soil macroorganisms and little vegetation, as without their contribution these ecosystems would have little chance to develop (Fischer et al., 2010a).

1.5.2 Agricultural soils

Another environment where bacterial polysaccharides should have a high relevance are agricultural soils. This is because while good soil structure for crop growth requires the presence of stable aggregates, those aggregates are regularly destroyed by agricultural management practices (Weil and Brady, 2017). One of the most destructive techniques used in agriculture is intensive tillage. In general, tillage is a physical loosening of soil in preparation for growing crops. However, depending on its intensity, different types of tillage can be distinguished (Finch et al., 2014). The most intensive is conventional tillage (CT), which refers to operations that turn over the upper soil layer to bury the crop residues. This approach exposes the biologically less active subsoil, while the biologically most active topsoil is moved down the soil profile (Townsend et al., 2016). Less invasive is conservation tillage, which adopts the idea that the unique biocenoses of different soil layers should not be disturbed. Conservation tillage was first introduced in the United States in the form of no-tillage, also known as zero-tillage or direct drilling (Mäder and Berner, 2012). In this technique, seeds are drilled directly into

the stubble that remains after harvesting the previous crop, and the level of soil disturbance is minimal (Townsend et al., 2016). However, no-tillage is better suited for warm and dry climates, and it requires the usage of herbicides to control weed infestation, which is not allowed in organic farming. As organic farming has been gaining a lot of popularity in Europe over the past years, and the dominant climate there is humid temperate, the conservation tillage technique most commonly used by European farmers is reduced tillage (RT). It involves shallow working depth without soil inversion and leaves crop residues on the fields (Mäder and Berner, 2012). The visual difference between fields after CT and RT can be observed in Figure 5.

Figure 5. Fields at the agricultural field trial in Frick (Switzerland) after CT (on the right) and RT (on the left). CT turned over the upper soil layer and buried the crop residues, whereas RT did not cause soil inversion and left the crop residues on the field.

Although RT gained recognition relatively recently, it is increasingly promoted over CT to protect the aggregated soil structure and prevent soil erosion (Mäder and Berner, 2012). In fact, several studies already showed that RT is superior to CT in respect to soil aggregate preservation (Jacobs et al., 2009; Mikha et al., 2013; Bartlova et al., 2015; Sheehy et al., 2015; Singh et al., 2016). However, our understanding of the underlying mechanisms of this phenomenon is still lacking. Moreover, some reports suggest that in certain conditions tillage intensity might have no influence on soil aggregation

(Asgari, 2014). While abiotic factors influencing soil aggregation such as soil texture have been investigated and recognized as possible determinants of soil response to different tillage intensities (Cooper et al., 2016), less attention has been focused on biotic factors such as bacteria or fungi, even though their importance for aggregate formation is well-known. Furthermore, no previous studies evaluated if soil parameters can change how tillage influences the bacterial and fungal ability to drive soil aggregation. A few independent reports suggest that intensive tillage disrupts hyphal networks and, thus, hinders the fungal soil aggregation capabilities (Beare et al., 1997; Wright et al., 1999; Cookson et al., 2008; Dai et al., 2015; Lu et al., 2018), but their number is insufficient to draw any definite conclusions on the interactions between fungi and soil parameters under tillage stress, and no relevant meta-study was performed. Information on bacteria responsible for soil aggregation are even scarcer, as all related studies were performed only on isolated strains (Costa et al., 2018). There seems to be a relation between tillage intensity and soil-aggregating bacteria, as Caesar-TonThat et al. (2007) found higher diversity of bacterial isolates able to stabilize artificial aggregates in no-tilled soil compared with conventionally tilled soil. Moreover, Caesar-TonThat et al. (2014) isolated higher proportion of bacteria with soil aggregation capabilities from soil under no-tillage than from soil under CT. This could be caused by the fact that EPS and LPS production requires high inputs of energy and carbon, while reduced soil organic carbon stocks have been frequently observed under CT compared with less intensive tillage (Gadermaier et al., 2012; Quintero and Comerford, 2013; Krauss et al., 2017). However, isolation approaches are strongly biased towards cultivable bacteria, and the aforementioned studies did not evaluate the mechanics of aggregate stabilization employed by the isolated strains. Therefore, no previous study investigated the impact of tillage on bacterial communities of adhesive polysaccharide producers, and further investigation needs to be aimed at closing this research gap.

1.6 Open questions, research aims and hypotheses

The research for this thesis was conducted as part of the project "Fertility Building Management Measures in Organic Cropping Systems" – FertilCrop, which aimed at developing efficient and sustainable agricultural management techniques that could be applied in organic farming to increase crop productivity. Especially vital for the sustainability of agro-ecosystems is the preservation of soil aggregates under the disruptive influence of tillage, and thus aggregate formation and stabilization was of great interest to the project. Our attention was directed particularly to the microbial production of compounds that act as "glue" for soil particles, as it is undoubtedly an important mechanism that drives soil aggregation. The most studied microbial "glue" is glomalin synthesized by the hyphae of mycorrhizal fungi belonging to *Glomeromycota* (Vlček and Pohanka, 2019). The extensive studies on this glycoprotein showed for example that its production can be influenced by different environmental factors such as various soil properties and climatic conditions (Hammer and Rillig, 2011; Wang et al., 2017) as well as by fungal species (Bedini et al., 2009). It was also demonstrated that aggregate stability is correlated with the concentration of glomalin in soil and that tillage can negatively impact soil aggregation by reducing glomalin production (Wright et al., 1999; Borie et al., 2006). However, glomalin contributes mainly to the formation and stabilization of macroaggregates, whereas both macro- and microaggregates are stabilized by bacterial polysaccharides (Lehmann et al., 2017b). Nevertheless, while a lot is known already about glomalin in the context of soil aggregation, the corresponding data on bacterial polysaccharides is still scarce. Therefore, we decided to contribute towards the overall goals of the FertilCrop project by deepening the knowledge on bacterial polysaccharides.

EPSs and LPSs are very diverse groups of compounds. In fact, even closely related bacteria can produce structurally different polysaccharides. Moreover, the properties of polysaccharides strongly depend on their structure (Berne et al., 2015). Therefore, EPSs and LPSs produced by different bacteria may vary greatly in their adhesiveness (Suresh Kumar et al., 2007). This could be crucial for soil aggregation in different environments and under differing conditions, as bacterial community composition can be influenced by many factors such as pH, nutrient content, soil texture, temperature or plant species (Fierer, 2017). Hence, different bacterial communities might differ in their soil aggregation efficiency. Surprisingly, many studies evaluated the effects of isolated bacterial strains on soil aggregation (de Caire et al., 1997; Caesar-TonThat et al., 2007; HuiXia et al., 2007; Caesar-TonThat et al., 2014; Colica et al., 2014; Kheirfam et al., 2017a; Mugnai et al., 2018), but less attention was put on the community dynamics of bacterial polysaccharide producers under natural conditions. Because of that, until recently little was known on: i) what part of bacterial community has the potential to produce EPSs and LPSs, ii) which are the potential key producers of these polysaccharides, iii) if the overall potential to produce EPSs and LPSs is shaped by the community composition or is conserved within a community even if the

community changes, and iv) if the answers to the above questions are universal or depend on different factors, in which case what are the factors and how do they influence these answers. To address these research questions, several studies were designed as part of this thesis to evaluate communities of bacteria with the potential to produce EPSs and LPSs in different environments.

We were especially interested in the bacterial potential to produce adhesive polysaccharides in agricultural soils, as the preservation of soil aggregates under different tillage systems is an urging issue that was addressed by the FertilCrop project, and very little is known about the effects of tillage intensity on the bacterial ability to influence soil aggregation through the formation of EPSs and LPSs. However, agricultural soils are complex systems, and thus we decided to investigate bacterial polysaccharide producers in biocrusts as well. That is because biocrusts are model organisms well-known for their ability to induce soil aggregation. Therefore, we expected to find a high number of potential producers of adhesive polysaccharides there. Moreover, biocrusts are found in many terrestrial ecosystems, including agricultural soils, and thus the same potential polysaccharide producers could possibly play a role in soil aggregation by biocrusts as well as in agricultural soils. Finally, investigating biocrusts is possible in controlled laboratory conditions, and conclusions drawn from such experiment could be helpful in explaining trends observed in more complex systems such as agricultural soils. Because of that, our study on potential polysaccharide producers in biocrusts took the form of a microcosm experiment (**P1**). Such format enabled us to follow the establishment of biocrusts on two different soil substrates that came from sites with different types of naturally occurring biocrusts. This design originated from our interest in observing the changes in the communities of potential EPS and LPS producers during the development of biocrusts as well as in comparing the potential polysaccharide producers in biocrusts composed of different bacterial communities.

Our investigation of potential producers of adhesive polysaccharides in agricultural soil was divided into two studies based on separate field sampling campaigns. In the first study (**P2**), we aimed to evaluate the communities of potential EPS and LPS producers under CT and RT, and we were especially interested in comparing the soil layers in and below the tillage horizon, as we expected that the biggest differences would be visible between the soil layers directly affected by tillage and those outside the reach of the disturbance caused by tillage. The second study (**P3**) was designed as a follow-up of the first one, as we wanted to delve further into the topic of the influence of CT and RT on the bacterial polysaccharide producers. However, the aim of this study was focused more on how environmental factors could change the way the different tillage intensities influence the potential producers of EPSs and LPSs. Therefore, while our first agricultural study (**P2**) encompassed one agricultural field trial sampled at different depths, the second (**P3**) was based on three agricultural field trials from which only the top soil

was sampled. While the tillage regimes tested at the three trials were comparable, the sites differed significantly in various other parameters. Aside from the comparable tillage regimes, the primary property based on the differences in which the three trials were selected was soil texture. That was because it was suggested that soil texture could be an important factor modifying how the intensity of tillage affects the microbial ability to aggregate soil (Babin et al., 2019). It is especially expected that the effects of tillage could be best studied in silty and sandy soils as opposed to clayey soils. Furthermore, while some investigations of the fungal aggregating capabilities in relation to tillage intensity were performed in soils with different textures (Beare et al., 1997; Wright et al., 1999; Cookson et al., 2008; Dai et al., 2015; Lu et al., 2018), corresponding experimental data for bacteria was missing.

In the frame of this work, the following hypotheses were tested in the above-mentioned experiments: i) the key players of EPS and LPS production differ in biocrusts composed of different bacterial communities, but the relative abundance of the genes involved in the formation of these compounds as well as of the potential polysaccharide producers generally increases in the biocrusts compared with the initial soil substrates regardless of the variability in the community composition of the potential polysaccharide producers (**P1**), ii) the relative abundance of the genes related to EPS and LPS synthesis as well as of the potential polysaccharide producers is lower under CT compared with RT, and the expected difference between the tillage intensities is more apparent in the tillage horizon rather than below it (**P2**), and iii) the relative abundance of the EPS and LPS genes and the community composition of the producers of these compounds differs under CT and RT especially in sandy and silty soils compared with clayey soils (**P3**). In addition to these experiment-specific hypotheses, we tested more general hypotheses that linked all our experiments: i) improved soil aggregation is connected with increased bacterial potential to produce EPSs and LPSs, ii) the community composition of potential polysaccharide producers is shaped by similar factors as the overall community composition, and iii) the potential of a bacterial community to produce EPSs and LPSs depends on the taxonomic composition of that community.

2 Materials and methods

This thesis focuses on the bacterial potential to produce EPSs and LPSs in specific environments: i) biocrusts from initial ecosystems and ii) tilled agricultural soils. Biocrusts were cultivated in a microcosm experiment (**P1**), and tilled soils were collected during two field sampling campaigns (**P1**, **P2**). Metagenomic analyses of genes specific for the formation of adhesive bacterial polysaccharides were complemented by measurements of parameters such as stable aggregate fraction (SAF), dissolved organic carbon (DOC) and nitrogen (DON), microbial biomass carbon (Cmic) and nitrogen (Nmic), soil organic carbon (SOC) and X-ray computed microtomography (XCMT).

2.1 Microcosm experiment (P1)

Soil substrates for the microcosm experiment were collected from two initial ecosystems: i) the artificial catchment Chicken Creek (51°36'18'' N, 14°15'58'' E) and ii) a moving sand dune near Lieberose (51°36'18'' N, 14°15'58'' E). The two sites are located approximately 37 km apart in the state of Brandenburg in eastern Germany. The soil there differs mainly in pH – it is slightly alkaline in Chicken Creek (~7.3) and rather acidic in Lieberose (~5.4). Moreover, the sites have different origins. The Chicken Creek catchment was constructed in 2005 in an opencast mine by dumping and contouring sand and loamy sand material originating from Pleistocene sediments. After the construction, no restoration was undertaken and the area was allowed to undergo natural succession. The mobile sand dune near Lieberose (composed of Pleistocene aeolian sand) is a result of extensive disturbances of the land surface by former military activities (until approximately 1992). Biocrusts occur naturally at both locations, but the communities of organisms forming them differ. At the Chicken Creek catchment, major members of biocrusts are cyanobacteria, while algae dominate biocrusts at the Lieberose sand dune.

Bulk soils from the two sites were passed through a 2 mm sieve, packed into plastic pots (10 cm x 10 cm x 10 cm) and compacted to the density of 1.6 g cm^{-3}. In total, the experiment comprised 18 pots (9 per site). The water content was adjusted to 50 % of the maximum water holding capacity and controlled on a weekly basis. The pots were incubated in a sun simulator with a 16:8 day:night cycle. The relative air humidity there was kept constant at 90-95 %, and the temperatures in the light and dark periods were 25 °C and 18 °C, respectively. After biocrusts developed successfully four months later, the pots were transferred to a greenhouse, where similar conditions of temperature, watering regime and night-day cycles were maintained.

Bulk soils without biocrust development were sampled at the beginning of the experiment (T0), whereas after 4 (T1) and 10 (T2) months of incubation, samples of biocrusts were taken from the top 2 mm. At each sampling time point, three independent pots per soil substrate were sampled and then discarded. One part of each sample was stored at -80 °C and used for DNA extraction, library preparation and sequencing, and the other was stored at 4 °C for biochemical analyses.

The biochemical analyses performed for all samples included DOC and DON measurements. Furthermore, bulk soil samples from T0 were used for the determination of pH, while water repellency was analyzed in biocrust samples from T2. Finally, biocrusts from T2 grown on soil substrate from Chicken Creek, as the only ones with sufficient thickness, were examined by means of an XCMT. The experimental design of the microcosm experiment is shown in Figure 6.

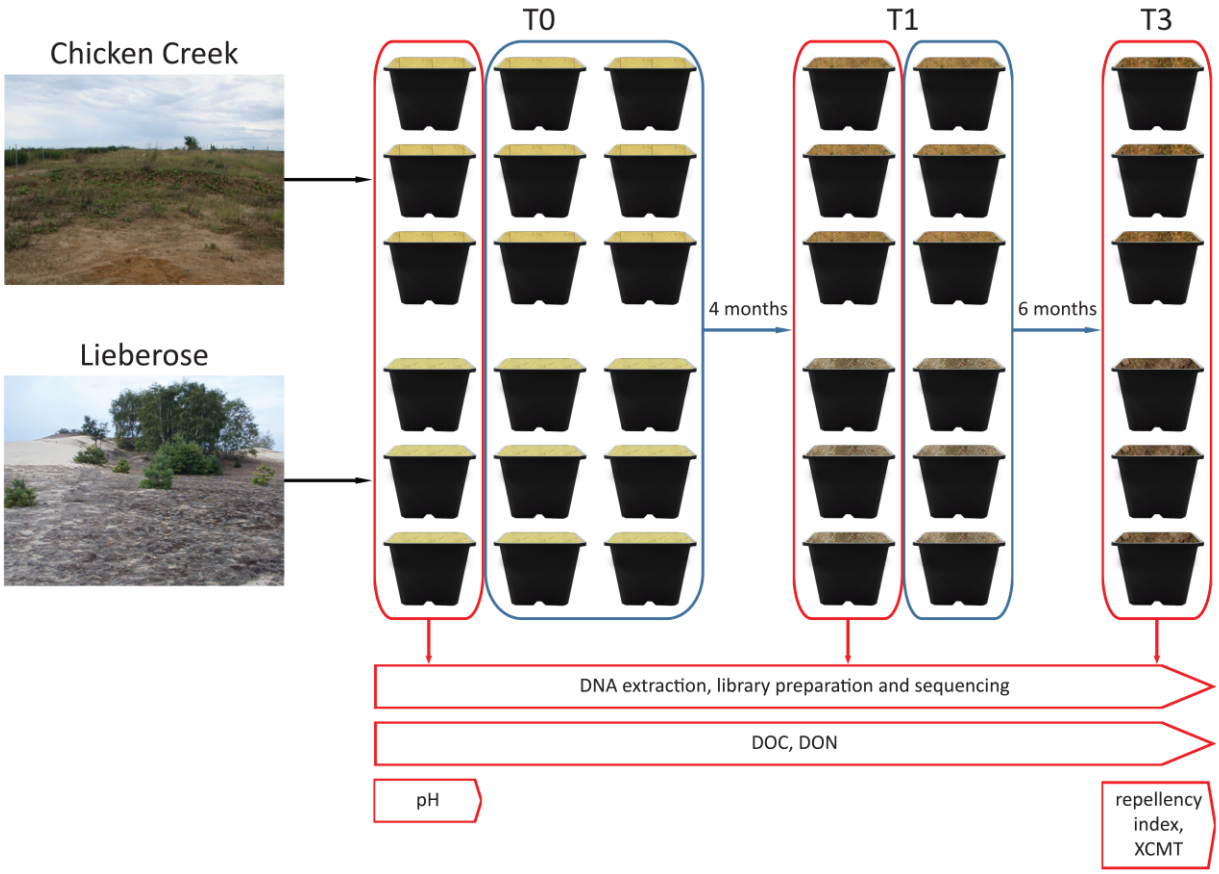


Figure 6. Experimental design of the microcosm experiment. Bulk soils were taken from two sites and put into pots. Samples of the bulk soils were taken at the beginning of the experiment at T0, whereas biocrusts were sampled after 4 and 10 months of incubation, at T1 and T2, respectively. Different analyses were performed on samples taken at selected time points.

2.2 Field sampling campaigns (P2, P3)

Samples of tilled agricultural soils were taken from the long-term organic field trials located in: i) Frick, Switzerland (47˚30'N, 8˚01'E, 350 m a.s.l.), ii) Moškanjci, Slovenia (46˚03'N, 15˚04'E, 225 m a.s.l.), and iii) Juchowo, Poland (53˚40'N, 16˚30'E, 160 m a.s.l.). The major distinction between the three sites is soil texture, but the sites differ also in other parameters (Table 1).

Table 1. Trial characteristics.

Trial	Frick	Moškanjci	Juchowo
Trial start	2002	1999	2010
Geographic coordinates	47˚30'N, 8˚01'E	46˚03'N, 15˚04'E	53˚40'N, 16˚30'E
Elevation [m a.s.l.]	350	225	160
Soil type	Stagnic Eutric Cambisol	Skeletic Eutric Cambisol	Haplic Arenosol
Soil texture	clayey	loamy	sandy
Climate type	temperate	continental	continental
Mean annual temperature [°C]	8.9	10.6	8.5
Mean annual precipitation [mm]	1000	913	750
Plant species in the last 2 years before the 2nd sampling campaign	spelt-grass & clover	winter rye-cover crops	spelt-lupine

The samples of two tillage treatments – CT and RT – were taken in spring, before tilling, from three replicated plots per treatment. The intensities of the CT and RT treatments at all three sites are comparable, even though the tillage treatments there are performed with slightly different equipment. At the Frick trial, CT is based on ploughing with a moldboard plough operating at 15-18 cm depth, while for RT, soil loosening is performed at a depth of 5-10 cm with a chisel and a skim plough, with occasional non-inversion loosening to 15-20 cm. In both systems, the seedbed preparation is done using a rotary harrow running at 5 cm depth. At the Moškanjci trial, for CT, a moldboard plough operating at 20 cm depth is used, followed by soil bed preparation with a rotary hoe. For RT, a special machine – 4-row disc harrow with individually suspended discs and a system for varying the working angle (and thus the tilling intensity) – is applied up to the depth of 10 cm in one or two passes to till the soil and prepare the seedbed. At the Juchowo trial, CT is performed by ploughing up to 30 cm deep with an Ecomat plough, while for RT, soil loosening up to 10 cm deep is done using a cultivator with goosefeet sweeps.

The first sampling campaign (**P2**) was carried out in 2015 only in Frick. At the time of sampling, a green manure mixture was growing on the plots. Samples were taken to a soil depth of 50 cm and divided into

three layers: 0-10 cm, 10-20 cm and 20-50 cm, producing 18 samples (3 depths x 2 tillage treatments x 3 plot replicates). During the second sampling campaign (**P3**), which took place in 2016 (in Juchowo) and 2018 (in Frick and Moškanjci), soil was taken to a depth of 10 cm from all three trials, resulting in 18 samples (3 trials x 2 tillage treatments x 3 replicated plots). At that time, grass-clover was growing in Frick, a mixture of cover crops – in Moškanjci, and lupine – in Juchowo. Both sampling campaigns were accomplished by taking 10 cores per plot using soil augers and homogenizing cores from the same plot. One part of each homogenized sample was stored at -20 °C before using for DNA extraction, library preparation and sequencing, and the other was stored at 4 °C before physicochemical measurements.

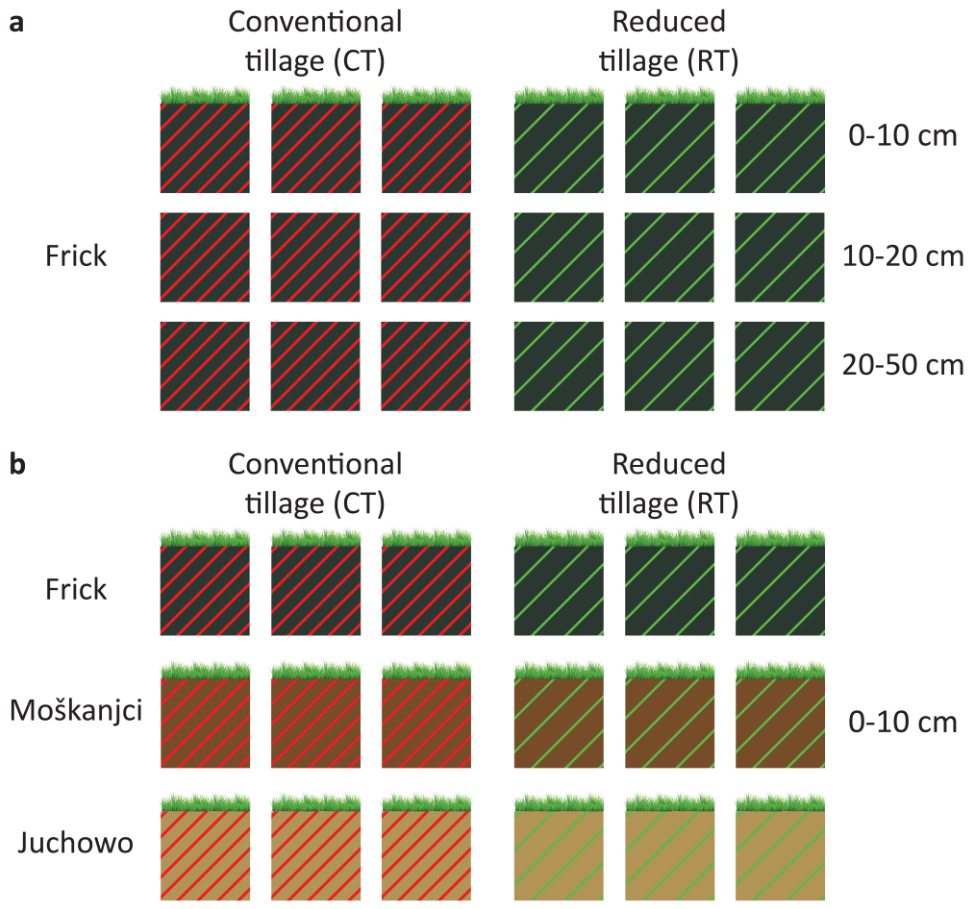


Figure 7. Experimental design of the two field sampling campaigns: a) the first field sampling campaign and b) the second field sampling campaign. During both field sampling campaigns, CT and RT were sampled. During the first field sampling campaign, three depths were sampled at one site, whereas during the second field sampling campaign, one depth was sampled at three sites.

The parameters analyzed in samples from the first sampling campaign were SAF, SOC, DOC and Cmic, while samples from the second sampling campaign were measured for clay, silt and sand content, SAF, pH, DOC, Cmic and Nmic. The experimental designs of the two field sampling campaigns are shown in Figure 7.

2.3 Physicochemical measurements

The physicochemical measurements described in this thesis were performed by cooperation partners. The determination of clay, silt and sand content was accomplished by means of a combined sieving and sedimentation method (ISO 11277 2009). SAF was measured using a wet sieving approach (Murer et al., 1993). DOC, DON, Cmic and Nmic were quantified by means of a chloroform fumigation-extraction method (Brookes et al., 1985; Vance et al., 1987). DOC and DON were determined in unfumigated samples, whereas Cmic and Nmic were estimated as a difference between fumigated and unfumigated samples (Joergensen, 1996; Joergensen and Mueller, 1996). The assessment of SOC was performed by wet oxidation (Krauss et al., 2017). Water repellency was evaluated by means of an ethanol/water microinfiltrometric sorptivity procedure (Fischer et al., 2010b). XCMT was employed to analyze pore connectivity and calculate Euler characteristics according to Köhne et al. (2011) and Vogel et al. (2010). The measurement of pH was achieved in $CaCl_2$ solution for samples from the microcosm experiment, and in demineralized water for the tilled soil samples (ISO 10390 2005).

2.4 DNA extraction, library preparation and sequencing

DNA extraction of samples from the microcosm experiment (**P1**) was carried out by means of the "Genomic DNA from soil" NucleoSpin Soil Kit (Macherey-Nagel, DE) following the producer's guideline. For these samples, lysis was accomplished using Buffer SL1. DNA from the tilled soil samples (**P2**, **P3**) was extracted according to the phenol-chloroform based DNA/RNA coextraction protocol described by Lueders et al. (2004). During this procedure, a Precellys24 homogenizer with CKMix tubes (Bertin Technologies, France) were used for sample lysis. Extracted DNA was checked for purity by means of a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, USA). The quantity was also verified using a SpectraMax Gemini EM microplate reader (Molecular Devices, USA) with a Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies, USA) (**P1**, **P2**), or a Qubit 4 Fluorometer with a Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, USA) (**P3**).

After extraction, DNA was sheared by employing an E220 Focused-ultrasonicator (Covaris, USA). Sheared DNA was used to construct metagenomic libraries by means of a NEBNext Ultra- (**P1**, **P2**) or NEBNext

Ultra II DNA Library Prep Kit for Illumina (**P3**), and NEBNext Multiplex Oligos for Illumina (New England Biolabs, UK), as described in the protocol of the manufacturer. Where applicable, Agencourt AMPure XP beads (Beckman Coulter, USA) were employed. Library size and concentration were assessed by means of a 2100 Bioanalyzer with a High Sensitivity DNA Analysis Kit (Agilent, USA) (**P1**, **P2**), or a Fragment Analyzer with a DNF-473 Standard Sensitivity NGS Fragment Analysis Kit (Advanced Analytical, USA) (**P3**).

After pooling equimolarily to 4 nM and spiking with PhiX, libraries were sequenced on a MiSeq sequencer using a MiSeq Reagent Kit v3 for 600 cycles (Illumina, USA). Raw sequencing data was uploaded to the Sequence Read Archive (SRA) under the accessions numbers PRJNA509545 (**P1**), PRJNA387672 (**P2**) and PRJNA555481 (**P3**), respectively. The processing of the samples is shown in Figure 8a.

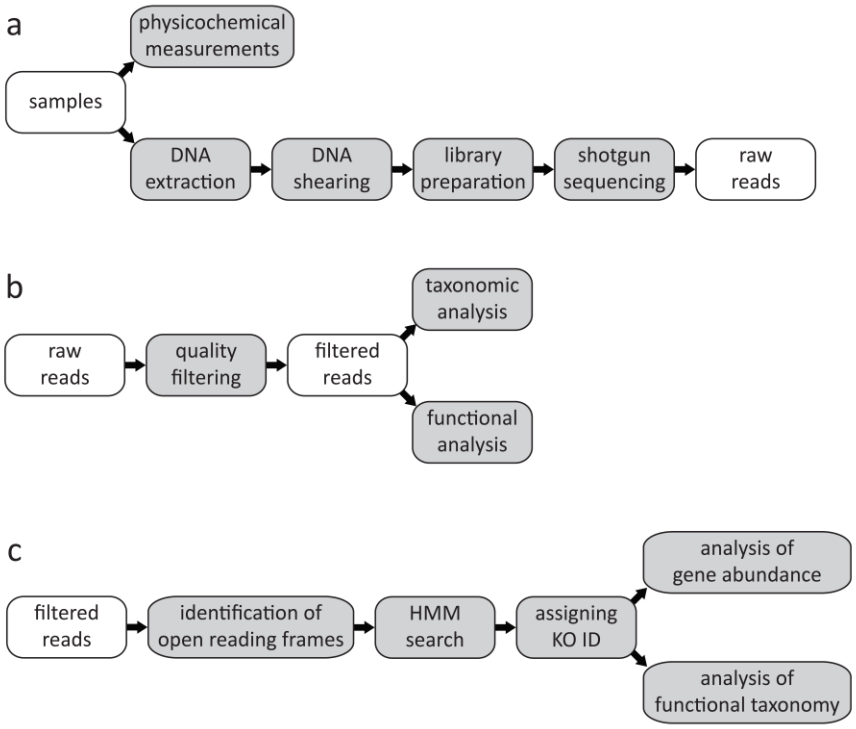


Figure 8. Schemes demonstrating the workflows of a) sample processing, b) general bioinformatical processing of sequencing data and c) functional analysis of genes specific for EPS and LPS biosynthesis.

2.5 Filtering and taxonomic analysis of sequencing data

Raw sequencing data was filtered by removing remnant adaptor sequences, trimming terminal nucleotides with Phred quality scores lower than 15 and discarding reads shorter than 50 bp. This was

accomplished by means of AdapterRemoval (Schubert et al., 2016). Additionally, reads with more than 1 % ambiguous bases (N) were eliminated from the microcosm sequencing data (**P1**) using PRINSEQ-lite (Schmieder and Edwards, 2011b). PhiX decontamination was performed by means of DeconSeq (Schmieder and Edwards, 2011a).

Filtered reads were taxonomically profiled against the National Center for Biotechnology Information's non-redundant (NCBI-nr) database (Sayers et al., 2019). This was achieved using Diamond (Buchfink et al., 2015) together with MEtaGenome ANalyzer (MEGAN) (Huson et al., 2011) (**P2**) or Kaiju (Menzel et al., 2016) (**P1, P3**). The results of the taxonomic profiling against the NCBI-nr database were confirmed by the SILVA's database (Glöckner et al., 2017) assignment of the 16S rRNA gene sequences identified by means of SortMeRNA (Kopylova et al., 2012). The general bioinformatic workflow is shown in Figure 8b.

2.6 Functional analysis of sequencing data

Genes specific for EPS and LPS biosynthesis were identified using a targeted pipeline combining Hidden Markov Model (HMM) searches with blasts. To create this pipeline, protein sequences associated with the production of adhesive polysaccharides were obtained from the online Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa and Goto, 2000) and examined for the presence of function-specific conserved domains by means of CD-search (Marchler-Bauer et al., 2015). KEGG Orthology (KO) entries that possessed such domains, as well as HMMs of the corresponding domains downloaded from the Pfam (Finn et al., 2016) and TIGRFAMs (Haft et al., 2013) databases were used to construct specific protein and HMM databases. Separate databases were created for sequences related to EPS and LPS formation, and their identification was performed independently. The pipeline was established using the sequencing data from the first field sampling campaign (**P2**). FragGeneScan (Rho et al., 2010) was applied on filtered reads to predict open-reading frames, which were then scanned with HMMER (Mistry et al., 2013). Reads matching the downloaded HMMs were subsequently blasted against the self-built KO databases using Diamond. KO numbers were assigned only to the reads with the top 25 blast results matching. The specificity of this pipeline was validated by using blastx on 25 randomly selected reads per KO number against the online NCBI-nr database. If all the 25 reads were assigned to the function of interest, the results were considered sufficiently specific. Otherwise, the KO entry and the corresponding HMM were removed from the pipeline. Out of 81 investigated KO numbers (67 for EPSs and 14 for LPSs), 14 were included in the final version of the pipeline. The final list of KO numbers and HMMs used for the analysis, can be found in Table 2. The workflow of the bioinformatical analysis used for the genes specific for EPS and LPS production is shown in Figure 8c. Additionally, COG (Clusters of Orthologous Groups) functional categories were assigned to the sequencing data from the

microcosm experiment (**P1**) based on the eggNOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups) database (Huerta-Cepas et al., 2015) by means of Diamond.

Table 2. KO numbers and HMM IDs used for the functional analysis with corresponding genes and proteins.

KO number	HMM ID	Gene	Protein
K01991	PF02563	*wza*	polysaccharide export outer membrane protein Wza
K03819	TIGR04016	*wcaB*	colanic acid biosynthesis acetyltransferase WcaB
K03818	TIGR04008	*wcaF*	colanic acid biosynthesis acetyltransferase WcaF
K16710	TIGR04006	*wcaK/amsJ*	colanic acid/amylovoran biosynthesis pyruvyl transferase WcaK/AmsJ
K10107	TIGR01010	*kpsE*	capsular polysaccharide export system permease KpsE
K19420	TIGR01006	*epsA*	exopolysaccharide biosynthesis tyrosine kinase modulator EpsA
K19419	PF14897	*epsG*	exopolysaccharide biosynthesis transmembrane protein EpsG
K16081	PF13372	*algE*	alginate export outer membrane protein AlgE
K19295	PF16822	*algJ*	alginate biosynthesis acetyltransferase AlgJ
K00692	PF02435	*sacB*	levansucrase SacB
K09691	PF14524	*wzt*	O-antigen ABC transporter ATP-binding protein Wzt
K11719	TIGR04409, PF06835	*lptC*	$LptB_2FGCADE$ lipopolysaccharide export complex inner membrane protein LptC
K07091	TIGR04407	*lptF*	$LptB_2FGCADE$ lipopolysaccharide export complex permease LptF
K11720	TIGR04408, PF03739	*lptG*	$LptB_2FGCADE$ lipopolysaccharide export complex permease LptG

2.7 Statistical analysis and data visualization

Prior to statistical analysis and data visualization, sequencing data was preprocessed into relative abundances. This was done by dividing the number of reads assigned to a gene or organisms by the total number of reads per sample. In case of the sequencing data from the microcosm experiment and the second field sampling campaign (**P1**, **P3**), only bacterial reads were analyzed. For the sequencing data from the second field sampling campaign (**P3**), absolute gene abundances were also estimated by multiplying the relative gene abundance with the Cmic value and dividing by 100 (Zhang et al., 2017).

Most of the following statistical analyses and data visualizations were performed using R (R Core Team, 2016).

The statistical tests were chosen based on the experimental design of the different experiments and the research questions that these experiments were meant to answer, and thus different tests were employed for different datasets. As the first field sampling campaign (**P2**) had a repeated-measures design, the data from this experiment was statistically analyzed by means of multilevel models. This was accomplished by employing the lme function from the nlme package (Pinheiro et al., 2018) and setting the following contrasts to identify differences between sampling depths: 0-20 cm vs 20-50 cm, and 0-10 cm vs 10-20 cm. The datasets from the microcosm experiment (**P1**) and the second field sampling campaign (**P3**) both included only two independent variables and, thus, were analyzed using a robust 2-way independent analysis of variance (ANOVA). This was achieved by means of the pbad2way function in the WRS package (Wilcox and Schönbrodt, 2014). Additionally, the pb2gen function from the same package was employed to perform a robust t-test for the data from the second field sampling campaign (**P3**) in order to analyze one of the variables in more detail. Both robust tests were based on the median as M-estimator and used 2000 bootstrap samples. To control the false discovery rate in the sequencing data, the Benjamini-Hochberg procedure was applied. For the data from the microcosm experiment (**P1**) and the second field sampling campaign (**P3**), effect sizes were calculated according to Field et al. (2012) in the form of omega squared (ω^2) and Pearson's correlation coefficient r where applicable.

Dissimilarity between samples was examined by means of principal component analysis (PCA) for soil data (**P3**), and non-metric multidimensional scaling (NMDS) (**P2**) or principal coordinates analysis (PCoA) (**P1**, **P3**) for sequencing data. The generation of PCA ordination was performed with the rda function of the vegan package (Oksanen et al., 2018). NMDS and PCoA ordinations were based on Bray-Curtis distances, and they were created by employing the metaMDS function from the vegan package or the pcoa function in the ape package (Paradis et al., 2004), respectively. Bray-Curtis dissimilarity matrices for PCoA were calculated using the vegdist function of the vegan package. Where necessary, negative eigenvalues were corrected by means of the Caillez procedure.

Within-sample diversity was measured as the nonpareil diversity index estimated using Nonpareil (Rodriguez-R and Konstantinidis, 2014) (**P1**) or the Shannon-Wiener index determined by means of the alpha.div function from the asbio package (Aho, 2015) (**P2**). Correlations within the data were explored using the cor.test function to calculate the Spearman's rank correlation coefficient ρ (**P1**). VENN diagrams were created with InteractiVenn (Heberle et al., 2015). The graphs were made by means of the functions ggplot, heatmap.2 and ordiplot in the packages ggplot2 (Wickham, 2009), gplots (Warnes et al., 2015) and vegan, respectively.

3 Discussion

3.1 Tools to analyze bacterial polysaccharides and their producers in soil

Adhesive bacterial polysaccharides, specifically EPSs and LPSs, are important drivers of soil aggregation. In turn, soil aggregates play a crucial role in preventing soil erosion. It is, therefore, essential to investigate the production of these polysaccharides in environments where soil is particularly prone to erosion. Especially high rates of erosion are observed in initial ecosystems, which are characterized by poorly established vegetation and low soil aggregation, and in agroecosystems, in which soil aggregates are regularly disturbed by agricultural management practices (Borrelli et al., 2017). The presence of bacteria with soil aggregation capabilities in biocrusts from initial ecosystems and in agricultural soils was proven before by isolation studies, but data on the whole communities of polysaccharide-producing bacteria in their natural environment was missing (Caesar-TonThat et al., 2007; HuiXia et al., 2007; Caesar-TonThat et al., 2014; Colica et al., 2014; Kheirfam et al., 2017a; Mugnai et al., 2018). For these reasons, the central aim of this thesis was to address the bacterial production of adhesive polysaccharides in biocrusts and agricultural soils. However, the measurement of polysaccharide content in these media is difficult.

While many methods could be used to directly determine the concentration of EPSs and LPSs in aqueous solutions, the structurally and chemically complex matrices of media such as soil or biocrust require for the polysaccharides to be extracted from these media before any measurements can be performed. Until now, the efforts to investigate LPSs were made mostly in the area of human health. Thus, the existing LPS extraction techniques were designed to obtain LPS molecules mainly from clinical samples, food products or pure cultures, for such purposes as serotyping, structure studies and epitope mapping. As a result, the available LPS extraction methodologies are not optimized for accurate measurements of LPS content in soil or biocrust, and more research needs to be done before these techniques can be reliably used for this purpose (Parker et al., 1982; Ford et al., 1985; Stromberg et al., 2017).

Conversely, more efforts were made to develop methods to extract EPSs from soil or biocrust. In this case, the main difficulty is to extract both free and bound extracellular polysaccharides from a given medium without damaging any cells and causing the leakage of intracellular polysaccharides. Recently, Redmile-Gordon et al. (2014) evaluated the suitability of different EPS extraction methods for soil but found all of them more or less biased. Similarly, Rossi et al. (2018) discussed the existing protocols of EPS extraction from biocrusts stressing that so far no universal methodology was established and further optimization of the available ones is required. Additionally, the existing methods do not allow for the distinction of bacterial EPSs from those of other origins (fungal, plant, etc.) that can be found in soil and

biocrust. This limited their usefulness for the purpose of the current thesis, whose main focus was on bacteria.

Another drawback of directly measuring the concentrations of EPSs and LPSs is that this approach provides only a snapshot of different microbial processes happening at a given moment. Specifically, polysaccharide levels in soil or biocrust depend not only on the activity of EPS and LPS producers but also of polysaccharide degraders (Colica et al., 2015). Moreover, the bacterial production of adhesive polysaccharides can be influenced by factors such as nutrient levels, aeration rates or temperature, and thus is highly dynamic (Suresh Kumar et al., 2007). At the same time, soil structure formation is a long-term process, which has been shown to be hardly affected by the dynamic changes in microbial activity (Redmile-Gordon et al., 2020). Therefore, gaining a broader perspective on the bacterial potential to shape the structure of soil via the production of adhesive polysaccharides requires measuring parameters that are more stable than EPS and LPS content.

Aside from directly measuring the content of EPSs and LPSs, it is also possible to indirectly estimate the production rates of these compounds by evaluating the results of their activity. As adhesive polysaccharides drive soil aggregation, the determination of stable aggregate fraction in soil is one of the available indirect methods (Murer et al., 1993). We applied it to demonstrate for example that the effects of different tillage types on aggregate stability depend on site-specific conditions (**P3**). Another useful technique is CT, which allows to observe the three-dimensional structure of soil (Vogel et al., 2010; Köhne et al., 2011). Unfortunately, this method is most appropriate for visualizing strong differences in soil aggregation, and thus its applications are limited. We used CT to prove that the biocrusts grown in our microcosm experiment had the ability to aggregate soil (**P1**). These approaches avoid the parameter stability issue of the direct measurements of adhesive polysaccharide content by targeting properties that change over a longer time period. However, they inform only about the overall soil aggregation without distinguishing between its drivers, and thus suffer from the same specificity limitation as the direct measurement methods of EPS content.

An indirect approach of assessing the bacterial production of adhesive polysaccharides in environmental samples that solves both the stability and specificity issues is metagenomics. In particular, shotgun sequencing (SGS) of total DNA is a powerful tool that not only provides insights into microbial community structure and functionality but also enables the linkage of specific functions to particular taxa. Therefore, SGS allows to investigate functions performed only by a specified group of organisms, as in the case of the current thesis – by bacteria. While this approach does not provide information on the actual activity of bacteria in a given environment, it can be used to identify and quantify bacterial genes and taxa that harbor them, and thus acquire information about the potential activity of bacterial

communities (Thomas et al., 2012). Moreover, gene levels are relatively stable within bacterial communities and, thus, allow for observing long-term changes in the bacterial potential to perform specific functions. Metagenomics alone generates relative abundance data, which explains the proportional importance of different taxa and genes within bacterial communities. Additionally, absolute abundances can be estimated by coupling this technique with microbial quantity measurements, such as microbial biomass carbon, in order to describe the potential significance of bacterial taxa and genes for the ecosystem functioning (Zhang et al., 2017). For these reasons, metagenomics is widely applied to better understand the functioning of bacterial communities and their role in different habitats (Nayfach and Pollard, 2016), but it was not used before to acquire knowledge on the bacterial production of adhesive polysaccharides.

However, the development of a novel analytical pipeline that targets a function not addressed previously in metagenomic studies faces the challenge of selecting the right marker genes. One obstacle is that adhesive bacterial polysaccharides, especially EPSs, are a very large and diverse group of compounds, which is reflected by the high number and diversity of the genes involved in their biosynthesis (Suresh Kumar et al., 2007; Schmid et al., 2015; Sengupta et al., 2018). Many of these genes are species or even strain specific and would require very high sequencing depth to be detected, if they were present in the sample at all, which disqualifies them as marker genes. Another obstacle is that many genes related to the EPS and LPS biosynthesis, in particular the ones responsible for the initial formation of precursor molecules, are involved not only in the production of EPSs and LPSs but also in the metabolism of intracellular sugars (Suresh Kumar et al., 2007; Pereira et al., 2015; Putker et al., 2015). For these reasons, we decided to analyze only those genes that were specific for the production of adhesive bacterial polysaccharides, most of which were responsible for the assembly and transport of our compounds of interest to the cell surface (in case of LPSs) or outside the cell (in case of EPSs) and were additionally conserved enough to be detectable at the applied sequencing depth.

3.2 Important pathways of adhesive polysaccharide biosynthesis

While we looked for the genes related to EPS and LPS production in samples as diverse as various agricultural soils, biocrusts as well as bulk soils from different initial ecosystems, we observed that the proportions between the investigated genes remained rather stable in all of the analyzed source materials (**P1**, **P2**, **P3**). Identified as the most or one of the most abundant amongst the investigated genes in all of the processed samples were *lptF* and *lptG* of the LPS synthesis pathway, which encode for the transmembrane proteins LptF and LptG of the $LptB_2FGCADE$ LPS export complex (Dong et al., 2017; Owens et al., 2019). The majority of bacterial reads from each experiment performed in the scope of

this thesis were assigned to phyla well-known for possessing LPSs, such as *Proteobacteria*, *Cyanobacteria*, *Acidobacteria*, and *Bacteroidetes* (Lagier et al., 2012). Furthermore, LPS producers were recently found in the *Fimicutes* phylum, whose members were previously considered as lacking LPSs (Antunes et al., 2016; Poppleton et al., 2017). As LptF and LptG are essential for LPS translocation to the outer membrane (Dong et al., 2017) and additionally highly conserved (Ruiz et al., 2008; Putker et al., 2015), the relatively high abundance of the genes encoding them in our datasets was expected. In comparison, the other investigated component of the $LptB_2FGCADE$ complex, LptC, is less conserved and likely not even necessary for LPS biosynthesis. In fact, its supportive role in the LPS transport was recently suggested after isolating LPS-producing mutants that lacked this protein (Benedet et al., 2016). This explains why the corresponding gene, *lptC*, was detected in a very low number of copies across all of the analyzed samples.

Another gene with a high relative abundance in our datasets was *wza* of the Wzx/Wzy-dependent EPS synthesis pathway. This gene encodes for an outer membrane protein Wza, which acts as a channel transporting many different EPSs across the outer membrane in a wide variety of bacterial taxa (Pereira et al., 2013). Moreover, Wza contains a highly conserved polysaccharide export sequence domain (Cuthbertson et al., 2009). This explains the relatively high abundance of the *wza* gene in the metagenomic data. In fact, this gene showed comparable relative abundance to *lptF* and *lptG* in most of the analyzed samples except for the bulk soils from initial ecosystems (**P1**). There, the majority of the genes related to EPS synthesis had low abundance, especially compared with the biocrusts that developed later on those bulk soils in the microcosm experiment. The increased relative abundance of the genes involved in EPS production in the biocrusts compared with the bulk soils is in agreement with the observed accumulation of dissolved organic carbon and increasing soil aggregation, all of which point to an intensified formation of EPSs in our biocrusts. This result was expected, as the biocrusts cultivated in the microcosm experiment were formed mainly of bacterial biofilms, for the development of which EPSs are essential (Maunders and Welch, 2017). The only gene associated with EPS synthesis that had higher relative abundance in the bulk soils (reached an intermediate level) compared with the biocrusts was *kpsE* of the ABC-dependent pathway. This gene was low abundant not only in the biocrusts (**P1**) but in the investigated agricultural soils as well (**P2**, **P3**). The fact that it is related to the formation of capsular polysaccharides, which help bacteria to survive in extreme environments, explains its, on one hand, relatively intermediate abundance in the nutrient-depleted bulk soils from initial ecosystems (**P1**), and, on the other hand, relatively low abundance in nutrient-rich biocrusts (**P1**) and agricultural soils (**P2**, **P3**) (Rendueles et al., 2017).

Aside from *kpsE*, the other genes that reached intermediate relative abundance were *wcaB* and *wcaF* of the Wzx/Wzy-dependent EPS synthesis pathway, as well as *wzt* of the LPS synthesis pathway. The *wcaB* and *wcaF* genes encode for the acetyltransferases WcaB and WcaF, which are involved in the formation of colanic acid (Schmid et al., 2015). The production of this EPS, also known as M antigen, characterizes bacteria of the family *Enterobacteriaceae*, including many human pathogens, although there is evidence that colanic acid could be produced also by other *Gammaproteobacteria* (Schmid et al., 2015; López et al., 2017). This EPS protects bacteria from such stresses as osmotic shock, acidity, desiccation, phages, extreme temperatures, destabilization of the outer membrane or oxidative stress (Kim et al., 2015; Pando et al., 2017). Moreover, M antigen was show to be important, or even required, for the biofilm formation in *Escherichia* coli and *Salmonella* (Danese et al., 2000; Ledeboer and Jones, 2005; Pando et al., 2017; Zhang and Poh, 2018). As a result, this EPS and the genetic machinery responsible for its production were extensively studied (Roca et al., 2015). Therefore, it is not surprising that a relatively high number of reads in our datasets was annotated to *wcaB* and *wcaF*. At the same time, the lower relative abundance of these genes compared with *wza*, which also participates in the production of colanic acid (as this EPS is synthesized via the Wzx/Wzy-dependent pathway), is in line with the product specificity of WcaB and WcaF as opposed to the universality of Wza. Similarly, the relatively high abundance of *wzt*, which encodes for the O-antigen ABC transporter ATP-binding protein Wzt, is not surprising. This protein participates in the processing of O-antigen polysaccharide, which is an optional component of LPS (Whitfield and Trent, 2014; DebRoy et al., 2016). The lower relative abundance of *wzt* compared with *lptF* and *lptG* is, thus, in agreement with the non-essentiality of O-antigen for LPS maturation as well as the fact that Wzt is involved in just one of the two main O-antigen formation pathways (the ABC transporter-dependent pathway as opposed to the Wzx/Wzy-dependent pathway (Greenfield and Whitfield, 2012).

The genes with the lowest relative abundance in our metagenomic datasets were *algE*, *algJ*, *epsA*, *epsG*, *sacB*, and *wcaK*/*amsJ*. As EPS synthesis pathways are often poorly conserved and specific only for single species or strains (Skorupska et al., 2006; Pereira et al., 2013; Schmid et al., 2015), the low relative abundance of some genes participating even in the formation of widely distributed EPSs, like in the case of *wcaK*/*amsJ* involved in colanic acid/amylovoran production, can be expected. The small presence of the genes related to EPSs synthesized only by a small group of bacteria in a diverse community is also natural. Alginate, in the formation of which participate *algE* and *algJ*, is characteristic for selected members of the genera *Azotobacter* and *Pseudomonas* (Hay et al., 2014). These genera are composed of many plant growth-promoting endophytes and rhizobacteria (Preston, 2004; Kandel et al., 2017), which was likely the reason why they were poorly represented in samples of bacterial biocrusts (**P1**) and bulk soils (**P1**, **P2**, **P3**). This explains the low relative abundance of the genes responsible for alginate

biosynthesis in our datasets. Similarly, the low relative abundance of *sacB*, *epsA* and *epsG*, which are involved in the formation of levan and other EPSs produced by Gram-positive bacteria (Wu et al., 2014; Schmid et al., 2015), is in line with the low proportion of typical Gram-positive to Gram-negative phyla in our samples (**P1**, **P2**, **P3**). However, even lowly abundant genes can play an important role, as the yield of a gene product depends not only on the abundance of the gene but also on its expression level (Abram, 2015).

3.3 Relation between the bacterial potential to produce EPSs and LPSs and soil aggregation

Aside from maintaining rather stable proportions between each other, the genes related to the biosynthesis of EPSs and LPSs displayed also remarkable stability in their relative abundance. As mentioned before, we found differences in the relative abundance of these genes mostly between initial soils and biocrusts grown on them in the microcosm experiment (**P1**). There, the increased relative abundance of the EPS and LPS genes in the biocrusts compared with the bulk soils was accompanied by improved soil aggregation, which suggests that soil aggregation could be influenced by the relative abundance of the genes involved in EPS and LPS production. However, little differences in the relative abundance of these genes were observed in the agricultural experiment comprising three field trials (**P3**), where the examined sites differed significantly in aggregate stability. This indicates that the relative abundance of the genes related to the biosynthesis of EPS and LPS is not the sole parameter shaping soil aggregation. Such result was expected, as it is known that the bacterial production of adhesive polysaccharides is not the only mechanism driving soil aggregation, and it does not contradict the hypothesis that the relative abundance of the EPS and LPS genes is important for soil aggregation. Yet the results of our experiments imply that another parameter related to the bacterial production of adhesive polysaccharides could be as or even more essential, namely the absolute abundance of the genes involved in EPS and LPS biosynthesis.

In the microcosm experiment, the aggregation of soil particles by bacterial biofilms was observed for all of our biocrusts. However, half of the biocrusts were thicker and more coherent than the rest. The thicker biocrusts were composed mostly of *Cyanobacteria*, whereas *Acidobacteria* and *Chloroflexi* were characteristic for the thinner biocrusts. *Cyanobacteria* are well-recognized for their ability to produce large amounts of polysaccharides, which they use e.g. for gliding movement and protection from extreme environmental conditions, and they synthesize LPSs as well (Rossi and De Philippis, 2015). They have been also repeatedly shown to improve aggregate stability and prevent soil erosion (de Caire et al., 1997; Kheirfam et al., 2017a; Kheirfam et al., 2017b; Rossi et al., 2017; Mugnai et al., 2018). Less

information is available on polysaccharide formation in *Chloroflexi* and *Acidobacteria*, although the latter phylum is known to produce LPSs (Lagier et al., 2012), and a recent report shows that some of its members might have the capability to synthesize large quantities of EPSs (Kielak et al., 2017). Therefore, both types of biocrusts in our experiment were composed of taxa potentially able to produce high amounts of adhesive polysaccharides. However, while *Cyanobacteria* grow relatively fast, with some strains capable of growth rates comparable to industrial yeasts, *Acidobacteria* and *Chloroflexi* are slow-growers (Davis et al., 2010; Yu et al., 2015). Therefore, as all of the biocrusts in our experiment were cultivated in the same conditions (except for the underlying substrate) for the same amount of time, the development of the cyanobacterial biocrusts was faster (which we confirmed by visual observation), and consequently, these biocrusts likely had higher biomass than the biocrusts with higher proportion of *Acidobacteria* and *Chloroflexi*. This would mean that despite the similarity in the relative abundance of the EPS and LPS genes, the absolute abundance of these genes was higher in the cyanobacteria-dominated biocrusts compared with the others. Higher absolute abundance of genes responsible for the production of specific compounds potentially translates to higher yields of these compounds (Subramaniam et al., 2018), and thus we hypothesized that the difference in the stability of our biocrusts resulted from the difference in the absolute abundance of the genes involved in the production of adhesive polysaccharides. While we were unable to immediately confirm this hypothesis due to the lack of biomass measurements that would support our visual observations of biocrust development, these measurements were included in our later experiment based on three agricultural trials (**P3**), providing us with proof to our assumption.

In the agricultural experiment involving three field trials (**P3**), the highest aggregate stability was found at the site with the highest microbial biomass, whereas the lowest aggregate stability characterized the site that had the lowest microbial biomass. Microbial communities at the investigated sites consisted mostly of bacteria, and thus we used microbial biomass values as a reference to estimate the absolute abundance of the genes related to EPS and LPS biosynthesis. As the relative abundance of most of the analyzed genes showed little differences between the sites, the absolute abundance of these genes followed the same trends as microbial biomass, and thus we observed that increased absolute abundance of the genes involved in the production of adhesive polysaccharides was connected to improved aggregate stability. However, the absolute abundance of the EPS and LPS genes was not the only parameter that could explain the variability in aggregate stability observed between the sites in our experiment. In particular, soil texture and soil organic carbon content seemed to affect aggregate stability as well. Specifically, the site with the highest aggregate stability had the highest clay and organic carbon content, whereas at the site where aggregate stability was the lowest, also the lowest clay and organic carbon contents were found. As clay particles and organic carbon are amongst the most

important binding agents for soil aggregation (Bronick and Lal, 2005; Weil and Brady, 2017; Totsche et al., 2018), this result was not surprising. Nevertheless, soil texture and organic carbon content could not explain all of the variability in aggregate stability detected in our experiment, as differences in this soil parameter were found not only between sites but also between tillage treatments. Similarly to other studies that report more stable aggregates under RT compared with CT (Jacobs et al., 2009; Mikha et al., 2013; Bartlova et al., 2015; Sheehy et al., 2015; Singh et al., 2016), we observed this trend at one of our sites. This was correlated with neither clay nor organic carbon content, as soil texture was not affected by tillage intensity at either site, and organic carbon content was higher under less intensive tillage at all of the three sites. Conversely, the highest number of the analyzed genes whose absolute abundance was influenced by tillage system was detected at the site where aggregate stability was impacted by tillage as well. Furthermore, all of the genes affected by tillage at that site had higher absolute abundance under RT compared with CT. This showed that the absolute abundance of genes related to the production of adhesive polysaccharides is a better predictor of aggregate stability than soil texture and organic carbon content alone.

The question remains why the absolute abundance of EPS and LPS genes reacted to tillage intensity only at one of the investigated agricultural sites. One possible explanation is trial duration, as it was shown before that the effects of tillage on soil parameters build up over time (Stockfisch et al., 1999; Grandy et al., 2006), and the trial where we detected tillage effects was the oldest one of the three. The fact that pH was yet another parameter that differed under CT and RT at that trial but not others supports that theory. Otherwise, the impact of tillage intensity on soil bacteria could be modulated by site characteristics. In that case, the parameter important for defining the interaction between tillage treatment and bacterial communities in soil could be soil texture, as the size of soil particles influences their susceptibility to be aggregated together not only by physicochemical forces but also by organic agents such as EPSs and LPSs (Bronick and Lal, 2005; Weil and Brady, 2017; Totsche et al., 2018). Therefore, the intensity of both RT and CT could be insufficient to disturb bacteria living inside soil aggregates stabilized by high clay content, whereas in sandy soils, where the aggregation is hindered by the lack of small building particles, even smallest disturbance could be enough to influence aggregate communities. In the end, the difference between the level of disturbance caused by CT and RT could matter only in soils with balanced content of big and small particles. This theory is supported by the fact that the site where the absolute abundance of the EPS and LPS genes was most affected had the most balanced content of clay, silt and sand, whereas at the other sites, either clay or sand dominated.

However, as three trials with multiple varying parameters are too small sample to make any definite conclusions on the interactions between soil bacteria, tillage and environmental factors, further

research needs to build on our findings if the full understanding of the complexities of soil aggregation is to be obtained. In order to reach this goal, future analyses should also include other soil biota, as bacteria are not the only organisms capable of producing gluing agents. In agricultural soils, especially fungi could be of high importance, as previous studies showed that improved soil aggregation under less intense tillage corresponded to increased fungal biomass and glomalin production (Beare et al., 1997; Wright et al., 1999; Cookson et al., 2008; Dai et al., 2015; Lu et al., 2018). Unfortunately, as we used a metagenomic approach to analyze bacterial communities of potential polysaccharide producers to counter the methodological issues of directly measuring EPS and LPS concentrations in soil, investigating fungi involved in soil aggregation suffers from the opposite problem. While the methods of measuring glomalin content in soil are relatively well-established (Wright et al., 1999; Rillig et al., 2002; Borie et al., 2006; Bedini et al., 2009; Hammer and Rillig, 2011; Dai et al., 2015; Wang et al., 2017), the sequence databases are still biased towards bacteria and, thus, lack gene sequences specific for the biosynthesis of eukaryotic gluing agents in general, although efforts have been recently made to identify the glomalin gene in different species of arbuscular mycorrhizal fungi (Magurno et al., 2019). Furthermore, the analysis of eukaryotes by means of short-read sequencing is problematic due to the high number of intronic sequences possessed by these organisms. In addition, while glomalin is the most prominent gluing agent synthesized by fungi, these eukaryotes are also able to release polysaccharides, which not only have similar gluing properties to bacterial EPSs and LPSs, but directly measuring their contents in soil entails the same methodological disadvantages. Therefore, future work needs to be aimed at overcoming these obstacles if a single analytical approach encompassing all biota with the ability to produce soil-aggregating compounds is to be developed. Nevertheless, our findings indicate that while the relative abundance of the genes involved in the production of bacterial polysaccharides could influence aggregate stability, the absolute abundance of these genes is more likely to shape this soil parameter, as it is more prone to be affected by environmental factors compared with the highly stable relative abundance of the EPS and LPS genes.

3.4 Factors shaping the community composition of bacterial polysaccharides producers

High variability applied not only to the absolute abundance of the genes related to EPS and LPS biosynthesis but to the taxa harboring them as well. This shows that the production of adhesive polysaccharides is not always performed by the same group of bacteria (like oxygenic photosynthesis found only in *Cyanobacteria* (Martiny et al., 2015)). Instead, the community composition of polysaccharide producers seems to be shaped by a wide variety of environmental factors. While many

researchers made efforts to build a better understanding of the influence of different factors on the structure of bacterial communities, their findings and the relative importance of the different factors for shaping the community composition of bacteria were comprehensively reviewed by Fierer (2017). We managed to observe the possible effects of several of those factors in our work as well.

According to Fierer (2017), the major factor determining the composition of bacterial communities is pH. We saw the possible influence of this factor in our microcosm experiment (**P1**). There, bacterial communities present in the bulk soils from two initial ecosystems were highly similar and underwent differentiation during the development of biocrusts. As mentioned before, the major potential producers of adhesive polysaccharides in the developed biocrusts were *Cyanobacteria*, *Chloroflexi* and *Acidobacteria*. All the three bacterial groups are commonly found in biocrusts as well as other communities that embed themselves in EPS matrix (e.g. biofilms, microbial mats) (Mogul et al., 2017; Prieto-Barajas et al., 2017; Rampadarath et al., 2017). Therefore, their dominance within the communities of potential polysaccharide producers in the biocrusts grown in our microcosm experiment was not surprising. However, these phyla showed strong preferences towards the two soil substrates used in our study, and those differed significantly in pH. *Cyanobacteria* prevailed in the biocrusts developed on the soil substrate with slightly alkaline pH (~7.3). The other soil substrate had rather acidic pH (~5.4), and the biocrusts grown on it were characterized by the high proportion of *Acidobacteria* and *Chloroflexi* instead. Incidentally, *Cyanobacteria* prefer alkaline habitats (Belnap and Lange, 2003), whereas the other two groups favor acidic environments (Jones et al., 2009; Lauber et al., 2009; Wilhelm et al., 2011; Santofimia et al., 2013; Jones et al., 2017). This correlation between the abundance of the aforementioned phyla and the pH of the soil substrates used for cultivating the biocrusts suggests that pH could be a major factor shaping not only the overall structure of bacterial communities but also the community composition of bacterial EPS and LPS producers. However, it needs to be taken into consideration that the biocrusts in our experiment developed from indigenous communities of the bulk soils. Therefore, the observed differentiation of bacterial communities could have been caused as well by rare species that had too low abundance to be detectable in the bulk soils but started dominating during the development of the biocrusts. Furthermore, different pH of the soil substrates suggests that some of their other parameters (e.g. micronutrient availability) differed as well (Lammel et al., 2018), and those could have been the main drivers shaping the community structure of potential polysaccharide producers instead. The proposed influence of pH would, thus, have to be confirmed by further experiments.

The microcosm experiment (**P1**) helped to identify also nutrient availability as another potentially important factor influencing the community composition of polysaccharide producers. In this

experiment, even though we observed the above-described differentiation of biocrust communities that was possibly driven by pH, the initial communities of the bulk soils remained highly similar despite the different parameters of the soil substrates. However, the initial soil substrates used for cultivating the biocrusts had in common very low concentrations of carbon and nitrogen. It is likely that these could have selected bacteria adapted to habitats with limited nutritional opportunities. Fierer (2017) also postulated that the availability of nutrients such as carbon and nitrogen is, after pH, one of the factors that have the strongest impact on the structure of bacterial communities. This is in line with the fact that the families of potential polysaccharide producers with the highest relative abundance in the bulk soils used in our microcosm experiment were *Burkholderiaceae*, *Comamonadaceae*, *Moraxellaceae* and *Flavobacteriaceae*. The metabolic versatility and ability of the members of these groups to degrade a wide range of compounds enables them to survive even in such nutrient-restricted environments as crude oil, desert soil, glacier ice or distilled water (Coenye, 2014; McBride, 2014; Teixeira and Merquior, 2014; Willems, 2014). In contrast, the diversity of potential polysaccharide producers increased in the biocrusts as did the carbon and nitrogen availability. Aside from the microcosm experiment, we observed that the communities of potential polysaccharide producers changed over the course of the reclamation of a post-mining area, where nutrient availability improved with time as well (**P4**). In the initial stage of the reclamation process, the communities of potential producers of EPSs and LPSs were dominated by typical colonizers of oligotrophic environments, which were replaced by bacteria characteristic for agricultural soils in the later agricultural phase of the reclamation. This shows that good availability of nutrients favors the establishment of diverse communities of polysaccharide producers, which is in line with the high dependency of polysaccharide production on sufficient nutrient supply. More evidence that nutrient availability is a major factor shaping the community structure of potential polysaccharide producers can be found in our second agricultural experiment (**P3**). One of the two families with the highest relative abundance of the genes involved in EPS and LPS formation discovered when comparing potential polysaccharide producers from three agricultural field trials in Switzerland, Slovenia and Poland was *Sphingomonadaceae*. This finding was not surprising, as the members of this group are well-known for the production of EPSs called sphingans (Glaeser and Kämpfer, 2014), and they have been previously shown to improve aggregate stability (Caesar-TonThat et al., 2007). They are also frequently isolated from soils and plant rhizosphere (Glaeser and Kämpfer, 2014), which is in agreement with their high relative abundance in our samples of agricultural soils. However, even though *Sphingomonadaceae* were a major family at all the three investigated trials, this group was most characteristic for the site where carbon and nitrogen availability was the lowest, which could be explained by the oligotrophic traits that this family has (Glaeser and Kämpfer, 2014).

The other family with the highest relative abundance of the genes related to EPS and LPS synthesis amongst the potential polysaccharide producers from the three agricultural field trials that we compared (**P3**) was *Bradyrhizobiaceae*. Its members are commonly found in soils, often in close association with plant roots (de Souza et al., 2014), and some of them establish symbiotic relationships with plants such as lupine (Reeve et al., 2013; de Souza et al., 2014). For this process, EPSs and LPSs have been previously shown to be essential (Quelas et al., 2010). Symbiotic *Bradyrhizobiaceae* use these compounds also for forming biofilms in order to survive in the absence of their hosts (Seneviratne and Jayasinghearachchi, 2005). Therefore, the high relative abundance and potential of *Bradyrhizobiaceae* to produce adhesive polysaccharides in our samples of agricultural soils was expected. However, this group was especially prevalent at the site where lupine was growing during the time of sampling. This is in agreement with Fierer's (2017) observation that some bacterial groups preferentially associate with certain plant species, although the importance of plant species as a factor shaping bacterial communities is not as high as of pH or nutrient availability. The example of *Bradyrhizobiaceae* in our study shows that plant species could be yet another factor that influences not only the overall structure of bacterial communities but also the community composition of potential polysaccharide producers.

Aside from *Sphingomonadaceae* and *Bradyrhizobiaceae*, other potential producers of EPSs and LPSs differed in their relative abundance between the three analyzed agricultural sites as well (**P3**). However, identifying which exact factors drove those differences is difficult. That is because the sites differed not only in nutrient availability and the identity of plant species cultivated there at the time of sampling but also other factors. One of the factors that differed significantly between the sites and was also listed by Fierer (2017) as important for shaping bacterial communities was soil texture. As described before, the texture varied between the sites from sandy through silty to clayish, and this variation could help explain some of the observed bacterial responses. For example, according to a recent report (Hemkemeyer et al., 2018), *Alphaproteobacteria* prefer sand particle fraction, which incidentally dominates at our site with growing lupine. As *Bradyrhizobiaceae* belong to *Alphaproteobacteria*, for their dominance at that site might have thus been responsible not only the presence of lupine but also soil texture. Therefore, similarly to the absolute abundance of the EPS and LPS genes, explaining what is the contribution of different factors to shaping the community structure of potential producers of adhesive polysaccharides is not always straightforward, as several factors can have similar effect.

Another factor that demonstrates the difficulty of interpreting bacterial responses is depth. When we compared potential producers of EPSs and LPSs in agricultural soils at three different depths in our first agricultural experiment (**P2**), we found that depth as a factor had a strong impact on the relative abundance of these microorganisms. Strong responses of bacterial taxa to this factor were found

especially amongst the major potential polysaccharide producers in that study, such as *Chitinophagaceae* and *Nitrospiraceae*. The ability to produce large amounts of adhesive polysaccharides was previously described in those families (Daims, 2014; Oh et al., 2019; Wang et al., 2019), and they were found to be potentially important EPS and LPS producers during the agricultural management phase of the reclamation of the post-mining site that we investigated as well (**P4**). In our agricultural study (**P2**), *Chitinophagaceae* had the highest relative abundance in the top soil layers, whereas *Nitrospiraceae* prevailed in the lowest sampled soil layer. The impact of depth on these taxa was not surprising, as different soil layers are known to be characterized by distinct conditions that select only the best-adapted bacteria (Fierer et al., 2003; Griffiths et al., 2003; Uksa et al., 2015). Specifically, nutrient concentrations already discussed above, but also oxygen levels considered by Fierer (2017) as an important factor in shaping bacterial communities as well, generally decrease down the soil profile. This is in line with the fact that most members of *Chitinophagaceae* are aerobic heterotrophs (Rosenberg, 2014; Oh et al., 2019), whereas *Nitrospiraceae* are a diverse group that consists of both heterotrophs and autotrophs as well as aerobes and anaerobes (Daims, 2014). Consequently, although depth clearly had an impact on the community composition of potential EPS and LPS producers in our study, understanding the mechanism of its influence is difficult, as this factor is a resultant of other variables interacting with each other.

Furthermore, the depth factor in our study (**P2**) was not linked only to the oxygen and nutrient gradients but also the level of exposure to agricultural practices, especially tillage. Tillage is another complex factor, which encompasses changes in physical and chemical parameters along the soil profile as well as regular disturbances to microbial habitats. In our study, the upper sampled soil layers were directly exposed to tillage, whereas the lowest layer was below the tillage horizon. The dominance of *Chitinophagaceae* in the top soil layers might have been, thus, connected as well to the ability of some members of this family to form microcysts that could protect them from the disturbance introduced by tillage (Feng et al., 2019) or to their ability to degrade cellulose, which should be more readily available closer to the soil surface (Rosenberg, 2014). *Nitrospiraceae*, on the other hand, consist of some specialized slow-growing bacteria that could possibly thrive better beyond the reach of disturbance caused by tillage (Mundinger et al., 2019). Therefore, the effects of depth observed in that study could be as well attributed to tillage, as these two factors were closely interconnected.

In the follow-up study comprising the three agricultural field trials in Switzerland, Slovenia and Poland (**P3**), we demonstrated that some effects of tillage can be observed independently of depth as well. There, as mentioned before, the relative abundance of most potential polysaccharide producers differed between the trials, which could be explained by the differences in site conditions such as nutrient

availability, plant species or soil texture. However, the relative abundance of some taxa in that study differed between the tillage intensities as well, and in some cases these responses were identical at all the three trials, which shows that the difference in tillage intensity can be important for shaping the community structure of potential polysaccharide producers regardless of site-specific conditions. All the families of potential polysaccharides producers that showed a consistent response to tillage intensity at all the investigated sites belonged to *Actinobacteria* and had higher relative abundance under RT compared with CT. *Actinobacteria* are one of the most common and abundant bacterial phyla in soil, likely due to the protection provided by EPSs produced by some of its members in large quantities (Kielak et al., 2017). They were second most abundant phylum, after *Proteobacteria*, in all our agricultural soil samples as well. The actinobacterial families affected by tillage (except for *Glycomycetaceae*) were also found amongst the communities of the potential polysaccharide producers from the biocrusts cultivated in the microcosm experiment (**P1**). Actinobacterial EPSs have been recently gaining a lot of interest, as they have not only important ecological functions but also potential applications in industry and human health, including cancer treatment (Kielak et al., 2017; Selim et al., 2018). The negative impact of intensive tillage on this bacterial group could have resulted from the formation of filaments by its members (Rosenberg et al., 2014), which might be sensitive to disturbance similarly to fungal hyphae (Beare et al., 1997; Wright et al., 1999; Borie et al., 2006; Wang et al., 2010; Kihara et al., 2012; Lu et al., 2018). Furthermore, we discovered more taxa influenced by tillage intensity when analyzing the three trials separately (**P3**). Specifically, some families had different potential to produce EPSs and LPSs under CT and RT but only at one or two of the investigated sites. Moreover, the responses of the families whose potential to synthesize adhesive polysaccharides was affected by tillage intensity at two sites were inconsistent. For example, the relative abundance of *Oxalobacteraceae* was higher under CT at one of the investigated sites but under RT at another. There is little information available on the production of adhesive polysaccharides by *Oxalobacteraceae* (Hiraishi et al., 1997), although this family is ubiquitous in soils, and some of its members found application in agriculture as plant growth-promoting agents (Baldani et al., 2014). This shows that while some effects of tillage intensity are consistent across different trials, others depend on site-specific conditions. The example of *Oxalobacteraceae* underlines also the necessity of metagenomic studies, which can help identify groups potentially important for soil aggregation that require closer investigation in laboratory studies.

3.5 Functional redundancy of EPS and LPS biosynthesis, and what it could mean for ecosystem services

Regardless of which factors were responsible for the differences in the community structure of potential polysaccharide producers across the different samples, it was surprising that these differences were not reflected more by the relative abundances of the genes involved in EPS and LPS synthesis. As mentioned before, most differences in the relative abundance of these genes were found between initial soils and biocrusts grown on them in the microcosm experiment (**P1**). However, while the changes in the potential to produce adhesive polysaccharides that occurred during the development of those biocrusts coincided with the shifts in the community composition of potential EPS and LPS producers, they seemed to be linked more to the sample type (biocrust vs. initial soil) rather than the community structure of the potential producers of adhesive polysaccharides. This is supported by the fact that the biocrusts in our experiment displayed comparable potential for EPS and LPS synthesis despite differing in the community composition of potential polysaccharide producers. Moreover, little differences in the relative abundance of the genes related to EPS and LPS production were observed between the samples of agricultural soils from three different trials even though the communities of the potential producers of adhesive polysaccharides in those samples differed in their structure (**P3**). We additionally observed that while the communities of potential polysaccharide producers changed over the course of the reclamation of the investigated post-mining area, the relative abundance of the genes involved in EPS and LPS synthesis in those communities remained constant (**P4**). These findings were unexpected because it has been repeatedly shown before that the functioning of bacterial communities strongly depends on their taxonomic composition (Langenheder et al., 2006; Strickland et al., 2009; Reed and Martiny, 2013; Logue et al., 2016).

However, other studies have demonstrated that the functional structure of different bacterial communities may be alike despite taxonomic differences between them if the communities inhabit similar habitats (Louca et al., 2018). These observations are consistent with the theory of functional redundancy, according to which important functions are conserved within bacterial communities even when their members change (Allison and Martiny, 2008). Similar conclusions were drawn by Fondi et al. (2016), who postulated that the functional composition of bacterial communities depends primarily on their broad ecological niches, e.g. sea or fresh water, soil, host and airborne. Therefore, the little difference in the relative abundance of the genes related to EPS and LPS formation observed in our experiments (**P1**, **P3**, **P4**) despite the high taxonomic variability of the potential polysaccharide producers could be explained if the potential to produce adhesive polysaccharides was an important trait for bacteria inhabiting agricultural soils and biocrusts from initial ecosystems. The promotion of the

bacterial potential to produce adhesive polysaccharides in these environments is especially probable, as many of EPSs' and LPSs' properties could increase the ecological fitness of bacteria living in these habitats, and most taxa detected in our samples of biocrusts and agricultural soils harbored the genes involved in the synthesis of these compounds. Finally, this hypothesis is supported by the lower number and relative abundance of potential polysaccharide producers coupled with the lower potential to form EPSs and LPSs in the initial soils from our microcosm experiment (**P1**) compared with both biocrust and agricultural soil samples.

However, even if the relative abundance of the genes related to the formation of adhesive polysaccharides remained mostly stable between samples, the bacterial ability to aggregate soil could differ anyway due to the great taxonomic variability of potential polysaccharide producers. That is because the properties of bacterial polysaccharides strongly depend on their structure, and even slight changes may result in considerably different parameters (Suresh Kumar et al., 2007). The structure of bacterial polysaccharides can differ between two strains of the same species and varies even more greatly between higher taxonomic levels (Sutherland and Thomson, 1975; Celik et al., 2008). Therefore, it is likely that polysaccharides produced by different taxa have also different aggregating properties. Moreover, gene induction and transcription pathways may differ between polysaccharide producers as well. It was shown before that homologous genes can be differentially expressed even in two strains belonging to the same species (Vital et al., 2015; Haryono et al., 2019). There, the differentially expressed functions included stress defense and carbohydrate metabolism. Therefore, it is probable that the genes related to EPS and LPS formation also have various expression patterns in different taxa. Lastly, as discussed before, the ecology of different bacterial species might influence the absolute abundance of the EPS and LPS genes, which we identified as a potentially important parameter for soil aggregation. Regardless of the mechanism, it can be expected that different polysaccharide producers differ in their soil aggregation capabilities. This is in line with the observation that the effect of pure bacterial cultures on soil aggregation depends on the bacterial species (Costa et al., 2018). In the end, the more different the composition of two bacterial communities is, the more likely the adhesive polysaccharide mixtures produced by them are to have different aggregating efficiency. Our experiments show that this could be important for evaluating the capacity to aggregate soil by bacterial communities from habitats with different environmental conditions, as those are most likely to differ in the composition of their polysaccharide producers.

4 Conclusions and outlook

The metagenomic pipeline that we established allowed us to investigate the bacterial potential to influence soil structure without the biases associated with traditional laboratory methodologies. Using this approach, we identified the key genes of EPS and LPS production. Amongst the most abundant ones were *wza*, *lptF* and *lptG*, encoding for the EPS export outer membrane protein Wza and the transmembrane proteins LptF and LptG of the LptB2FGCADE LPS export complex. We found that the relative abundance of the genes related to the synthesis of adhesive polysaccharides can differ between sample types (e.g. biocrust vs. initial bulk soil), and the increased relative abundance of the EPS and LPS genes is accompanied in those cases by improved soil aggregation. Amongst samples of the same type, the relative abundance of the genes involved in polysaccharide production remains constant, and differences in aggregate stability can be driven either by different taxonomic composition of the communities of potential polysaccharide producers or differences in the absolute abundance of the genes involved in EPS and LPS biosynthesis. We discovered that the taxa harboring the EPS and LPS genes are highly diverse even in samples of the same type, and their number and community composition changes in response to a variety of different factors. Amongst those factors are pH, nutrient availability, soil texture and tillage. While low nutrient content seems to be the most important factor limiting the number and diversity of the potential producers of EPSs and LPSs, in the habitats where nutrient availability does not limit bacterial growth and functioning, other factors seem to play a more important role in shaping the communities of potential polysaccharide producers. We proposed the mechanisms of the observed responses of the potential producers of adhesive polysaccharide to these factors, but confirming them would require further investigation. That is because identifying with certainty connections between bacterial responses and environmental factors that caused them in multivariate studies such as ours is difficult, if not impossible, especially when those factors have similar effects or interact with each other. Therefore, future work should focus on explaining the complexities of the influences of the discussed factors. This could be accomplished in small-scale experiments by minimizing the number of varying factors and using the same initial communities or single isolates. Several key groups of potential polysaccharide producers that could be of interest for this purpose were proposed in our work. These include phyla such as *Cyanobacteria*, *Acidobacteria*, *Chloroflexi* and *Actinobacteria*, as well as families such as *Burkholderiaceae*, *Comamonadaceae*, *Moraxellaceae*, *Flavobacteriaceae*, *Sphingomonadaceae*, *Bradyrhizobiaceae*, *Chitinophagaceae*, *Nitrospiraceae* and *Oxalobacteraceae*. Not only the susceptibility of the potential producers of EPSs and LPSs to different factors should be measured, but the adhesive properties of polysaccharides produced by them under different conditions should be assessed as well. Experiments on a larger scale should take a form of meta-studies, which would allow for performing proper multivariate analyses. Further investigation of

other habitats aside from agricultural soils and biocrusts would also help explaining if the promotion of the potential to produce adhesive polysaccharides observed in our experiments applies to soils in general as well as to other environments. In addition, changing the sampling strategy taking into account the fast turnover of mRNA would allow to investigate if the expected differences in gene expression patterns of different communities are affected primarily by fluctuations of different factors, and if they could have lasting long-term effects on soil aggregation. Finally, further effort should be made to develop a single approach to analyze gluing agents of different origins (bacterial, fungal, etc.). This would allow for true meta-studies encompassing all organisms involved in soil aggregation and bring us much closer to reaching the full understanding of the complexities of this essential process. However, while a lot of work is still required to accomplish this research goal, our findings could be used for the benefit of world soils even now. Specifically, educating agricultural practitioners on the influence of different tillage intensities on the bacterial population that improves soil structure could help boosting the awareness of the causal link between intensive tillage and soil erosion and speed up the transition to more sustainable agricultural practices. Moreover, the knowledge of the taxa potentially capable of producing high amounts of adhesive polysaccharides could be used not only for designing more targeted isolation strategies of new strains with scientific and industrial value but also developing bioinocula that would improve the capacity of the autochthonous microflora to stabilize soil structure. Such measures might act as a mitigation to the damage caused by erosion while further research will build on our findings to detangle the complexities of polysaccharide-producing bacterial communities and factors affecting them.

References

Abram F. (2015). Systems-based approaches to unravel multi-species microbial community functioning. *Comput Struct Biotechnol J* 13**,** 24-32.

Aho K. (2015). "asbio: a collection of statistical tools for biologists". R package version 1.1-5.

Allison S.D., and Martiny J.B.H. (2008). Resistance, resilience, and redundancy in microbial communities. *P Natl Acad Sci USA* 105**,** 11512-11519.

Amezketa E. (1999). Soil aggregate stability: a review. *J Sustain Agr* 14**,** 83-151.

Antunes L.C.S., Poppleton D., Klingl A., Criscuolo A., Dupuy B., Brochier-Armanet C., Beloin C., and Gribaldo S. (2016). Phylogenomic analysis supports the ancestral presence of LPS-outer membranes in the Firmicutes. *Elife* 5**,** e14589.

Arriaga F.J., Guzman J., and Lowery B. (2017). "Conventional agricultural production systems and soil functions" in *Soil health and intensification of agroecosytems,* eds. M.M. Al-Kaisi & B. Lowery. (London: Elsevier), 109-125.

Asgari H. (2014). Effect of agronomic practices on the aggregate stability and organic carbon of soil (case study: the northern of Aq Qala). *Env Resour Res* 2**,** 95-106.

Ates O. (2015). Systems biology of microbial exopolysaccharides production. *Front Bioeng Biotechnol* 3**,** 200.

Babin D., Deubel A., Jacquiod S., Sørensen S.J., Geistlinger J., Grosch R., and Smalla K. (2019). Impact of long-term agricultural management practices on soil prokaryotic communities. *Soil Biol Biochem* 129**,** 17-28.

Bach E., and Wall D. (2018). "Trends in global biodiversity: soil biota and processes" in *Encyclopedia of the anthropocene,* eds. D.A. Dellasala & M.I. Goldstein. (Oxford: Elsevier), 125-130.

Baldani J.I., Rouws L., Cruz L.M., Olivares F.L., Schmid M., and Hartmann A. (2014). "The family Oxalobacteraceae" in *The prokaryotes: Alphaproteobacteria and Betaproteobacteria,* ed. E. Rosenberg. (Heidelberg: Springer), 919-974.

Bar-On Y.M., Phillips R., and Milo R. (2018). The biomass distribution on Earth. *P Natl Acad Sci USA* 115**,** 6506-6511.

Bartlova J., Badalikova B., Pospíšilová L., Pokorný E., and Šarapatka B. (2015). Water stability of soil aggregates in different systems of tillage. *Soil Water Res* 10**,** 147-154.

Beare M., Hu S., Coleman D., and Hendrix P. (1997). Influences of mycelial fungi on soil aggregation and organic matter storage in conventional and no-tillage soils. *Appl Soil Ecol* 5**,** 211-219.

Bedini S., Pellegrino E., Avio L., Pellegrini S., Bazzoffi P., Argese E., and Giovannetti M. (2009). Changes in soil aggregation and glomalin-related soil protein content as affected by the arbuscular mycorrhizal fungal species Glomus mosseae and Glomus intraradices. *Soil Biol Biochem* 41**,** 1491-1496.

Belnap J., and Gardner J.S. (1993). Soil microstructure in soils of the Colorado Plateau: the role of the cyanobacterium Microcoleus vaginatus. *Great Basin Nat* 53**,** 6.

Belnap J., and Lange O.L. (2003). *Biological soil crusts: structure, function, and management.* New York: Springer.

Benard P., Zarebanadkouki M., Brax M., Kaltenbach R., Jerjen I., Marone F., Couradeau E., Felde V.J., Kaestner A., and Carminati A. (2019). Microhydrological Niches in Soils: How Mucilage and EPS Alter the Biophysical Properties of the Rhizosphere and Other Biological Hotspots. *Vadose Zone J* 18.

Benedet M., Falchi F.A., Puccio S., Di Benedetto C., Peano C., Polissi A., and Dehò G. (2016). The lack of the essential LptC protein in the trans-envelope lipopolysaccharide transport machine is circumvented by suppressor mutations in LptF, an inner membrane component of the Escherichia coli transporter. *PloS One* 11**,** e0161354.

Berne C., Ducret A., Hardy G.G., and Brun Y.V. (2015). Adhesins involved in attachment to abiotic surfaces by Gram-negative bacteria. *Microbiol Spectr* 3.

Birch J., Van Calsteren M.-R., Pérez S., and Svensson B. (2019). The exopolysaccharide properties and structures database: EPS-DB. Application to bacterial exopolysaccharides. *Carbohyd Polym* 205**,** 565-570.

Bird S.B., Herrick J.E., Wander M.M., and Murray L. (2007). Multi-scale variability in soil aggregate stability: implications for understanding and predicting semi-arid grassland degradation. *Geoderma* 140**,** 106-118.

Bohl T.E., and Aihara H. (2018). Current progress in the structural and biochemical characterization of proteins involved in the assembly of lipopolysaccharide. *Int J Microbiol* 2018**,** Article ID 5319146.

Borie F., Rubio R., Rouanet J., Morales A., Borie G., and Rojas C. (2006). Effects of tillage systems on soil characteristics, glomalin and mycorrhizal propagules in a Chilean Ultisol. *Soil Tillage Res* 88**,** 253-261.

Borrelli P., Robinson D.A., Fleischer L.R., Lugato E., Ballabio C., Alewell C., Meusburger K., Modugno S., Schütt B., and Ferro V. (2017). An assessment of the global impact of 21st century land use change on soil erosion. *Nat Commun* 8**,** 2013.

Bronick C.J., and Lal R. (2005). Soil structure and management: a review. *Geoderma* 124, 3–22.

Brookes P., Landman A., Pruden G., and Jenkinson D. (1985). Chloroform fumigation and the release of soil nitrogen: a rapid direct extraction method to measure microbial biomass nitrogen in soil. *Soil Biol Biochem* 17**,** 837-842.

Buchfink B., Xie C., and Huson D.H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12**,** 59-60.

Büdel B., Darienko T., Deutschewitz K., Dojani S., Friedl T., Mohr K.I., Salisch M., Reisser W., and Weber B. (2009). Southern African biological soil crusts are ubiquitous and highly diverse in drylands, being restricted by rainfall frequency. *Microb Ecol* 57**,** 229-247.

Caesar-Tonthat T., Caesar A.J., Gaskin J.F., Sainju U.M., and Busscher W. (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 T., Stevens W.B., Sainju U.M., Caesar A.J., West M., and Gaskin J.F. (2014). Soil-aggregating bacterial community as affected by irrigation, tillage, and cropping system in the northern great plains. *Soil Sci* 179**,** 11-20.

Celik G.Y., Aslim B., and Beyatli Y. (2008). Characterization and production of the exopolysaccharide (EPS) from Pseudomonas aeruginosa G1 and Pseudomonas putida G12 strains. *Carbohyd Polym* 73**,** 178-182.

Coenye T. (2014). "The family Burkholderiaceae" in *The prokaryotes: Alphaproteobacteria and Betaproteobacteria,* ed. E. Rosenberg. (Heidelberg: Springer), 759-776.

Colica G., Li H., Rossi F., Li D., Liu Y., and De Philippis R. (2014). Microbial secreted exopolysaccharides affect the hydrological behavior of induced biological soil crusts in desert sandy soils. *Soil Biol Biochem* 68**,** 62-70.

Colica G., Li H., Rossi F., Philippis R.D., and Liu Y. (2015). Differentiation of the characteristics of excreted extracellular polysaccharides reveals the heterogeneous primary succession of induced biological soil crusts. *J Appl Phycol* 27**,** 24-32.

Cookson W.R., Murphy D.V., and Roper M.M. (2008). Characterizing the relationships between soil organic matter components and microbial function and composition along a tillage disturbance gradient. *Soil Biol Biochem* 40**,** 763-777.

Cooper J., Baranski M., Stewart G., Nobel-De Lange M., Bàrberi P., Fließbach A., Peigné J., Berner A., Brock C., and Casagrande M. (2016). Shallow non-inversion tillage in organic farming maintains crop yields and increases soil C stocks: a meta-analysis. *Agron Sustain Dev* 36**,** 1-20.

Costa O.Y., Raaijmakers J.M., and Kuramae E.E. (2018). Microbial extracellular polymeric substances: ecological function and impact on soil aggregation. *Front Microbiol* 9**,** 1-14.

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

Dai J., Hu J., Zhu A., Bai J., Wang J., and Lin X. (2015). No tillage enhances arbuscular mycorrhizal fungal population, glomalin-related soil protein content, and organic carbon accumulation in soil macroaggregates. *J Soil Sediment* 15**,** 1055-1062.

Daims H. (2014). "The family Nitrospiraceae" in *The prokaryotes: other major lineages of bacteria and the archaea,* ed. E. Rosenberg. (Berlin, Heidelberg: Springer), 733-749.

Danese P.N., Pratt L.A., and Kolter R. (2000). Exopolysaccharide production is required for development of Escherichia coli K-12 biofilm architecture. *J Bacteriol* 182**,** 3593-3596.

Davis K.E.R, Sangwan P., and Janssen P.H. (2011). Acidobacteria, Rubrobacteridae and Chloroflexi are abundant among very slow-growing and mini-colony-forming soil bacteria. *Environ Microbiol* 13, 798-805.

De Caire G.Z., De Cano M.S., De Mule M.Z., Palma R., and Colombo K. (1997). Exopolysaccharide of Nostoc muscorum (Cyanobacteria) in the aggregation of soil particles. *J Appl Phycol* 9**,** 249-253.

De Souza J.A.M., Alves L.M.C., De Mello Varani A., and De Macedo Lemos E.G. (2014). "The family Bradyrhizobiaceae" in *The prokaryotes: Alphaproteobacteria and Betaproteobacteria,* ed. R. E. (Heidelber.g: Springer), 135-154.

Debroy C., Fratamico P.M., Yan X., Baranzoni G., Liu Y., Needleman D.S., Tebbs R., O'Connell C.D., Allred A., and Swimley M. (2016). Comparison of O-antigen gene clusters of all O-serogroups of Escherichia coli and proposal for adopting a new nomenclature for O-typing. *PLoS One* 11**,** e0147434.

Delbarre-Ladrat C., Sinquin C., Lebellenger L., Zykwinska A., and Colliec-Jouault S. (2014). Exopolysaccharides produced by marine bacteria and their applications as glycosaminoglycan-like molecules. *Front Chem* 2**,** 85.

Dertli E., Colquhoun I.J., Gunning A.P., Bongaerts R.J., Le Gall G., Bonev B.B., Mayer M.J., and Narbad A. (2013). Structure and biosynthesis of two exopolysaccharides produced by Lactobacillus johnsonii FI9785. *J Biol Chem* 288**,** 31938-31951.

Dong H., Zhang Z., Tang X., Paterson N.G., and Dong C. (2017). Structural and functional insights into the lipopolysaccharide ABC transporter LptB 2 FG. *Nat Commun* 8**,** 222.

Duncan M.C., Forbes J.C., Nguyen Y., Shull L.M., Gillette R.K., Lazinski D.W., Ali A., Shanks R.M., Kadouri D.E., and Camilli A. (2018). Vibrio cholerae motility exerts drag force to impede attack by the bacterial predator Bdellovibrio bacteriovorus. *Nat Commun* 9**,** 4757.

Enagbonma B.J., and Babalola O.O. (2019). Potentials of termite mound soil bacteria in ecosystem engineering for sustainable agriculture. *Ann Microbiol* 69**,** 211-219.

Falkowski P.G., Fenchel T., and Delong E.F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science* 320**,** 1034-1039.

FAO (2011). *The state of the world's land and water resources for food and agriculture (SOLAW) – managing systems at risk.* London: Rome and Earthscan.

FAO (2019). *Soil erosion: the greatest challenge to sustainable soil management.* Rome: Food and Agriculture Organization of the United Nations.

FAO, and ITPS (2015). *Status of the world's soil resources (SWSR) – main report.* Rome: Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils.

Feng R., Zhao G., Yang Y., Xu M., Huang S., Sun G., Guo J., and Li J. (2019). Enhanced biological removal of intermittent VOCs and deciphering the roles of sodium alginate and polyvinyl alcohol in biofilm formation. *PloS One* 14.

Field A., Miles J., and Field Z. (2012). *Discovering statistics using R.* Thousand Oaks, California: Sage publications.
Fierer N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* 15**,** 579-590.
Fierer N., Schimel J.P., and Holden P.A. (2003). Variations in microbial community composition through two soil depth profiles. *Soil Biol Biochem* 35**,** 167-176.
Finch H., Samuel A., and Lane G. (2014). *Lockhart & Wiseman's crop husbandry including grassland.* London: Elsevier.
Finn R.D., Coggill P., Eberhardt R.Y., Eddy S.R., Mistry J., Mitchell A.L., Potter S.C., Punta M., Qureshi M., and Sangrador-Vegas A. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44**,** D279-D285.
Fischer T., Veste M., Schaaf W., Dümig A., Kögel-Knabner I., Wiehe W., Bens O., and Hüttl R.F. (2010a). Initial pedogenesis in a topsoil crust 3 years after construction of an artificial catchment in Brandenburg, NE Germany. *Biogeochemistry* 101**,** 165-176.
Fischer T., Veste M., Wiehe W., and Lange P. (2010b). Water repellency and pore clogging at early successional stages of microbiotic crusts on inland dunes, Brandenburg, NE Germany. *Catena* 80**,** 47-52.
Flachowsky G., Meyer U., and Südekum K.-H. (2017). Land use for edible protein of animal origin - a review. *Animals* 7**,** 25.
Fondi M., Karkman A., Tamminen M.V., Bosi E., Virta M., Fani R., Alm E., and Mcinerney J.O. (2016). "Every gene is everywhere but the environment selects": Global geolocalization of gene sharing in environmental samples through network analysis. *Genome Biol Evol* 8**,** 1388-1400.
Ford S.R., Webster J.J., and Leach F.R. (1985). Extraction of bacterial lipopolysaccharides from soil samples and their determination using the Limulus amebocyte lysate reaction. *Soil Biol Biochem* 17**,** 811-814.
Gadermaier F., Berner A., Fließbach A., Friedel J.K., and Mäder P. (2012). Impact of reduced tillage on soil organic carbon and nutrient budgets under organic farming. *Renew Agr Food Syst* 27**,** 68-80.
Garmiri P., Coles K.E., Humphrey T.J., and Cogan T.A. (2008). Role of outer membrane lipopolysaccharides in the protection of Salmonella enterica serovar Typhimurium from desiccation damage. *FEMS Microbiol Lett* 281**,** 155-159.
Gilley J. (2005). "Erosion water-induced" in *Encyclopedia of soils in the environment,* ed. D. Hillel. (London: Elsevier), 463-469.
Glaeser S.P., and Kämpfer P. (2014). "The family Sphingomonadaceae" in *The prokaryotes: Alphaproteobacteria and Betaproteobacteria,* ed. R. E. (Heidelberg: Springer), 641-707.
Glöckner F.O., Yilmaz P., Quast C., Gerken J., Beccati A., Ciuprina A., Bruns G., Yarza P., Peplies J., and Westram R. (2017). 25 years of serving the community with ribosomal RNA gene reference databases and tools. *J Biotechnol* 261**,** 169-176.
Grandy A.S., Robertson G.P., and Thelen K.D. (2006). Do productivity and environmental trade-offs justify periodically cultivating no-till cropping systems? *Agron. J* 98, 1377–1383.
Greenfield L.K., and Whitfield C. (2012). Synthesis of lipopolysaccharide O-antigens by ABC transporter-dependent pathways. *Carbohyd Res* 356**,** 12-24.
Griffiths R.I., Whiteley A.S., O'Donnell A.G., and Bailey M.J. (2003). Influence of depth and sampling time on bacterial community structure in an upland grassland soil. *FEMS Microbiol Ecol* 43**,** 35-43.
Haft D.H., Selengut J.D., Richter R.A., Harkins D., Basu M.K., and Beck E. (2013). TIGRFAMs and genome properties in 2013. *Nucleic Acids Res* 41**,** D387-D395.
Hammer E.C., and Rillig M.C. (2011). The influence of different stresses on glomalin levels in an arbuscular mycorrhizal fungus—salinity increases glomalin content. *PLoS One* 6.

Haryono M., Cho S.-T., Fang M.-J., Chen A.-P., Chou S.-J., Lai E.-M., and Kuo C.-H. (2019). Differentiations in gene content and expression response to virulence induction between two Agrobacterium strains. *Front Microbiol* 10**,** 1554.

Hay I.D., Wang Y., Moradali M.F., Rehman Z.U., and Rehm B.H.A. (2014). Genetics and regulation of bacterial alginate production. *Environ Microbiol* 16**,** 2997-3011.

Heberle H., Meirelles G.V., Da Silva F.R., Telles G.P., and Minghim R. (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* 16**,** 169.

Hemkemeyer M., Dohrmann A.B., Christensen B.T., and Tebbe C.C. (2018). Bacterial preferences for specific soil particle size fractions revealed by community analyses. *Front Microbiol* 9**,** 149.

Hillel D. (2008). *Soil in the environment.* Amsterdam: Elsevier.

Hiraishi A., Shin Y.K., and Sugiyama J. (1997). Proposal To Reclassify Zoogloea ramigera IAM 12670 (PR Dugan 115) as Duganella zoogloeoides gen. nov., sp. nov. *Int J Syst Evol Micr* 47**,** 1249-1252.

Huerta-Cepas J., Szklarczyk D., Forslund K., Cook H., Heller D., Walter M.C., Rattei T., Mende D.R., Sunagawa S., and Kuhn M. (2015). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* 44**,** D286-D293.

Huixia P., Zhengming C., Xuemei Z., Shuyong M., Xiaoling Q., and Fang W. (2007). A study on an oligotrophic bacteria and its ecological characteristics in an arid desert area. *Sci China Ser D* 50**,** 128-134.

Huson D.H., Mitra S., Ruscheweyh H.-J., Weber N., and Schuster S.C. (2011). Integrative analysis of environmental sequences using MEGAN4. *Genome Res* 21**,** 1552-1560.

Jacobs A., Rauber R., and Ludwig B. (2009). Impact of reduced tillage on carbon and nitrogen storage of two Haplic Luvisols after 40 years. *Soil Till Res* 102**,** 158-164.

Jacques M. (1996). Role of lipo-oligosaccharides and lipopolysaccharides in bacterial adherence. *Trends Microbiol* 4**,** 408-410.

Joergensen R.G. (1996). The fumigation-extraction method to estimate soil microbial biomass: calibration of the kEC value. *Soil Biol Biochem* 28**,** 25-31.

Joergensen R.G., and Mueller T. (1996). The fumigation-extraction method to estimate soil microbial biomass: calibration of the kEN value. *Soil Biol Biochem* 28**,** 33-37.

Jones D.S., Lapakko K.A., Wenz Z.J., Olson M.C., Roepke E.W., Sadowsky M.J., Novak P.J., and Bailey J.V. (2017). Novel microbial assemblages dominate weathered sulfide-bearing rock from copper-nickel deposits in the Duluth complex, Minnesota, USA. *Appl Environ Microbiol* 83**,** e00909-00917.

Jones R.T., Robeson M.S., Lauber C.L., Hamady M., Knight R., and Fierer N. (2009). A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *ISME J* 3**,** 442-453.

Jouquet P., Bottinelli N., Shanbhag R.R., Bourguignon T., Traoré S., and Abbasi S.A. (2016). Termites: the neglected soil engineers of tropical soils. *Soil Sci* 181**,** 157-165.

Kalev S.D., and Toor G.S. (2018). "The composition of soils and sediments" in *Green chemistry: an inclusive approach,* eds. B. Torok & T. Dransfield. (Amsterdam: Elsevier), 339-357.

Kandel S.L., Joubert P.M., and Doty S.L. (2017). Bacterial endophyte colonization and distribution within plants. *Microorganisms* 5**,** 77.

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

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

Kheirfam H., Sadeghi S.H., Darki B.Z., and Homaee M. (2017a). Controlling rainfall-induced soil loss from small experimental plots through inoculation of bacteria and cyanobacteria. *Catena* 152**,** 40-46.

Kheirfam H., Sadeghi S.H., Homaee M., and Darki B.Z. (2017b). Quality improvement of an erosion-prone soil through microbial enrichment. *Soil Till Res* 165**,** 230-238.
Kielak A.M., Castellane T.C., Campanharo J.C., Colnago L.A., Costa O.Y., Da Silva M.L.C., Van Veen J.A., Lemos E.G., and Kuramae E.E. (2017). Characterization of novel Acidobacteria exopolysaccharides with potential industrial and ecological applications. *Sci Rep-UK* 7**,** 41193.
Kierek K., and Watnick P.I. (2003). The Vibrio cholerae O139 O-antigen polysaccharide is essential for Ca2+-dependent biofilm development in sea water. *P Natl Acad Sci USA* 100**,** 14357-14362.
Kihara J., Martius C., Bationo A., Thuita M., Lesueur D., Herrmann, L., Amelung, W., and Vlek, P.L. (2012). Soil aggregation and total diversity of bacteria and fungi in various tillage systems of sub-humid and semi-arid Kenya. *Appl Soil Ecol* 58**,** 12-20.
Kim M.S., Kim Y.D., Hong S.S., Park K., Ko K.S., and Myung, H. (2015). Phage-encoded colanic acid-degrading enzyme permits lytic phage infection of a capsule-forming resistant mutant Escherichia coli strain. *Appl Environ Microbiol* 81**,** 900-909.
Köhne J.M., Schlüter S., and Vogel H.-J. (2011). Predicting solute transport in structured soil using pore network models. *Vadose Zone J* 10**,** 1082-1096.
Kopylova E., Noé L., and Touzet H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28**,** 3211-3217.
Krauss M., Ruser R., Müller T., Hansen S., Mäder P., and Gattinger A. (2017). Impact of reduced tillage on greenhouse gas emissions and soil carbon stocks in an organic grass-clover ley-winter wheat cropping sequence. *Agric Ecosyst Environ* 239**,** 324-333.
Kwon K.-J., Park K.-J., Kim J.-D., Kong J.-Y., and Kong I.-S. (1994). Isolation of two different polysaccharides from halophilic Zoogloea sp. *Biotechnol Lett* 16**,** 783-788.
Lagier J.-C., Million M., Hugon P., Armougom F., and Raoult D. (2012). Human gut microbiota: repertoire and variations. *Front Cell Infect Mi* 2**,** 136.
Lal R. (2001). Soil degradation by erosion. *Land Degrad Dev* 12**,** 519-539.
Lammel D.R., Barth G., Ovaskainen O., Cruz L.M., Zanatta J.A., Ryo M., De Souza E.M., and Pedrosa F.O. (2018). Direct and indirect effects of a pH gradient bring insights into the mechanisms driving prokaryotic community structures. *Microbiome* 6**,** 106.
Lan S., Wu L., Zhang D., and Hu C. (2012). Successional stages of biological soil crusts and their microstructure variability in Shapotou region (China). *Environ Earth Sci* 65**,** 77-88.
Langenheder S., Lindström E.S., and Tranvik L.J. (2006). Structure and function of bacterial communities emerging from different sources under identical conditions. *Appl Environ Microbiol* 72**,** 212-220.
Langley S., and Beveridge T. (1999). Effect of O-side-chain-lipopolysaccharide chemistry on metal binding. *Appl Environ Microbiol* 65**,** 489-498.
Lauber C.L., Hamady M., Knight R., and Fierer N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl Environ Microbiol* 75**,** 5111-5120.
Ledeboer N.A., and Jones B.D. (2005). Exopolysaccharide sugars contribute to biofilm formation by Salmonella enterica serovar Typhimurium on HEp-2 cells and chicken intestinal epithelium. *J Bacteriol* 187**,** 3214-3226.
Lehmann A., Leifheit E., and Rillig M.C. (2017a). "Mycorrhizas and soil aggregation" in *Mycorrhizal mediation of soil: fertility, structure, and carbon storage,* eds. N.C. Johnson, C. Gehring & J. Jansa. (Amsterdam: Elsevier).
Lehmann A., Zheng W., and Rillig M.C. (2017b). Soil biota contributions to soil aggregation. *Nat Ecol Evol* 1**,** 1828-1835.
Lerouge I., and Vanderleyden J. (2002). O-antigen structural variation: mechanisms and possible roles in animal/plant–microbe interactions. *FEMS Microbiol Rev* 26**,** 17-47.

Lindhout T., Lau P.C.Y., Brewer D., and Lam J.S. (2009). Truncation in the core oligosaccharide of lipopolysaccharide affects flagella-mediated motility in Pseudomonas aeruginosa PAO1 via modulation of cell surface attachment. *Microbiology+* 155**,** 3449-3460.

Logue J.B., Stedmon C.A., Kellerman A.M., Nielsen N.J., Andersson A.F., Laudon H., Lindström E.S., and Kritzberg E.S. (2016). Experimental insights into the importance of aquatic bacterial community composition to the degradation of dissolved organic matter. *ISME J* 10**,** 533.

Louca S., Polz M.F., Mazel F., Albright M.B., Huber J.A., O'Connor M.I., Ackermann M., Hahn A.S., Srivastava D.S., and Crowe S.A. (2018). Function and functional redundancy in microbial systems. *Nat Ecol Evol* 2**,** 936.

Low K.E., and Howell P.L. (2018). Gram-negative synthase-dependent exopolysaccharide biosynthetic machines. *Curr Opin Struc Biol* 53**,** 32-44.

López G., Diaz-Cárdenas C., Shapiro N., Woyke T., Kyrpides N.C., Alzate J.D., González L.N., Restrepo S., and Baena S. (2017). Draft genome sequence of Pseudomonas extremaustralis strain USBA-GBX 515 isolated from Superparamo soil samples in Colombian Andes. *Stand Genomic Sci* 12**,** 78.

Lu X., Lu X., and Liao Y. (2018). Effect of tillage treatment on the diversity of soil arbuscular mycorrhizal fungal and soil aggregate-associated carbon content. *Front Microbiol* 9**,** 1-10.

Lueders T., Manefield M., and Friedrich M.W. (2004). Enhanced sensitivity of DNA-and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. *Environ Microbiol* 6**,** 73-78.

Mäder P., and Berner A. (2012). Development of reduced tillage systems in organic farming in Europe. *Renew Agr Food Syst* 27**,** 7-11.

Magurno F., Malicka M., Posta K., Woźniak G., Lumini E., and Piotrowska-Seget Z. (2019). Glomalin gene as molecular marker for functional diversity of arbuscular mycorrhizal fungi in soil. *Biol Fert Soils* 55, 411–417.

Marchler-Bauer A., Derbyshire M.K., Gonzales N.R., Lu S., Chitsaz F., Geer L.Y., Geer R.C., He J., Gwadz M., and Hurwitz D.I. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43**,** D222-D226.

Martiny J.B., Jones S.E., Lennon J.T., and Martiny A.C. (2015). Microbiomes in light of traits: a phylogenetic perspective. *Science* 350**,** aac9323.

Matsuura M. (2013). Structural modifications of bacterial lipopolysaccharide that facilitate gram-negative bacteria evasion of host innate immunity. *Front Immunol* 4**,** 109.

Matsuyama H., Kamesaki T., Sasaki R., Minami H., and Yumoto I. (2003). Production of two types of exopolysaccharide by Novosphingobium rosa. *J Biosci Bioeng* 95**,** 152-156.

Maunders E., and Welch M. (2017). Matrix exopolysaccharides; the sticky side of biofilm formation. *FEMS Microbiol Lett* 364**,** fnx120.

Mazor G., Kidron G.J., Vonshak A., and Abeliovich A. (1996). The role of cyanobacterial exopolysaccharides in structuring desert microbial crusts. *FEMS Microbiol Ecol* 21**,** 121-130.

Mcbride M.J. (2014). "The family Flavobacteriaceae" in *The prokaryotes: other major lineages of bacteria and the archaea,* ed. E. Rosenberg. (Heidelberg: Springer), 643-676.

Menzel P., Ng K.L., and Krogh A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7**,** 11257.

Mikha M.M., Vigil M.F., and Benjamin J.G. (2013). Long-term tillage impacts on soil aggregation and carbon dynamics under wheat-fallow in the central Great Plains. *Soil Sci Soc Am J* 77**,** 594-605.

Mistry J., Finn R.D., Eddy S.R., Bateman A., and Punta M. (2013). Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res* 41**,** e121.

Mogul R., Vaishampayan P., Bashir M., Mckay C.P., Schubert K., Bornaccorsi R., Gomez E., Tharayil S., Payton G., and Capra J. (2017). Microbial community and biochemical dynamics of biological soil

crusts across a gradient of surface coverage in the central Mojave desert. *Front Microbiol* 8**,** 1974.
Montanarella L. (2016). "The importance of land restoration for achieving a land degradation–neutral world" in *Land restoration,* eds. I. Chabay, M. Frick & J. Helgeson. (London: Elsevier), 249-258.
Montgomery D.R. (2007). Soil erosion and agricultural sustainability. *P Natl Acad Sci USA* 104**,** 13268-13272.
Mugnai G., Rossi F., Felde V.J.M.N.L., Colesie C., Büdel B., Peth S., Kaplan A., and De Philippis R. (2018). The potential of the cyanobacterium Leptolyngbya ohadii as inoculum for stabilizing bare sandy substrates. *Soil Biol Biochem* 127**,** 318-328.
Mundinger A.B., Lawson C.E., Jetten M.S., Koch H., and Lücker S. (2019). Cultivation and transcriptional analysis of a canonical Nitrospira under stable growth conditions. *Front Microbiol* 10**,** 1325.
Murer E.J., Baumgarten A., Eder G., Gerzabek M.H., Kandeler E., and Rampazzo N. (1993). "An improved sieving machine for estimation of soil aggregate stability (SAS)" in *Soil structure/soil biota interrelationships,* eds. L. Brussaard, M.J. Kooistra. (Amsterdam: Elsevier), 539-547.
Nayfach S., and Pollard K.S. (2016). Toward accurate and quantitative comparative metagenomics. *Cell* 166**,** 1103-1116.
Nwodo U.U., Green E., and Okoh A.I. (2012). Bacterial exopolysaccharides: functionality and prospects. *Int J Mol Sci* 13**,** 14002-14015.
Oades J., and Waters A. (1991). Aggregate hierarchy in soils. *Soil Res* 29**,** 815-828.
Oh S., Choi D., and Cha C.-J. (2019). Ecological processes underpinning microbial community structure during exposure to subinhibitory level of triclosan. *Sci Rep* 9**,** 1-12.
Oksanen J., Blanchet F.G., Friendly M., Kindt R., Legendre P., Mcglinn D., Minchin P.R., O'Hara R.B., Simpson G.L., Solymos P., Stevens M.H.H., Szoecs E., and Wagner H. (2018). "vegan: community ecology package". R package version 2.5-1.
Otero A., and Vincenzini M. (2004). Nostoc (Cyanophyceae) goes nude: extracellular polysaccharides serve as a sink for reducing power under unbalanced C/N metabolism. *J Phycol* 40**,** 74-81.
Owens T.W., Taylor R.J., Pahil K.S., Bertani B.R., Ruiz N., Kruse A.C., and Kahne D. (2019). Structural basis of unidirectional export of lipopolysaccharide to the cell surface. *Nature* 567**,** 550.
Pando J.M., Karlinsey J.E., Lara J.C., Libby S.J., and Fang F.C. (2017). The Rcs-regulated colanic acid capsule maintains membrane potential in Salmonella enterica serovar Typhimurium. *MBio* 8**,** e00808-00817.
Papo N., and Shai Y. (2005). A molecular mechanism for lipopolysaccharide protection of Gram-negative bacteria from antimicrobial peptides. *J Biol Chem* 280**,** 10378-10387.
Paradis E., Claude J., and Strimmer K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20**,** 289-290.
Parker J., Smith G., Fredrickson H., Vestal J., and White D. (1982). Sensitive assay, based on hydroxy fatty acids from lipopolysaccharide lipid A, for Gram-negative bacteria in sediments. *Appl Environ Microbiol* 44**,** 1170-1177.
Pereira S.B., Mota R., Santos C.L., De Philippis R., and Tamagnini P. (2013). Assembly and export of extracellular polymeric substances (EPS) in cyanobacteria: a phylogenomic approach. *Adv Bot Res* 65**,** 235-279.
Pereira S.B., Mota R., Vieira C.P., Vieira J., and Tamagnini P. (2015). Phylum-wide analysis of genes/proteins related to the last steps of assembly and export of extracellular polymeric substances (EPS) in cyanobacteria. *Sci Rep-UK* 5**,** 14835.
Pérez-Burgos M., García-Romero I., Jung J., Valvano M.A., Søgaard-Andersen L. (2019). Identification of the lipopolysaccharide O-antigen biosynthesis priming enzyme and the O-antigen ligase in Myxococcus xanthus: critical role of LPS O-antigen in motility and development. *Mol Microbiol* 112, 1178-1198.

Pinheiro J., Bates D., Debroy S., and Sarkar D. (2018). "nlme: Linear and nonlinear mixed effects models". R package version 3.1-131.1.
Poppleton D.I., Duchateau M., Hourdel V., Matondo M., Flechsler J., Klingl A., Beloin C., and Gribaldo S. (2017). Outer membrane proteome of Veillonella parvula: a diderm firmicute of the human microbiome. *Front Microbiol* 8**,** 1215.
Preston G.M. (2004). Plant perceptions of plant growth-promoting Pseudomonas. *Philos T Roy Soc A* 359**,** 907-918.
Prieto-Barajas C.M., Valencia-Cantero E., and Santoyo G. (2017). Microbial mat ecosystems: Structure types, functional diversity, and biotechnological application. *Electron J Biotechn* 31**,** 48-56.
Putker F., Bos M.P., and Tommassen J. (2015). Transport of lipopolysaccharide to the Gram-negative bacterial cell surface. *FEMS Microbiol Rev* 39**,** 985-1002.
Quelas J.I., Mongiardini E.J., Casabuono A., López-García S.L., Althabegoiti M.J., Covelli J.M., Pérez-Giménez J., Couto A., and Lodeiro A.R. (2010). Lack of galactose or galacturonic acid in Bradyrhizobium japonicum USDA 110 exopolysaccharide leads to different symbiotic responses in soybean. *Mol Plant Microbe Int* 23**,** 1592-1604.
Quintero M., and Comerford N.B. (2013). Effects of conservation tillage on total and aggregated soil organic carbon in the Andes. *Open J Soil Sci* 3**,** 361-373.
Rampadarath S., Bandhoa K., Puchooa D., Jeewon R., and Bal S. (2017). Early bacterial biofilm colonizers in the coastal waters of Mauritius. *Electron J Biotechn* 29**,** 13-21.
Redmile-Gordon M., Brookes P., Evershed R., Goulding K., and Hirsch P. (2014). Measuring the soil-microbial interface: extraction of extracellular polymeric substances (EPS) from soil biofilms. *Soil Biol Biochem* 72**,** 163-171.
Redmile-Gordon M., Gregory A., White R., and Watts C. (2020). Soil organic carbon, extracellular polymeric substances (EPS), and soil structural stability as affected by previous and current land-use. *Geoderma* 363**,** 114143.
Reed H.E., and Martiny J.B. (2013). Microbial composition affects the functioning of estuarine sediments. *ISME J* 7**,** 868.
Reeve W., Terpolilli J., Melino V., Ardley J., Tian R., De Meyer S., Tiwari R., Yates R., O'Hara G., Howieson J., Ninawi M., Teshima H., Bruce D., Detter C., Tapia R., Han, C., Wei C., Huntemann M., Han J., Chen I., Mavrommatis K., Markowitz V., Ivanova N., Ovchinnikova G., Pagani I., Pati A., Goodwin L., Peters L., Woyke T., and Kyrpides N. (2013). Genome sequence of the lupin-nodulating Bradyrhizobium sp. strain WSM1417. *Stand Genomic Sci* 9**,** 273-282.
Rendueles O., Garcia-Garcerà M., Néron B., Touchon M., and Rocha E.P. (2017). Abundance and co-occurrence of extracellular capsules increase environmental breadth: Implications for the emergence of pathogens. *PLoS Pathog* 13**,** e1006525.
Rho M., Tang H., and Ye Y. (2010). FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res* 38**,** e191-e191.
Rillig M.C., Wright S.F., and Eviner V.T. (2002). The role of arbuscular mycorrhizal fungi and glomalin in soil aggregation: comparing effects of five plant species. *Plant Soil* 238**,** 325-333.
Roca C., Alves V.D., Freitas F., and Reis M.A. (2015). Exopolysaccharides enriched in rare sugars: bacterial sources, production, and applications. *Front Microbiol* 6**,** 288.
Rodriguez-R L.M., and Konstantinidis K.T. (2014). Estimating coverage in metagenomic data sets and why it matters. *ISME J* 8**,** 2349-2351.
Rosenberg E. (2014). "The family Chitinophagaceae" in *The prokaryotes: other major lineages of bacteria and the archaea,* ed. E. Rosenberg. (Berlin, Heidelberg: Springer), 493-496.
Rosenberg E., Delong E.F., Lory S., Stackebrandt E., and Thompson F. (2014). *The Prokaryotes: Actinobacteria.* Heidelberg: Springer.

Rossi F., and De Philippis R. (2015). Role of cyanobacterial exopolysaccharides in phototrophic biofilms and in complex microbial mats. *Life* 5**,** 1218-1238.
Rossi F., Li H., Liu Y., and De Philippis R. (2017). Cyanobacterial inoculation (cyanobacterisation): perspectives for the development of a standardized multifunctional technology for soil fertilization and desertification reversal. *Earth-Sci Rev* 171**,** 28-43.
Rossi F., Mugnai G., and De Philippis R. (2018). Complex role of the polymeric matrix in biological soil crusts. *Plant Soil* 429**,** 19-34.
Ruiz N., Gronenberg L.S., Kahne D., and Silhavy T.J. (2008). Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of Escherichia coli. *P Natl Acad Sci USA* 105**,** 5537-5542.
Sancho L.G., Maestre F.T., and Büdel B. (2014). Biological soil crusts in a changing world: introduction to the special issue. *Biodivers Conserv* 23**,** 1611–1617.
Sande C., Wear S., and Whitfield C. (2019). "Capsules and secreted extracellular polysaccharides" in Encyclopedia of microbiology, ed. T.M. Schmidt. (New York: Academic Press), 604-612.
Santofimia E., González-Toril E., López-Pamo E., Gomariz M., Amils R., and Aguilera Á. (2013). Microbial diversity and its relationship to physicochemical characteristics of the water in two extreme acidic pit lakes from the Iberian Pyrite Belt (SW Spain). *PLoS One* 8**,** e66746.
Sartori M., Philippidis G., Ferrari E., Borrelli P., Lugato E., Montanarella, L., and Panagos P. (2019). A linkage between the biophysical and the economic: Assessing the global market impacts of soil erosion. *Land Use Policy* 86**,** 299-312.
Sayers E., Agarwala R., Bolton E., Brister J., Canese K., Clark K., Connor R., Fiorini N., Funk K., Hefferon T., Holmes J., Kim S., Kimchi A., Kitts P., Lathrop S., Lu Z., Madden T., Marchler-Bauer A., Phan L., Schneider V., Schoch C., Pruitt K., and Ostell J. (2019). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 47**,** D23-D28.
Schmid J. (2018). Recent insights in microbial exopolysaccharide biosynthesis and engineering strategies. *Curr Opin Biotechnol* 53**,** 130-136.
Schmid J., Sieber V., and Rehm B. (2015). Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. *Front Microbiol* 6**,** 496.
Schmieder R., and Edwards R. (2011a). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PloS One* 6**,** e17288.
Schmieder R., and Edwards R. (2011b). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27**,** 863-864.
Schubert M., Lindgreen S., and Orlando L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes* 9**,** 88.
Selim M.S., Amer S.K., Mohamed S.S., Mounier M.M., and Rifaat H.M. (2018). Production and characterisation of exopolysaccharide from Streptomyces carpaticus isolated from marine sediments in Egypt and its effect on breast and colon cell lines. *J Genet Eng Biotechnol* 16**,** 23-28.
Seneviratne G., and Jayasinghearachchi H. (2005). A rhizobial biofilm with nitrogenase activity alters nutrient availability in a soil. *Soil Biol Biochem* 37**,** 1975-1978.
Sengupta D., Datta S., and Biswas D. (2018). Towards a better production of bacterial exopolysaccharides by controlling genetic as well as physico-chemical parameters. *Appl Microbiol Biot* 102**,** 1587-1598.
Sheehy J., Regina K., Alakukku L., and Six J. (2015). Impact of no-till and reduced tillage on aggregation and aggregate-associated carbon in Northern European agroecosystems. *Soil Till Res* 150**,** 107-113.
Silipo A., and Molinaro A. (2010). "The diversity of the core oligosaccharide in lipopolysaccharides" in *Endotoxins: structure, function and recognition,* eds. X. Wang & P.J. Quinn. (Dordrecht: Springer), 69-99.

Singh P., Heikkinen J., Mitra S., and Regina K. (2016). "Impact of tillage and residue management practices on soil aggregates and soil organic carbon" in *Geostatistical and geospatial approaches for the characterization of natural resources in the environment,* ed. N.J. Raju. (Cham: Springer), 381-384.

Six J., Bossuyt H., Degryze S., and Denef K. (2004). A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. *Soil Till Res* 79**,** 7-31.

Skorupska A., Janczarek M., Marczak M., Mazur A., and Król J. (2006). Rhizobial exopolysaccharides: genetic control and symbiotic functions. *Microb Cell Fact* 5**,** 7.

Smith S., Abed R.M.M., and Gercia-Pichel F. (2004). Biological soil crusts of sand dunes in Cape Cod National Seashore, Massachusetts, USA. *Microb Ecol* 48**,** 200-208.

Steimle A., Autenrieth I.B., and Frick J.-S. (2016). Structure and function: Lipid A modifications in commensals and pathogens. *Int J Med Microbiol* 306**,** 290-301.

Stockfisch N., Forstreuter T., and Ehlers W. (1999). Ploughing effects on soil organic matter after twenty years of conservation tillage in Lower Saxony, Germany. *Soil Till Res* 52, 91–101.

Strickland M.S., Lauber C., Fierer N., and Bradford M.A. (2009). Testing the functional significance of microbial community composition. *Ecology* 90**,** 441-451.

Stromberg L.R., Mendez H.M., and Mukundan H. (2017). "Detection methods for lipopolysaccharides: past and present" in *Escherichia coli - recent advances on physiology, pathogenesis and biotechnological applications,* ed. A. Samie. (IntechOpen).

Subramaniam R., Thirumal V., Chistoserdov A., Bajpai R.,Bader J., and Popovic M. (2018). High-density cultivation in the production of microbial products. *Chem Biochem Eng Q* 32, 451-464.

Suresh Kumar A., Mody K., and Jha B. (2007). Bacterial exopolysaccharides–a perception. *J Basic Microb* 47**,** 103-117.

Sutcliffe I.C. (2010). A phylum level perspective on bacterial cell envelope architecture. *Trends Microbiol* 18**,** 464-470.

Sutherland I.W. (1972). "Bacterial exopolysaccharides" in *Advances in microbial physiology,* ed. A. Rose. (New York: Academic Press), 143-213.

Sutherland I.W., and Thomson, S. (1975). Comparison of polysaccharides produced by Myxococcus strains. *Microbiology* 89**,** 124-132.

Sutherland I.W. (2001a). Biofilm exopolysaccharides: a strong and sticky framework. *Microbiology* 147**,** 3-9.

Sutherland, I.W. (2001b). Microbial polysaccharides from Gram-negative bacteria. *Int Dairy J* 11**,** 663-674.

R Core Team. (2016). "R: a language and environment for statistical computing". (Vienna: R Foundation for Statistical Computing).

Teixeira L.M., and Merquior V.L.C. (2014). "The family Moraxellaceae" in *The prokaryotes: Gammaproteobacteria,* ed. E. Rosenberg. (Heidelberg: Springer), 443-476.

Thomas T., Gilbert J., and Meyer F. (2012). Metagenomics - a guide from sampling to data analysis. *Microb Inform Exp* 2**,** 3.

Tisdall J.M., and Oades J. (1982). Organic matter and water-stable aggregates in soils. *J Soil Sci* 33**,** 141-163.

Torri D., Ciampalini R., and Gil P.A. (1998). "The role of soil aggregates in soil erosion processes" in *Modelling soil eosion by water,* eds. J. Boardman & D. Favis-Mortlock. (Berlin: Springer), 247-257.

Totsche K.U., Amelung W., Gerzabek M.H., Guggenberger G., Klump, E., Knief C., Lehndorff E., Mikutta R., Peth S., and Prechtel A. (2018). Microaggregates in soils. *J Plant Nutr Soil Sc* 181**,** 104-136.

Townsend T.J., Ramsden S.J., and Wilson P. (2016). How do we cultivate in England? Tillage practices in crop production systems. *Soil Use Manage* 32**,** 106-117.

Tytgat H.L., and Lebeer S. (2014). The sweet tooth of bacteria: common themes in bacterial glycoconjugates. *Microbiol Mol Biol Rev* 78**,** 372-417.
Uksa M., Schloter M., Endesfelder D., Kublik S., Engel M., Kautz T., Köpke U., and Fischer D. (2015). Prokaryotes in subsoil—evidence for a strong spatial separation of different phyla by analysing co-occurrence networks. *Front Microbiol* 6**,** 1269.
Vance E.D., Brookes P.C., and Jenkinson D.S. (1987). An extraction method for measuring soil microbial biomass C. *Soil Biol Biochem* 19**,** 703-707.
Veste M. (2005). The importance of biological soil crusts for rehabilitation of degraded arid and semi-arid ecosystems. *Sci Soil Water Conserv* 3**,** 42-47.
Vital M., Chai B., Østman B., Cole J., Konstantinidis K.T., and Tiedje J.M. (2015). Gene expression analysis of E. coli strains provides insights into the role of gene regulation in diversification. *ISME J* 9**,** 1130-1140.
Vlček V., and Pohanka M. (2019). Glomalin–an interesting protein part of the soil organic matter. *Soil Water Res*.
Vogel H.-J., Weller U., and Schlüter S. (2010). Quantification of soil structure based on Minkowski functions. *Comput Geosci* 36**,** 1236-1245.
Vogeleer P., Tremblay Y.D., Mafu A.A., Jacques M., and Harel J. (2014). Life on the outside: role of biofilms in environmental persistence of Shiga-toxin producing Escherichia coli. *Front Microbiol* 5**,** 317.
Voss B.J., and Trent M.S. (2018). LPS transport: flipping out over MsbA. *Curr Biol* 28**,** R30-R33.
Wagg C., Bender S.F., Widmer F., and Van Der Heijden M.G. (2014). Soil biodiversity and soil community composition determine ecosystem multifunctionality. *P Natl Acad Sci USA* 111**,** 5266-5270.
Wang C., Zhang R., Liu B.-T., Liu C.-L., and Du Z.-J. (2019). Paracnuella aquatica gen. nov., sp. nov., a member of the family Chitinophagaceae isolated from a hot spring. *Int J Syst Evol Micr* 69**,** 2360-2366.
Wang W., Zhong Z., Wang Q., Wang H., Fu Y., and He X. (2017). Glomalin contributed more to carbon, nutrients in deeper soils, and differently associated with climates and soil properties in vertical profiles. *Sci Rep* 7**,** 1-13.
Wang X., Ribeiro A.A., Guan Z., Mcgrath S.C., Cotter R.J., and Raetz C.R. (2006). Structure and biosynthesis of free lipid A molecules that replace lipopolysaccharide in Francisella tularensis subsp. novicida. *Biochemistry* 45**,** 14427-14440.
Wang Y., Xu J., Shen J., Luo Y., Scheu S., and Ke X. (2010). Tillage, residue burning and crop rotation alter soil fungal community and water-stable aggregation in arable fields. *Soil Tillage Res* 107**,** 71-79.
Warnes G.R., Bolker B., Bonebakker L., Gentleman R., Liaw W.H.A., Lumley T., Maechler M., Magnusson A., Moeller S., Schwartz M., and Venables B. (2015). "gplots: various R Programming Tools for Plotting Data". R package version 2.17.0.
Weber B., Büdel B., and Belnap J. (2016). *Biological soil crusts: an organizing principle in drylands.* New York: Springer.
Weil R.R., and Brady N.C. (2017). *The nature and properties of soils.* Harlow: Pearson Education.
Whitfield C. (2010). Polymerases: glycan chain-length control. *Nat Chem Biol* 6**,** 403-404.
Whitfield C., and Larue K. (2008). Stop and go: regulation of chain length in the biosynthesis of bacterial polysaccharides. *Nat Struct Mol Biol* 15**,** 121-123.
Whitfield C., and Trent M.S. (2014). Biosynthesis and export of bacterial lipopolysaccharides. *Annu Rev Biochem* 83**,** 99-128.
Whitney J., and Howell P. (2013). Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria. *Trends Microbiol* 21**,** 63-72.
Wickham H. (2009). *ggplot2: elegant graphics for data analysis.* Springer.

Wilcox R.R., and Schönbrodt F.D. (2014). "The WRS package for robust statistics in R". R package version 0.24.

Wilhelm R.C., Niederberger T.D., Greer C., and Whyte L.G. (2011). Microbial diversity of active layer and permafrost in an acidic wetland from the Canadian High Arctic. *Can J Microbiol* 57**,** 303-315.

Willems A. (2014). "The family Comamonadaceae" in *The prokaryotes: Alphaproteobacteria and Betaproteobacteria,* ed. E. Rosenberg. (Heidelberg: Springer), 777-851.

Williams L., Borchhardt N., Colesie C., Baum C., Komsic-Buchmann K., Rippin M., Becker B., Karsten U., and Büdel B. (2017). Biological soil crusts of Arctic Svalbard and of Livingston Island, Antarctica. *Polar Biol* 40**,** 399-411.

Wright S., Starr J., and Paltineanu I. (1999). Changes in aggregate stability and concentration of glomalin during tillage management transition. *Soil Sci Soc Am J* 63**,** 1825-1829.

Wu Q., Tun H.M., Leung F.C.-C., and Shah N.P. (2014). Genomic insights into high exopolysaccharide-producing dairy starter bacterium Streptococcus thermophilus ASCC 1275. *Sci Rep* 4**,** 1-8.

Yolcubal I., Brusseau M.L., Artiola J.F., Wierenga P.J., and Wilson L.G. (2004). "Environmental physical properties and processes" in *Environmental monitoring and characterization,* eds. J.F. Artiola, I.L. Pepper & M.L. Brusseau. (Amsterdam: Elsevier), 207-239.

Yu J., Liberton M., Cliften P.F., Head R.D., Jacobs J.M., Smith R.D., Koppenaal D.W., Brand J.J., and Pakrasi H.B. (2015). Synechococcus elongatus UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO_2. *Sci Rep* 5, 8132.

Zeidan A.A., Poulsen V.K., Janzen T., Buldo P., Derkx P.M., Øregaard G., and Neves A.R. (2017). Polysaccharide production by lactic acid bacteria: from genes to industrial applications. *FEMS Microbiol Rev* 41**,** S168-S200.

Zhang G., Meredith T.C., and Kahne D. (2013). On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Curr Opin Microbiol* 16**,** 779-785.

Zhang J., and Poh C.L. (2018). Regulating exopolysaccharide gene wcaF allows control of Escherichia coli biofilm formation. *Sci Rep* 8**,** 13127.

Zhang Z., Qu Y., Li S., Feng K., Wang S., Cai,W., Liang Y., Li H., Xu M., and Yin H. (2017). Soil bacterial quantification approaches coupling with relative abundances reflecting the changes of taxa. *Sci Rep* 7**,** 4837.

List of abbreviations

ANOVA	analysis of variance
c-di-GMP	bis-(3’-5’)-cyclic dimeric guanosine monophosphate
Cmic	microbial biomass carbon
COG	Clusters of Orthologous Groups
CT	conventional tillage
DOC	dissolved organic carbon
DON	dissolved organic nitrogen
dTDP	deoxythymidine diphosphate
eggNOG	evolutionary genealogy of genes: Non-supervised Orthologous Groups
EPS	exopolysaccharide
EPS-DB	Bacterial Exopolysaccharide Properties and Structures Database
ETA	ethanolamine
GDP	guanosine diphosphate
GT	glycosyltransferase
HMM	Hidden Markov Model
Kdo	poly-2-keto-3-deoxyoctulosonic acid
KEGG	Kyoto Encyclopedia of Genes and Genomes
LPS	lipopolysaccharide
N	ambiguous base
NCBI	National Center for Biotechnology Information
NMDS	non-metric multidimensional scaling
Nmic	microbial biomass nitrogen
nr	non-redundant
OPX	outer membrane polysaccharide export
OS	oligosaccharide

P	phosphate
P1-4	publication 1-4
PCA	principal component analysis
PCoA	principal component analysis
PCP	polysaccharide copolymerase
RT	reduced tillage
SAF	stable aggregate fraction
SGS	shotgun sequencing
SOC	soil organic carbon
SRA	Sequence Read Archive
TPR	tetratricopeptide repeat
UDP	uridine diphosphate
Und	undecaprenyl
USDA	U.S. Department of Agriculture
XCMT	X-ray computed microtomography

Acknowledgments

Completing this PhD study has been a truly life-changing experience for me and would not have been possible without the efforts and support of several people to whom I am extremely grateful. First and foremost, I would like to address my supervisors Prof. Michael Schloter and Dr. Stefanie Schulz. Thank you for your patience, motivation and immense knowledge. It was a privilege to work with you both.

Michael, thank you for accepting me into your group and giving me many opportunities for which I will be ever grateful. Your constructive critique helped me to do better, and your support for publishing my work defined the success of this thesis.

Steffi, I am sincerely grateful for all the help and guidance that I have received from you for the past five years. It was a long and, at times, rough journey, but I would not have been able to get this far without you. Your door was always open to discuss any problem, and you also gave me space to work things out at my own speed when I needed it.

Many thanks to my co-authors for their contribution to the studies presented in this thesis. I am profoundly grateful for their hard work and insightful feedback.

I gratefully acknowledge the funding received from the Federal Office for Agriculture and Food. I would also like to express my appreciation to Core Organic Plus Funding Bodies being partners of the EP7 ERA-Net project CORE Organic Plus for funding the project “Fertility Building Management Measures in Organic Cropping Systems” – FertilCrop. Participating in this project allowed me to meet many wonderful people and grow not only as a scientist but also as a person.

Pragnę podziękować moim rodzicom za to, że zawsze we mnie wierzyli i wspierali w pogoni za marzeniami. Mamuś, Tatuś, powinniście być z siebie dumni, bo bez Was ta praca doktorska nigdy by nie powstała. Dziękuję Wam za wszystko i kocham Was z całego serca.

Special thanks to the person who holds a special place in my heart. Giaco, you being by my side means more to me than you can imagine. Thank you for being you and for making me happy.

Last but not least, I would like to say heartfelt thank you to all the other people who supported me during this challenging period. You are too many to mention by name, but I am indebted to all of you.

Appendices

A Publication 1

Microbial Ecology
https://doi.org/10.1007/s00248-019-01415-6

ENVIRONMENTAL MICROBIOLOGY

Biological Soil Crusts from Different Soil Substrates Harbor Distinct Bacterial Groups with the Potential to Produce Exopolysaccharides and Lipopolysaccharides

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

Received: 20 March 2019 / Accepted: 16 July 2019

Abstract

Biological soil crusts (biocrusts) play an important role in improving soil stability and resistance to erosion by promoting aggregation of soil particles. During initial development, biocrusts are dominated by bacteria. Some bacterial members of the biocrusts can contribute to the formation of soil aggregates by producing exopolysaccharides and lipopolysaccharides that act as "glue" for soil particles. However, little is known about the dynamics of "soil glue" producers during the initial development of biocrusts. We hypothesized that different types of initial biocrusts harbor distinct producers of adhesive polysaccharides. To investigate this, we performed a microcosm experiment, cultivating biocrusts on two soil substrates. High-throughput shotgun sequencing was used to obtain metagenomic information on microbiomes of bulk soils from the beginning of the experiment, and biocrusts sampled after 4 and 10 months of incubation. We discovered that the relative abundance of genes involved in the biosynthesis of exopolysaccharides and lipopolysaccharides increased in biocrusts compared with bulk soils. At the same time, communities of potential "soil glue" producers that were highly similar in bulk soils underwent differentiation once biocrusts started to develop. In the bulk soils, the investigated genes were harbored mainly by *Betaproteobacteria*, whereas in the biocrusts, the major potential producers of adhesive polysaccharides were, aside from *Alphaproteobacteria*, either *Cyanobacteria* or *Chloroflexi* and *Acidobacteria*. Overall, our results indicate that the potential to form exopolysaccharides and lipopolysaccharides is an important bacterial trait for initial biocrusts and is maintained despite the shifts in bacterial community composition during biocrust development.

Keywords Biological soil crusts · Exopolysaccharides · Lipopolysaccharides · Microbiome · Metagenomics

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00248-019-01415-6) contains supplementary material, which is available to authorized users.

✉ Stefanie Schulz
stefanie.schulz@helmholtz-muenchen.de

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

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

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

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

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

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

Published online: 01 August 2019

Introduction

Biological soil crusts (biocrusts) are important biotic components of many terrestrial ecosystems [1, 2]. They consist of highly specialized and complex communities of algae, mosses, lichens, fungi, cyanobacteria, and other prokaryotes [3]. These organisms live in a close association with soil particles, forming a coherent layer within the uppermost few millimeters of the topsoil, or directly on the soil surface [1]. An important structural element of biocrusts is the extracellular polymeric matrix (EPM) which is composed mostly of polysaccharides and connects organisms and soil particles [4]. EPM ensures biocrust integrity, provides protection from external harmful agents, and alters moisture content as well as nutrient availability [4]. EPM also fosters the stabilization of soil aggregates and protects soils from erosion by wind or water [5–8]. Among organisms forming biocrusts, the best-studied producers of polysaccharides are cyanobacteria and algae [1]. However, although not as thoroughly studied, also non-photosynthetic microbial members of biocrusts, including fungi, proteobacteria, and actinobacteria, are prominent producers of these compounds [9–11].

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

In this respect, bacterial polysaccharides, specifically exopolysaccharides (EPSs) and lipopolysaccharides (LPSs), are of great interest, as cyanobacteria and non-phototrophic bacteria form biocrusts in the initial stage of biocrust development [25]. EPSs are either synthesized intracellularly and excreted by one of three different pathways: the Wzy-dependent pathway, the ABC transporter-dependent pathway, and the synthase-dependent pathway, or synthesized extracellularly [26, 27]. In contrast, parts of LPSs are initially synthesized inside a cell, then ligated together at the inner membrane and transported to the cell surface as mature molecules [28, 29]. While LPSs are present in most Gram-negative bacteria [30], EPSs are exuded by a wide range of taxa [16]. Among the most-recognized producers of EPSs are cyanobacterial members of *Oscillatoria*, *Nostoc*, *Lyngbya*, and *Schizothrix*, as well as bacterial members of *Microbacterium*, *Pseudomonas*, *Bacillus*, *Paenibacillus*, and *Streptomyces* [31]. These microorganisms are the first colonizers of bare soils, and their EPSs as well as LPSs are considered as essential for the initial consolidation of soil particles and the preparation of conditions for the establishment of cryptogamic surface cover in the later stages of biocrust development [32]. Thus, a better understanding of the dynamics of polysaccharide-producing organisms during the initial development of biocrusts requires more in-depth knowledge on cyanobacteria and other bacteria that initialize biocrust establishment [4].

Many researchers studied polysaccharide-producing bacterial strains that were isolated from biocrusts at different stages of development [33–36]. However, data on the community dynamics of bacterial EPS and LPS producers under natural conditions is missing. Thus, our aim was to investigate polysaccharide-producing bacterial communities during the initial stage of biocrust development. We assumed that the relative abundance of genes related to EPS and LPS formation would increase once biocrust development starts. Moreover, we hypothesized that different types of initial biocrusts would harbor different communities of potential EPS and LPS producers. To test our hypotheses, we performed a microcosm experiment cultivating biocrusts on two different soil substrates. As the soil substrates came from sites with different types of naturally occurring biocrusts, we expected that the biocrusts cultivated in the microcosm experiment would also consist of distinct microbial communities. To address our research questions, we used a high-throughput shotgun sequencing of DNA extracted from bulk soils from the beginning of the experiment, as well as initial biocrusts sampled after 4 and 10 months of incubation. We employed a bioinformatics pipeline described by Cania et al. [37], which targets genes specific for EPS and LPS production, to obtain information on bacteria potentially involved in the production of adhesive polysaccharides.

Materials and Methods

Site Description

Soils for the incubation experiment were collected in 2011 from two sites at the initial stages of ecosystem development. One was the artificial catchment Chicken Creek (51°36′18″ N, 14°15′58″ E) and the other was an initial moving sand dune close to Lieberose (51°55′49″ N, 14°22′22″ E). Both sites are located in the state of Brandenburg in eastern Germany,

approximately 37 km apart. The climate of the region is temperate continental with a mean air temperature of 8.9 °C and mean rainfall of 569 mm a^{-1} [19].

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

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

Sampling and Incubation Experiment

Bulk soil from the Chicken Creek catchment and the Lieberose sand dune was used to establish a microcosm experiment. A total amount of 100 kg of soil was taken per site from the top 20 cm. At Lieberose, the soil was collected from five spots on the top of the dune, where no plants were growing, while at Chicken Creek, five plant-free spots were used for the soil sampling. The soil was transported and afterwards stored in the dark at room temperature for approximately 6 months before the incubation experiment. During that time, pre-experiments to adjust the incubation conditions for biocrust growth were performed.

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

Biocrusts were sampled after 4 (T1) and 10 (T2) months of incubation from three independent pots per soil substrate and sampling time point. Only the upper 2 mm was considered as biocrust. In addition, samples of bulk soil without biocrust development were taken at the beginning of the experiment (T0). In total, 18 samples were collected (3 sampling times × 2 sites × 3 replicates). Samples for DNA analyses were directly

frozen at − 80 °C, while samples for biochemical analyses were stored at 4 °C until further processing. For the determination of water repellency and the computed tomography analysis, undisturbed samples from the end of the experiment were taken using Petri dishes and stored at 4 °C until further analysis.

Physicochemical Measurements

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

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

DNA Extraction, Library Preparation, and Sequencing

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

Bioinformatical Analysis of Sequencing Data

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

Metagenomes obtained from bulk soils (T0) comprised reads on average 106 bp shorter than metagenomes created from biocrusts (T1 and T2). To test whether the difference in read length affects the accuracy of annotations, T1 and T2 reads were trimmed in silico in a randomized manner to resemble the length distributions of T0 reads. A comparison of the length distributions of exemplary "short" and "long reads" metagenomes, before and after trimming, is presented in Supplementary Material 1 Fig. S2. The metagenomes with

trimmed sequences were analyzed taxonomically together with the original metagenomes. Principal coordinate analysis (PCoA) ordination plots (Supplementary Material 1 Fig. S3) showed that the taxonomic annotations were not notably biased by the difference in read length. Consequently, further analyses were performed on the metagenomes with original read lengths.

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

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

Statistical Analysis and Data Visualization

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

Statistical analyses and data visualization were conducted using R (version 3.4.4) [66]. Effects of soil substrate, incubation time, and their possible interaction were determined according to Field et al. [67]. Briefly, significant differences were detected by a robust 2-way independent analysis of variance (ANOVA) based on the median as M-estimator, with 2000 bootstrap samples. For this purpose, the pbad2way function from the WRS package [68] was used. The influence was counted as significant if the p value was below 5% ($p < 0.05$). The Benjamini-Hochberg procedure was used to control the false discovery rate in data derived from the metagenomic datasets. Omega squared (ω^2) was calculated as an effect size to estimate the magnitude of observed influences of the analyzed factors. It can be interpreted as the percentage of variation in the dependent variable explained by the independent variable [69].

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

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

Results

Initial Soil Substrate Parameters

Bulk soils collected from both sites had similar low contents of DOC and DON. DOC values were in the range of 4.57 ± 1.67 μg/g in samples collected from Chicken Creek, and 6.63 ± 0.46 μg/g in those from Lieberose, while DON was below detection limit in samples from both sites. Conversely, pH values differed between soils from both locations. Soil from Chicken Creek was slightly alkaline (7.31 ± 0.30), whereas soil from Lieberose was rather acidic (pH = 5.42 ± 0.39). Initial soil substrate parameters are presented in Table 2.

Biocrust Development

Biocrusts that developed in the microcosm experiment were in the initial stage of development. They consisted mostly of bacterial and algal biofilms, which enmeshed soil particles and formed patches on the soil surface. Mosses were also observed, but they did not form a dense surface cover yet. For biocrusts developed on the Chicken Creek soil, mosses and algae were already visible after the first 4 months of

Table 1 Proteins related to exo- and lipopolysaccharide production with corresponding KO numbers, HMM IDs, and genes

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

incubation (T1). For the Lieberose soil, mostly biofilms around single soil particles were visible at T1, whereas mosses and distinct biocrust structures appeared after 10 months of incubation (T2). Representative pictures are presented in Supplementary Material 1 Fig. S4.

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

The exemplary CT images (Supplementary Material 1 Fig. S5A-D) of biocrusts from T2 grown on soil from Chicken Creek showed a layer of smaller particles in the crust horizon compared to the underlying soil substrate. Positive Euler numbers (Supplementary Material 1 Fig. S5E) for both biocrusts and the underlying soil indicate more isolated pores than connections in the pore network. The connectivity of the pore space was lower for the biocrusts, especially when small pores (46 μm) were considered (Euler number of 8.6 mm^{-3}). In the underlying soil, the connectivity was the lowest for pore size class of 107 μm (Euler number of 4.6 mm^{-3}). The connectivity then increased towards larger pore sizes as indicated by decreasing Euler numbers.

Major Characteristics of the Shotgun Sequencing Libraries

Shotgun sequencing of 18 libraries made from bulk soils from the beginning of the experiment (T0) and biocrusts from the 4-month (T1) and 10-month (T2) samplings generated 18.3 Gbases of data in total. This corresponded to 59,710,640 filtered reads. The number of filtered reads per sample varied between 2.1 and 5.3 million. Mean lengths of sequences after trimming ranged from 120 to 250 bp. Details of the raw and filtered sequencing data are summarized in Supplementary Material 2 Table S1.

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

Taxonomic Analysis

42.83% of all metagenomic reads were assigned to *Bacteria*, which could be further differentiated into 366 families. Only these reads were further analyzed, as the main focus of this study was on EPS and LPS producers of bacterial origin, and molecular data on other microbial polysaccharide producers in the employed databases is poor. The principal coordinate analysis (PCoA) ordination plot (Fig. 1(a)) showed that bacterial communities were highly similar at the family level in bulk soils, and underwent differentiation during the development of biocrusts. Dominant families were identified by selecting the five most abundant families from each location at each time point, and sorting them according to their relative abundance of all metagenomes. Relative abundances of the dominant families are shown in Supplementary Material 1 Fig. S6. As confirmed by ANOVA, the most characteristic families for T0 were *Burkholderiaceae*, *Comamonadaceae*, and

Table 2 DNA concentration, dissolved organic carbon (DOC) and nitrogen (DON), pH, and water repellency index values. The repellency index has no dimension

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

The en dash (–) signifies that the parameter was not measured for the respective samples; *bdl*, below detection limit

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

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

Functional Annotation of Metagenomic Datasets

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

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

ANOVA revealed that the relative abundances of most investigated genes changed mainly between T0 and T1. However, the differences in the relative numbers of gene copies were also driven by the underlying soil substrate (Supplementary Material 2 Table S5). In particular, the genes *wza* and *wcaF* increased at T1, and the increase was more pronounced in samples originating from Chicken Creek compared with that in those from Lieberose. Moreover, *wzt* increased in biocrusts grown on soil substrate from Chicken Creek already at T1, while the increase was observed in biocrusts grown on bulk soil taken from Lieberose only at T2. Conversely, *kpsE* and *lptC* decreased at T1. Additionally, *kpsE* was relatively more abundant in samples from Lieberose, whereas *lptG* was dominating in samples from Chicken Creek. Finally, *wcaK*/amsJ, *algJ*, *sacB*, and *lptF* were not significantly affected by either incubation time or soil substrate.

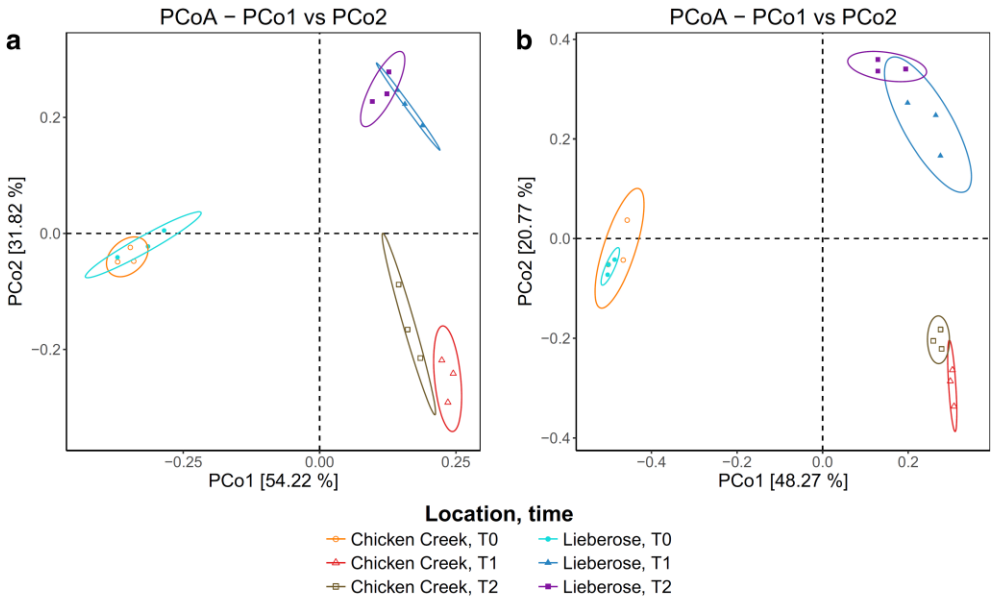


Fig. 1 PCoA plots depicting differences on the family level in (a) bacterial community composition and (b) taxonomic affiliation of genes related to EPS and LPS formation. Ellipses drawn around triplicates represent a 95% confidence level

Investigation of Potential EPS/LPS Producers

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

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

ANOVA identified a significant impact on the overall relative abundance of the investigated genes caused by location alone in three families, by time alone in 14 families, and by interaction of both factors in 23 families. The full list of affected families can be taken from Supplementary Material 2 Table S3. The PCoA plot (Fig. 1(b)) indicated that the distribution pattern of the analyzed genes among bacterial families resembled that of the total bacterial community (Fig. 1(a)). In fact, Spearman's rank correlation analysis revealed a positive correlation between the total abundance of a given family and the amount of sequences related to EPS and LPS formation harbored by that family for 57 families of potential EPS and LPS producers (average Rho = 0.69, minimum Rho = 0.47, maximum Rho = 0.97). Three families showed a negative correlation (average Rho = 0.56, minimum Rho = 0.49, maximum Rho = 0.61), and 150 exhibited no correlation (average Rho = 0.23, minimum Rho = 0.00, maximum Rho = 0.46).

Of the 57 families that showed a positive correlation, 25 exceeded an abundance of 1%, and encompassed altogether 43.26% of all bacterial reads. Both the relative abundance as well as the potential for EPS and LPS synthesis and export of these families were strongly influenced by both incubation time and underlying soil substrate (Fig. 4). In fact, these factors selected the key producers of EPSs and LPSs already at the phylum level. *Betaproteobacteria* (especially *Burkholderiaceae*), as well as *Gammaproteobacteria* (*Moraxellaceae*) and *Bacteroidetes* (*Flavobacteriaceae*), were prevalent at T0, although most of their members were found also at T1 and T2. *Deltaproteobacteria*

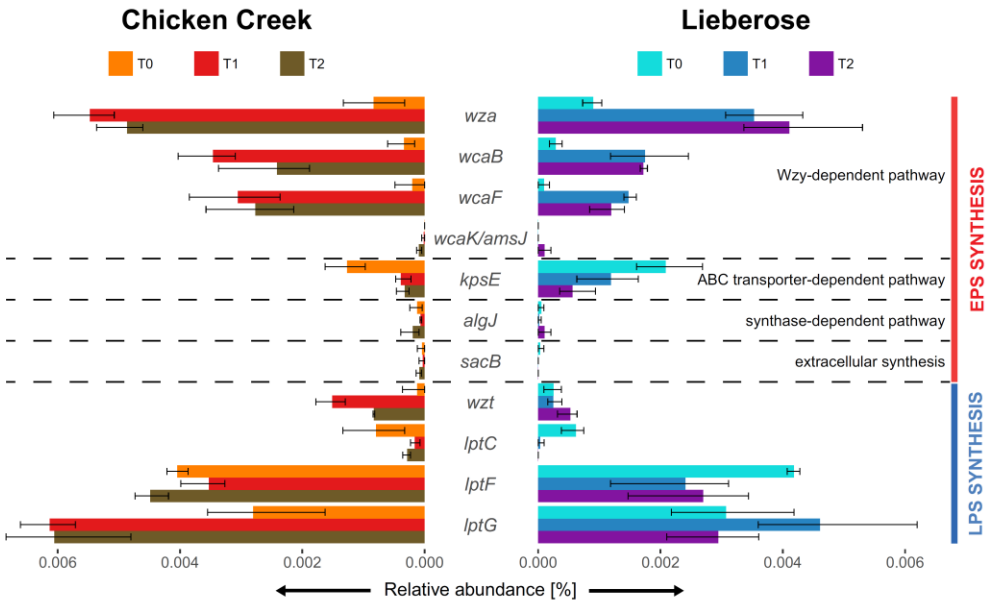


Fig. 2 Relative abundances of genes specific for the formation of EPSs and LPSs. Error bars show standard deviations of triplicates

(*Myxococcaceae* and *Archangiaceae*) and *Planctomycetes* (*Gemmataceae* and *Planctomycetaceae*) occurred mainly at T1 and T2 in biocrusts grown on soil taken from Chicken Creek. However, *Gemmataceae* were relatively abundant also at T2 in biocrusts originating from Lieberose. *Cyanobacteria* were characteristic for Chicken Creek samples from T1 and T2, but some of their members (*Oscillatoriaceae* and *Leptolyngbyaceae*) could also be important for EPS and LPS production in Lieberose samples from T1 and T2. Typical for Lieberose samples from T1 and T2 were *Chloroflexi* (*Ktedonobacteraceae* and *Thermogemmatisporaceae*) and *Acidobacteria* (*Acidobacteriaceae*). *Alphaproteobacteria* were prevalent at T1 and T2 in general, but some of their members were more characteristic for one of the underlying substrates (e.g., *Sphingomonadaceae* for soil from Chicken Creek, and *Acetobacteraceae* for that from Lieberose).

Discussion

Bacterial Communities of Initial Soils

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

Influence of Initial Biocrusts on Soil Stability and Hydrological Properties

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

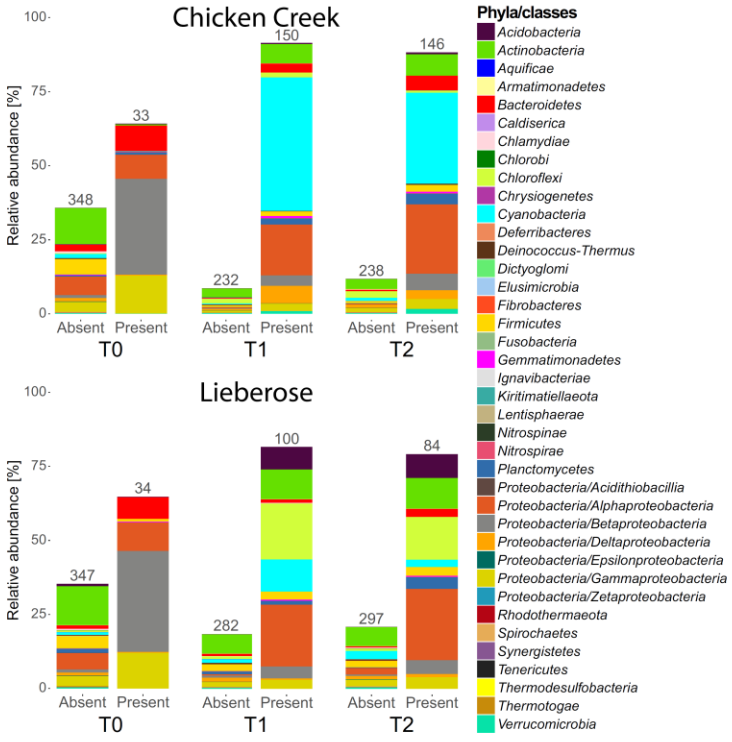


Fig. 3 Comparison of relative abundances of bacteria with and without the potential for EPS and LPS formation (labeled as "Present" and "Absent"). The distinction between the potential producers and non-producers was performed on the level of family. The families were then pooled according to their respective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

an important role in improving soil stability, especially in initial biocrusts that harbor large amounts of bacteria, like in our study. Bacterial polysaccharides adhere around soil particles, connecting them and cementing into larger aggregates [82]. Several studies demonstrated that bacterial polysaccharides increased the amount of stable soil aggregates [83–85] and reduced rainfall-induced erosion up to 98% [35]. Using the exemplary XCMT images of the 10-month-old samples from Chicken Creek, we also confirmed the ability of initial biocrusts to trap surface soil particles. Similar activity of cyanobacterial crusts was captured on XCMT images, for example, by Raanan et al. [86]. Moreover, the increase of the potential for EPS and LPS formation in the initial biocrusts compared with the bulk soils was correlated in our study with an accumulation of dissolved organic carbon (data not shown). Altogether, these point to an increased production of adhesive bacterial polysaccharides in our biocrusts.

Additionally, we measured the influence of biocrusts on soil hydrological properties, as the key role in altering soil moisture dynamics seems to be played by polysaccharides [1]. On one hand, they tend to clog pores through swelling, which may reduce soil infiltrability [32, 34, 49, 87]. On the other hand, they can increase soil porosity, which is known to positively affect water penetration [88, 89]. Some researchers also postulate that polysaccharides alter the hydrophobicity of biocrust surfaces [90]. In our study, the water repellency of biocrusts incubated for 10 months on both substrates was close to ideal wettability. Similar water repellency was reported for very young biocrusts also in other studies [49, 91]. In initial biocrusts, the effect on hydrological processes highly depends on the transient amount and chemical nature of polysaccharides building the bacterial biofilms [92]. For example, water molecules as well as nutrients are bound mainly by the hydrophilic polysaccharide fractions, while the hydrophobic fractions increase the stability of biocrusts and their ability to adhere to solid surfaces [93]. Furthermore, polysaccharides in bacterial biofilms are subjected to constant modification and degradation processes, both enzymatic and abiotic [4]. Colica et al. [94] underlined that polysaccharide content cannot be directly correlated with biocrust age, as the transient amount of polysaccharides in biocrusts depends on the activity of both polysaccharide producers as well as chemoheterotrophic organisms that use polymeric carbohydrates as a carbon source. Thus, the hydrological properties of biocrusts are highly dynamic and may fluctuate during biocrust development, as shown previously [49]. Comparing the structure of bacterial communities in biocrusts with the composition and chemical properties of bacterial polysaccharides throughout the whole development of biocrusts would surely shed more light on this issue. However, more research on the methods of extracting

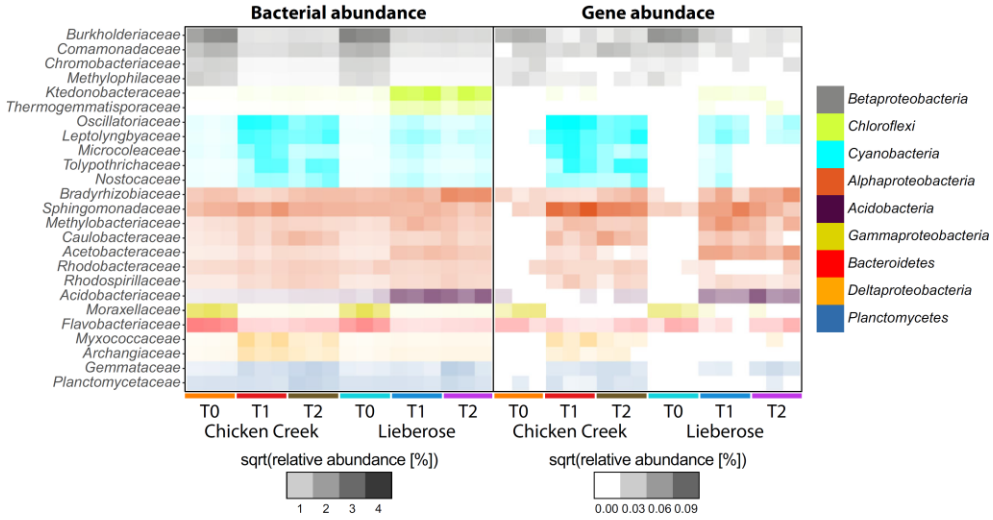


Fig. 4 Potential key families of EPS and LPS formation based on their relative abundance and the relative abundance of their genes related to EPS and LPS biosynthesis. Note the different color intensities between bacterial and gene abundances. The replicates (n = 3) are presented in separate columns

bacterial polysaccharides from biocrust needs to be done before such measurements will be reliable and give additional information compared to the repellency index [4, 95].

Genes Related to EPS and LPS Formation

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

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

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

Differentiation of Potential Key Producers of EPSs and LPSs During Initial Development of Biocrusts on Different Soil Substrates

Even though the soil substrate had little impact on the relative abundance of the investigated genes, it shaped the composition of bacterial communities in the developing biocrusts. In fact, bacterial communities that were highly similar in the bulk soils underwent differentiation once biocrusts started to develop. Furthermore, the taxonomic affiliation of the investigated

genes reflected the overall composition of the bacterial communities in our study, and thus the differentiation of the overall bacterial communities was accompanied by the differentiation of the communities of potential producers of adhesive polysaccharides. This is in line with the theory about functional redundancy, which states that important functions are preserved by a community even if the community changes its composition [108]. Our results indicate that the potential to form EPSs and LPSs is an important trait for initial biocrusts, as it is maintained despite the different development of bacterial communities on the two investigated substrates.

The importance of the potential to produce "soil glue" in the initial stage of biocrust development is further underlined by the fact that the highest numbers of sequences related to EPS and LPS biosynthesis were harbored by the families dominating the initial biocrusts. The potential key producers of adhesive polysaccharides found in biocrusts grown on soil from the Chicken Creek catchment and the Lieberose sand dune were distinct already at the phylum level. In the Chicken Creek biocrusts, the most abundant potential producers of EPSs and LPSs were *Cyanobacteria*. They are well known for their capability to form external polysaccharidic layers that enable them to survive in extreme environments [13]. In fact, the genetic machinery of the LPS synthesis and the Wzy-dependent pathway of the EPS synthesis were both found in *Cyanobacteria* before [27]. This explains the dominance of these particular polysaccharide biosynthesis pathways in the metagenomes from the Chicken Creek biocrusts. However, *Cyanobacteria* played only a minor role in the community of potential EPS and LPS producers in the Lieberose biocrusts, possibly because they prefer alkaline environments [1]. In the biocrusts grown on the soil from Lieberose, *Cyanobacteria* were replaced by *Chloroflexi* and *Acidobacteria*, which favor acidic habitats [109–113]. While *Chloroflexi* lack the ability to synthesize LPSs, *Acidobacteria* are known LPS producers [96]. Furthermore, even though the information on the proficiency of both phyla in EPS formation is still limited, sequences related to EPS synthesis were previously found in *Acidobacteria*, and a recent report suggests that some members of this phylum produce large amount of EPSs [114]. *Acidobacteria* and *Chloroflexi* are also common members of communities that embed themselves in an EPS matrix, such as biofilms, microbial mats, and biocrusts [115–117]. The low relative abundance of *Cyanobacteria* in biocrusts grown on the soil from Lieberose suggests that, besides *Chloroflexi*, the other major phototrophic organisms there could have been algae, which are also well-known producers of EPSs. Algae dominate acidic soils and are major components of the natural biocrusts found at Lieberose, except for the terminal successional stage that is dominated by mosses and fungi [1, 44, 45, 49]. However, the identification of eukaryotes involved in polysaccharide production is difficult using short-read shotgun sequencing and would require a different approach [118, 119]. In any case, our results show that potential producers of EPSs and LPSs dominate bacterial communities of biocrusts during the initial stage of biocrust development. Consequently, the differentiation of overall bacterial communities leads to the emergence of distinct potential key producers of "soil glue."

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

Conclusions

Our study indicates that the potential to produce EPSs and LPSs is an important trait for bacterial communities forming biocrusts in the initial stage of biocrust development, as (i) the relative abundance of genes related to the biosynthesis of adhesive polysaccharides increases in the bacterial communities of initial biocrusts compared with the indigenous bacterial communities of bulk soils, (ii) the relative abundances of EPS and LPS genes remain similar in initial biocrusts with different compositions of bacterial communities, and (iii) the highest numbers of sequences related to the "soil glue" production is found in families dominating initial biocrusts. Furthermore, we demonstrate that the community composition of potential producers of EPSs and LPSs reflects the overall structure of bacterial communities in initial biocrusts, and thus, initial biocrusts with different bacterial community compositions harbor distinct potential key producers of adhesive polysaccharides. Whether the ability of biocrusts to improve soil development in the long term is compromised by differences in the efficiency of polysaccharide formation, or the adhesive properties of EPSs and LPSs produced by different taxa, needs further investigation. Similarly, whether the differentiation of bacterial communities during the initial development of biocrusts is primarily triggered by soil

properties, or results from various rare species present in the initial bacterial community of bulk soil, remains to be determined.

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

Compliance with Ethical Standards

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

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

References

1. Belnap J, Lange OL (2003) Biological soil crusts: structure, function, and management. Springer, New York
2. Sancho LG, Maestre FT, Büdel B (2014) Biological soil crusts in a changing world: introduction to the special issue. Biodivers Conserv 23:1611–1617
3. Wu Y, Rao B, Wu P, Liu Y, Li G, Li D (2013) Development of artificially induced biological soil crusts in fields and their effects on top soil. Plant Soil 370:115–124
4. Rossi F, Mugnai G, De Philippis R (2018) Complex role of the polymeric matrix in biological soil crusts. Plant Soil 429:19–34
5. Totsche KU, Amelung W, Gerzabek MH, Guggenberger G, Klumpp E, Knief C, Lehndorff E, Mikutta R, Peth S, Prechtel A (2018) Microaggregates in soils. J Plant Nutr Soil Sc 181:104–136
6. Six J, Bossuyt H, Degryze S, Denef K (2004) A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. Soil Till Res 79:7–31
7. Abu-Lail NI, Camesano TA (2003) Role of lipopolysaccharides in the adhesion, retention, and transport of Escherichia coli JM109. Environ Sci Technol 37:2173–2183
8. Veste M, Littmann T, Breckle S-W, Yair A (2001) The role of biological soil crusts on desert sand dunes in the northwestern Negev, Israel. In: Breckle S-W, Veste M, Wucherer W (eds) Sustainable land use in deserts. Springer, Heidelberg, pp 357–367
9. Selbmann L, Stingele F, Petruccioli M (2003) Exopolysaccharide production by filamentous fungi: the example of Botryosphaeria rhodina. Anton Leeuw Int J G 84:135–145
10. Martínez-Cánovas MJ, Quesada E, Martínez-Checa F, del Moral A, Bejar V (2004) Salipiger mucescens gen. nov., sp. nov., a moderately halophilic, exopolysaccharide-producing bacterium isolated from hypersaline soil, belonging to the α-proteobacteria. Int J Syst Evol Micr 54:1735–1740
11. Suela Silva M, Naves Sales A, Teixeira Magalhães-Guedes K, Ribeiro Dias D, Schwan RF (2013) Brazilian Cerrado soil Actinobacteria ecology. Biomed Res Int 2013:503805
12. Wu N, Zhang Y, Pan H, Zhang J (2010) The role of nonphotosynthetic microbes in the recovery of biological soil crusts in the Gurbantunggut Desert, Northwestern China. Arid Land Res Manag 24:42–56
13. Pereira S, Zille A, Micheletti E, Moradas-Ferreira P, De Philippis R, Tamagnini P (2009) Complexity of cyanobacterial exopolysaccharides: composition, structures, inducing factors and putative genes involved in their biosynthesis and assembly. FEMS Microbiol Rev 33:917–941
14. Seviour R, Stasinopoulos S, Auer D, Gibbs P (1992) Production of pullulan and other exopolysaccharides by filamentous fungi. Crc Cr Rev Biotechn 12:279–298
15. Mahapatra S, Banerjee D (2013) Fungal exopolysaccharide: production, composition and applications. Microbiol Insights 6:1–16
16. Suresh Kumar A, Mody K, Jha B (2007) Bacterial exopolysaccharides–a perception. J Basic Microb 47:103–117
17. Zaady E, Kuhn U, Wilske B, Sandoval-Soto L, Kesselmeier J (2000) Patterns of CO2 exchange in biological soil crusts of successional age. Soil Biol Biochem 32:959–966
18. Fischer T, Gypser S, Subbotina M, Veste M (2014) Synergic hydraulic and nutritional feedback mechanisms control surface patchiness of biological soil crusts on tertiary sands at a post-mining site. J Hydrol Hydromech 62:293–302
19. Fischer T, Veste M (2018) Carbon cycling of biological soil crusts mirrors ecological maturity along a Central European inland dune catena. Catena 160:68–75
20. West NE (1990) Structure and function of microphytic soil crusts in wildland ecosystems of arid to semi-arid regions. Adv Ecol Res 20:179–223
21. Eldridge D, Greene R (1994) Microbiotic soil crusts-a review of their roles in soil and ecological processes in the rangelands of Australia. Soil Res 32:389–415
22. Kidron G, Barzilay E, Sachs E (2000) Microclimate control upon sand microbiotic crusts, western Negev Desert, Israel. Geomorphology 36:1–18
23. Yair A, Almog R, Veste M (2011) Differential hydrological response of biological topsoil crusts along a rainfall gradient in a sandy arid area: Northern Negev desert, Israel. Catena 87:326–333
24. Kidron GJ, Vonshak A (2012) The use of microbiotic crusts as biomarkers for ponding, subsurface flow and soil moisture content and duration. Geoderma 181:56–64
25. Belnap J (2006) The potential roles of biological soil crusts in dryland hydrologic cycles. Hydrol Process 20:3159–3178
26. Schmid J, Sieber V, Rehm B (2015) Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. Front Microbiol 6:496
27. Pereira SB, Mota R, Santos CL, De Philippis R, Tamagnini P (2013) Assembly and export of extracellular polymeric substances (EPS) in cyanobacteria: a phylogenomic approach. Adv Bot Res 65:235–279
28. Hunt F (1985) Patterns of LPS synthesis in gram negative bacteria. J Theor Biol 115:213–219
29. Wang X, Quinn PJ (2010) Lipopolysaccharide: biosynthetic pathway and structure modification. Prog Lipid Res 49:97–107
30. Whitfield C, Trent MS (2014) Biosynthesis and export of bacterial lipopolysaccharides. Annu Rev Biochem 83:99–128
31. Costa OY, Raaijmakers JM, Kuramae EE (2018) Microbial extracellular polymeric substances: ecological function and impact on soil aggregation. Front Microbiol 9:1–14
32. Mazor G, Kidron GJ, Vonshak A, Abeliovich A (1996) The role of cyanobacterial exopolysaccharides in structuring desert microbial crusts. FEMS Microbiol Ecol 21:121–130

33. HuiXia P, ZhengMing C, XueMei Z, ShuYong M, XiaoLing Q, Fang W (2007) A study on an oligotrophic bacteria and its ecological characteristics in an arid desert area. Sci China Ser D 50: 128–134
34. Colica G, Li H, Rossi F, Li D, Liu Y, De Philippis R (2014) Microbial secreted exopolysaccharides affect the hydrological behavior of induced biological soil crusts in desert sandy soils. Soil Biol Biochem 68:62–70
35. Kheirfam H, Sadeghi SH, Darki BZ, Homaee M (2017) Controlling rainfall-induced soil loss from small experimental plots through inoculation of bacteria and cyanobacteria. Catena 152:40–46
36. Mugnai G, Rossi F, Felde VJMNL, Colesie C, Büdel B, Peth S, Kaplan A, De Philippis R (2018) The potential of the cyanobacterium Leptolyngbya ohadii as inoculum for stabilizing bare sandy substrates. Soil Biol Biochem 127:318–328
37. Cania B, Vestergaard G, Krauss M, Fliessbach A, Schloter M, Schulz S (2019) A long-term field experiment demonstrates the influence of tillage on the bacterial potential to produce soil structure-stabilizing agents such as exopolysaccharides and lipopolysaccharides. Environ Microbiome 1:1–14
38. Gerwin W, Schaaf W, Biemelt D, Fischer A, Winter S, Hüttl RF (2009) The artificial catchment "Chicken Creek"(Lusatia, Germany)—a landscape laboratory for interdisciplinary studies of initial ecosystem development. Ecol Eng 35:1786–1796
39. Russell DJ, Hohberg K, Elmer M (2010) Primary colonisation of newly formed soils by actinedid mites. Soil Org 82:237–251
40. Zaplata MK, Winter S, Fischer A, Kollmann J, Ulrich W (2012) Species-driven phases and increasing structure in early-successional plant communities. Am Nat 181:E17–E27
41. Lukešová A (2001) Soil algae in brown coal and lignite post-mining areas in central Europe (Czech Republic and Germany). Restor Ecol 9:341–350
42. Gypser S, Herppich WB, Fischer T, Lange P, Veste M (2016) Photosynthetic characteristics and their spatial variance on biological soil crusts covering initial soils of post-mining sites in Lower Lusatia, NE Germany. Flora 220:103–116
43. Dümig A, Veste M, Hagedorn F, Fischer T, Lange P, Spröte R, Kögel-Knabner I (2014) Organic matter from biological soil crusts induces the initial formation of sandy temperate soils. Catena 122: 196–208
44. Fischer T, Veste M, Bens O, Hüttl RF (2012) Dew formation on the surface of biological soil crusts in central European sand ecosystems. Biogeosciences 9:4621–4628
45. Fischer T, Veste M, Eisele A, Bens O, Spyra W, Hüttl RF (2012) Small scale spatial heterogeneity of normalized difference vegetation indices (NDVIs) and hot spots of photosynthesis in biological soil crusts. Flora 207:159–167
46. Döhring T, Koefferlein M, Thiel S, Seidlitz HK (1996) Spectral shaping of artificial UV-B irradiation for vegetation stress research. J Plant Physiol 148:115–119
47. Thiel S, Döhring T, Köfferlein M, Kosak A, Martin P, Seidlitz HK (1996) A phytotron for plant stress research: how far can artificial lighting compare to natural sunlight? J Plant physiol 148:456–463
48. Brankatschk R, Töwe S, Kleineidam K, Schloter M, Zeyer J (2011) Abundances and potential activities of nitrogen cycling microbial communities along a chronosequence of a glacier forefield. ISME J 5:1025–1037
49. Fischer T, Veste M, Wiehe W, Lange P (2010) Water repellency and pore clogging at early successional stages of microbiotic crusts on inland dunes, Brandenburg, NE Germany. Catena 80: 47–52
50. Hallett P, Young I (1999) Changes to water repellence of soil aggregates caused by substrate-induced microbial activity. Eur J Soil Sci 50:35–40
51. Urbanek E, Hallett P, Feeney D, Horn R (2007) Water repellency and distribution of hydrophilic and hydrophobic compounds in soil aggregates from different tillage systems. Geoderma 140: 147–155
52. Köhne JM, Schlüter S, Vogel H-J (2011) Predicting solute transport in structured soil using pore network models. Vadose Zone J 10:1082–1096
53. Vogel H-J, Weller U, Schlüter S (2010) Quantification of soil structure based on Minkowski functions. Comput Geosci 36: 1236–1245
54. Vestergaard G, Schulz S, Schöler A, Schloter M (2017) Making big data smart—how to use metagenomics to understand soil quality. Biol Fert Soils 53:479–484. https://doi.org/10.1007/s00374-017-1191-3
55. Schubert M, Lindgreen S, Orlando L (2016) AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res Notes 9:88
56. Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863–864
57. Schmieder R, Edwards R (2011) Fast identification and removal of sequence contamination from genomic and metagenomic datasets. PloS One 6:e17288
58. Rodriguez-R LM, Konstantinidis KT (2014) Estimating coverage in metagenomic data sets and why it matters. ISME J 8:2349–2351
59. Menzel P, Ng KL, Krogh A (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun 7:11257
60. Kopylova E, Noé L, Touzet H (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28:3211–3217
61. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M (2015) eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res 44:D286–D293
62. Haft DH, Selengut JD, Richter RA, Harkins D, Basu MK, Beck E (2013) TIGRFAMs and genome properties in 2013. Nucleic Acids Res 41:D387–D395
63. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A (2016) The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279–D285
64. Rho M, Tang H, Ye Y (2010) FragGeneScan: predicting genes in short and error-prone reads. Nucleic Acids Res 38:e191–e191
65. Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60
66. R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
67. Field A, Miles J, Field Z (2012) Discovering statistics using R. Sage publications, Thousand Oaks
68. Wilcox RR, Schönbrodt FD (2014) The WRS package for robust statistics in R. R package version 0.24. Retrieved from http://r-forge.r-project.org/projects/wrs/
69. Tunks T (1978) The use of omega squared in interpreting statistical significance. B Coun Res Music Ed 57:28–34.
70. Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290
71. Legendre P, Legendre LF (2012) Numerical ecology. Elsevier, Amsterdam
72. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2018) Vegan: community ecology package. R package version 2:5–1
73. Sul WJ, Asuming-Brempong S, Wang Q, Tourlousse DM, Penton CR, Deng Y, Rodrigues JL, Adiku SG, Jones JW, Zhou J (2013) Tropical agricultural land management influences on soil

microbial communities through its effect on soil organic carbon. Soil Biol Biochem 65:33–38
74. Cederlund H, Wessén E, Enwall K, Jones CM, Juhanson J, Pell M, Philippot L, Hallin S (2014) Soil carbon quality and nitrogen fertilization structure bacterial communities with predictable responses of major bacterial phyla. Appl Soil Ecol 84:62–68
75. Coenye T (2014) The family Burkholderiaceae. In: Rosenberg E (ed) The prokaryotes: Alphaproteobacteria and Betaproteobacteria. Springer, Heidelberg, pp 759–776
76. Willems A (2014) The family Comamonadaceae. In: Rosenberg E (ed) The prokaryotes: Alphaproteobacteria and Betaproteobacteria. Springer, Heidelberg, pp 777–851
77. Teixeira LM, Merquior VLC (2014) The family Moraxellaceae. In: Rosenberg E (ed) The prokaryotes: Gammaproteobacteria. Springer, Heidelberg, pp 443–476
78. McBride MJ (2014) The family Flavobacteriaceae. In: Rosenberg E (ed) The prokaryotes: other major lineages of bacteria and the archaea. Springer, Heidelberg, pp 643–676
79. Vogeleer P, Tremblay YD, Mafu AA, Jacques M, Harel J (2014) Life on the outside: role of biofilms in environmental persistence of Shiga-toxin producing Escherichia coli. Front Microbiol 5:317
80. Lindhout T, Lau PCY, Brewer D, Lam JS (2009) Truncation in the core oligosaccharide of lipopolysaccharide affects flagella-mediated motility in Pseudomonas aeruginosa PAO1 via modulation of cell surface attachment. Microbiology +155:3449–3460
81. Kierek K, Watnick PI (2003) The Vibrio cholerae O139 O-antigen polysaccharide is essential for Ca2+-dependent biofilm development in sea water. Proc Natl Acad Sci U S A 100:14357–14362
82. Huang Q, Wu H, Cai P, Fein JB, Chen W (2015) Atomic force microscopy measurements of bacterial adhesion and biofilm formation onto clay-sized particles. Sci Rep-UK 5:16857
83. de Caire GZ, De Cano MS, De Mule MZ, Palma R, Colombo K (1997) Exopolysaccharide of Nostoc muscorum (Cyanobacteria) in the aggregation of soil particles. J Appl Phycol 9:249–253
84. Carrasco L, Caravaca F, Azcón R, Roldán A (2009) Soil acidity determines the effectiveness of an organic amendment and a native bacterium for increasing soil stabilisation in semiarid mine tailings. Chemosphere 74:239–244
85. Rossi F, Li H, Liu Y, De Philippis R (2017) Cyanobacterial inoculation (cyanobacterisation): perspectives for the development of a standardized multifunctional technology for soil fertilization and desertification reversal. Earth-Sci Rev 171:28–43
86. Raanan H, Felde VJ, Peth S, Drahorad S, Ionescu D, Eshkol G, Treves H, Felix-Henningsen P, Berkowicz SM, Keren N (2016) Three-dimensional structure and cyanobacterial activity within a desert biological soil crust. Environ Microbiol 18:372–383
87. Issa OM, Défarge C, Trichet J, Valentin C, Rajot J-L (2009) Microbiotic soil crusts in the Sahel of Western Niger and their influence on soil porosity and water dynamics. Catena 77:48–55
88. Greene R (1992) Soil physical properties of three geomorphic zones in a semi-arid mulga woodland [Acacia aneura]. Aust J Soil Res 30:55–69
89. Eldridge DJ (2003) Biological soil crusts and water relations in Australian deserts. In: Belnap J, Lange OL (eds) Biological soil crusts: structure, function, and management. Springer, Berlin, pp 327–337
90. Felde VJMNL, Rossi F, Colesie C, Uteau-Puschmann D, Horne R, Felix-Henningsen P, De Philippis R, Peth S (2016) Pore characteristics in biological soil crusts are independent of extracellular polymeric substances. Soil Biol Biochem 103:294–299
91. Pluis J (1994) Algal crust formation in the inland dune area, Laarder Wasmeer, the Netherlands. Vegetation 113:41–51
92. Flemming H-C, Wingender J (2010) The biofilm matrix. Nat Rev Microbiol 8:623–633
93. Rossi F, Micheletti E, Bruno L, Adhikary SP, Albertano P, De Philippis R (2012) Characteristics and role of the exocellular polysaccharides produced by five cyanobacteria isolated from phototrophic biofilms growing on stone monuments. Biofouling 28:215–224
94. Colica G, Li H, Rossi F, Philippis RD, Liu Y (2015) Differentiation of the characteristics of excreted extracellular polysaccharides reveals the heterogeneous primary succession of induced biological soil crusts. J Appl Phycol 27:24–32
95. Redmile-Gordon M, Brookes P, Evershed R, Goulding K, Hirsch P (2014) Measuring the soil-microbial interface: extraction of extracellular polymeric substances (EPS) from soil biofilms. Soil Biol Biochem 72:163–171
96. Lagier J-C, Million M, Hugon P, Armougom F, Raoult D (2012) Human gut microbiota: repertoire and variations. Front Cell Infect Mi 2:136
97. Antunes LCS, Poppleton D, Klingl A, Criscuolo A, Dupuy B, Brochier-Armanet C, Beloin C, Gribaldo S (2016) Phylogenomic analysis supports the ancestral presence of LPS-outer membranes in the Firmicutes. Elife 5:e14589
98. Poppleton DI, Duchateau M, Hourdel V, Matondo M, Flechsler J, Klingl A, Beloin C, Gribaldo S (2017) Outer membrane proteome of Veillonella parvula: a diderm Firmicute of the human microbiome. Front Microbiol 8:1215
99. Whitfield C, Larue K (2008) Stop and go: regulation of chain length in the biosynthesis of bacterial polysaccharides. Nat Struct Mol Biol 15:121–123
100. Whitfield C (2010) Polymerases: glycan chain-length control. Nat Chem Biol 6:403–404
101. Rosenow C, Esumeh F, Roberts IS, Jann K (1995) Characterization and localization of the KpsE protein of Escherichia coli K5, which is involved in polysaccharide export. J Bacteriol 177:1137–1143
102. Bachtiar BM, Coloe PJ, Fry BN (2007) Knockout mutagenesis of the kpsE gene of Campylobacter jejuni 81116 and its involvement in bacterium–host interactions. FEMS Immunol Med Mic 49:149–154
103. Muhammadi AN (2007) Genetics of bacterial alginate: alginate genes distribution, organization and biosynthesis in bacteria. Curr Genomics 8:191–202
104. Rendueles O, Garcia-Garcerà M, Néron B, Touchon M, Rocha EP (2017) Abundance and co-occurrence of extracellular capsules increase environmental breadth: implications for the emergence of pathogens. PLoS Pathog 13:e1006525
105. Putker F, Bos MP, Tommassen J (2015) Transport of lipopolysaccharide to the Gram-negative bacterial cell surface. FEMS Microbiol Rev 39:985–1002
106. Ruiz N, Gronenberg LS, Kahne D, Silhavy TJ (2008) Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of Escherichia coli. Proc Natl Acad Sci U S A 105:5537–5542
107. Benedet M, Falchi FA, Puccio S, Di Benedetto C, Peano C, Polissi A, Dehò G (2016) The lack of the essential LptC protein in the trans-envelope lipopolysaccharide transport machine is circumvented by suppressor mutations in LptF, an inner membrane component of the Escherichia coli transporter. PloS One 11:e0161354
108. Allison SD, Martiny JBH (2008) Resistance, resilience, and redundancy in microbial communities. Proc Natl Acad Sci U S A 105:11512–11519
109. Jones RT, Robeson MS, Lauber CL, Hamady M, Knight R, Fierer N (2009) A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. ISME J 3:442–453
110. Lauber CL, Hamady M, Knight R, Fierer N (2009) Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl Environ Microbiol 75:5111–5120

111. Wilhelm RC, Niederberger TD, Greer C, Whyte LG (2011) Microbial diversity of active layer and permafrost in an acidic wetland from the Canadian High Arctic. Can J Microbiol 57: 303–315
112. Santofimia E, González-Toril E, López-Pamo E, Gomariz M, Amils R, Aguilera Á (2013) Microbial diversity and its relationship to physicochemical characteristics of the water in two extreme acidic pit lakes from the Iberian Pyrite Belt (SW Spain). PLoS One 8:e66746
113. Jones DS, Lapakko KA, Wenz ZJ, Olson MC, Roepke EW, Sadowsky MJ, Novak PJ, Bailey JV (2017) Novel microbial assemblages dominate weathered sulfide-bearing rock from copper-nickel deposits in the Duluth complex, Minnesota, USA. Appl Environ Microbiol 83:e00909–e00917
114. Kielak AM, Castellane TC, Campanharo JC, Colnago LA, Costa OY, Da Silva MLC, Van Veen JA, Lemos EG, Kuramae EE (2017) Characterization of novel Acidobacteria exopolysaccharides with potential industrial and ecological applications. Sci Rep-UK 7: 41193
115. Rampadarath S, Bandhoa K, Puchooa D, Jeewon R, Bal S (2017) Early bacterial biofilm colonizers in the coastal waters of Mauritius. Electron J Biotechn 29:13–21
116. Prieto-Barajas CM, Valencia-Cantero E, Santoyo G (2017) Microbial mat ecosystems: structure types, functional diversity, and biotechnological application. Electron J Biotechn 31:48–56
117. Mogul R, Vaishampayan P, Bashir M, McKay CP, Schubert K, Bornaccorsi R, Gomez E, Tharayil S, Payton G, Capra J (2017) Microbial community and biochemical dynamics of biological soil crusts across a gradient of surface coverage in the central Mojave desert. Front Microbiol 8:1974
118. De Vries M, Schöler A, Ertl J, Xu Z, Schloter M (2015) Metagenomic analyses reveal no differences in genes involved in cellulose degradation under different tillage treatments. FEMS Microbiol Ecol 91:fiv069
119. Wooley JC, Godzik A, Friedberg I (2010) A primer on metagenomics. PLoS Comput Biol 6:1–13
120. Fierer N (2017) Embracing the unknown: disentangling the complexities of the soil microbiome. Nat Rev Microbiol 15:579–590
121. Lammel DR, Barth G, Ovaskainen O, Cruz LM, Zanatta JA, Ryo M, de Souza EM, Pedrosa FO (2018) Direct and indirect effects of a pH gradient bring insights into the mechanisms driving prokaryotic community structures. Microbiome 6:106

Supplementary Materials for Publication 1

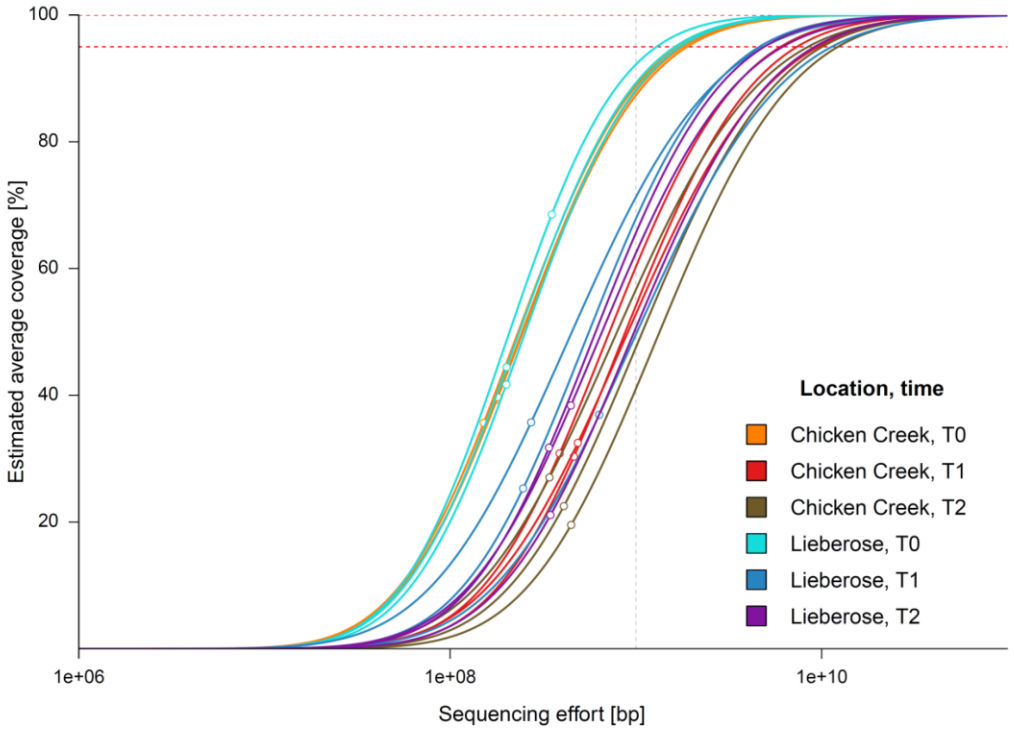


Fig. S1 Nonpareil estimation of diversity coverage in available metagenomic datasets. Nonpareil employs read redundancy to assess dataset complexity and the amount of sequences needed to achieve a desired level of coverage. Datasets that are more complex require larger sequencing effort to achieve the same level of coverage as less complex datasets. Thus, curves located rightward in the plot correspond to more diverse datasets than curves positioned on the left. Circles on curves represent the actual level of coverage achieved by the sequencing depth of each dataset.

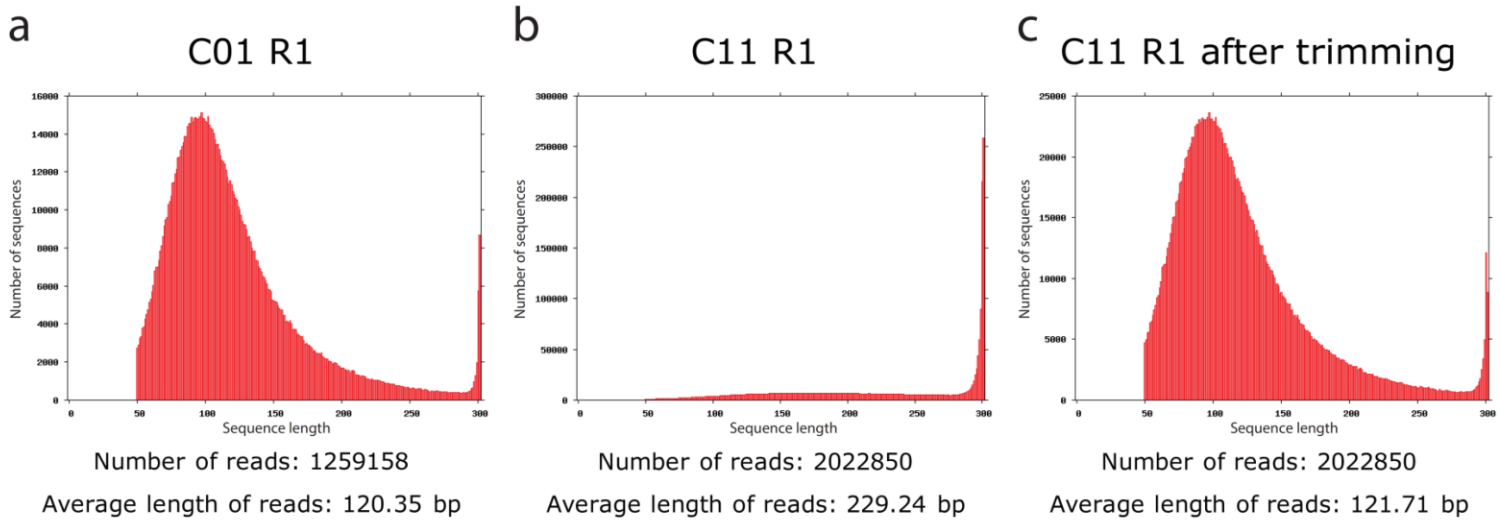


Fig. S2 Read length distribution of exemplary metagenomes before and after trimming. The trimming was performed to test if the read length significantly affects metagenomic annotations. Histogram “a” shows a metagenome obtained from bulk soil taken from Chicken Creek, which comprised relatively short sequences. Histogram “b” demonstrates a metagenome of biocrusts grown for 4 months on the soil taken from Chicken Creek. This metagenome contained sequences that were relatively long. Histogram “c” shows the same metagenome as histogram “b”, after trimming it to mimic the length distribution of reads in the metagenome presented in histogram “a”. The number of reads and the average length of reads in each metagenome are displayed underneath the respective histograms.

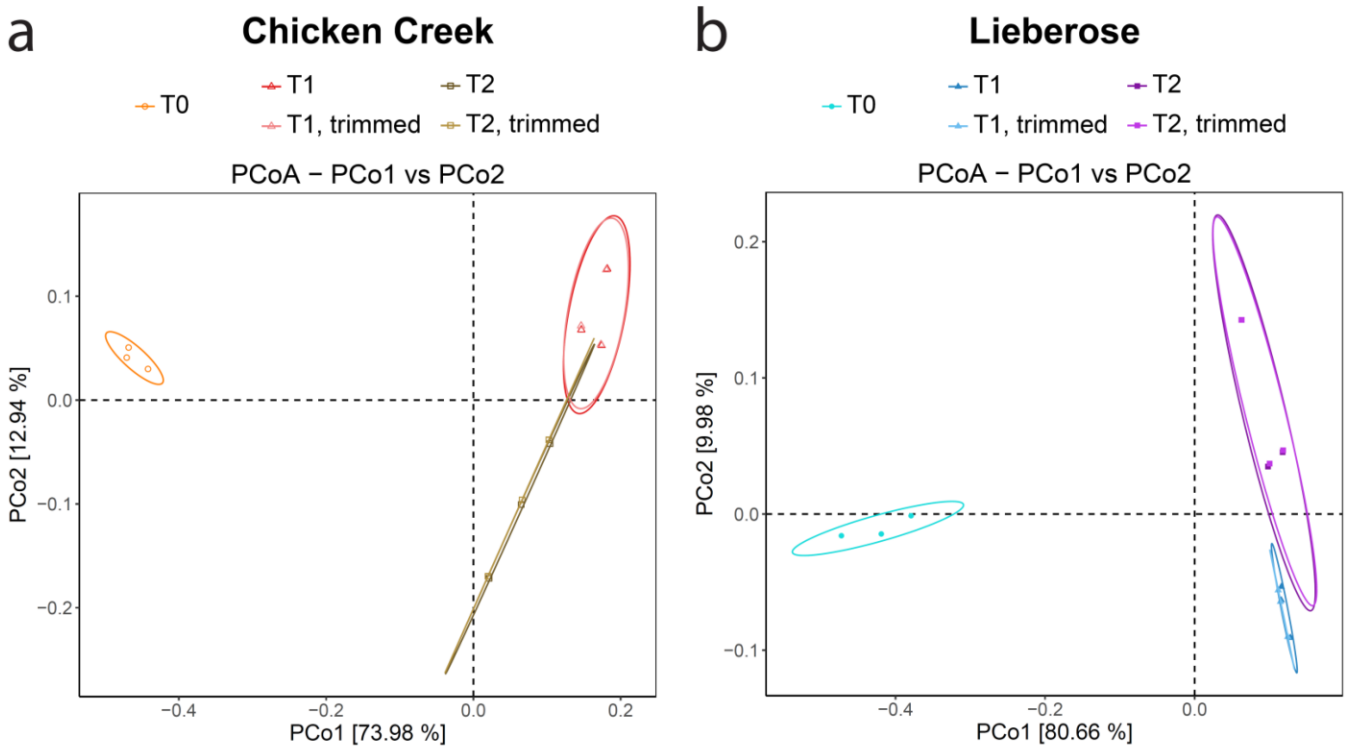


Fig. S3 PCoA ordination plots depicting taxonomic profiles of bacteria at the family level in bulk soil and biocrusts originating from a – Chicken Creek and b – Lieberose sand dune. Compared are metagenomes obtained from samples taken at T0, which comprised relatively short reads, metagenomes created from samples taken at T1 and T2, which contained relatively long reads, and T1 and T2 metagenomes, whose sequences were trimmed to resemble the length distribution of reads in T0 metagenomes. Ellipses drawn around triplicates represent a 95 % confidence level.

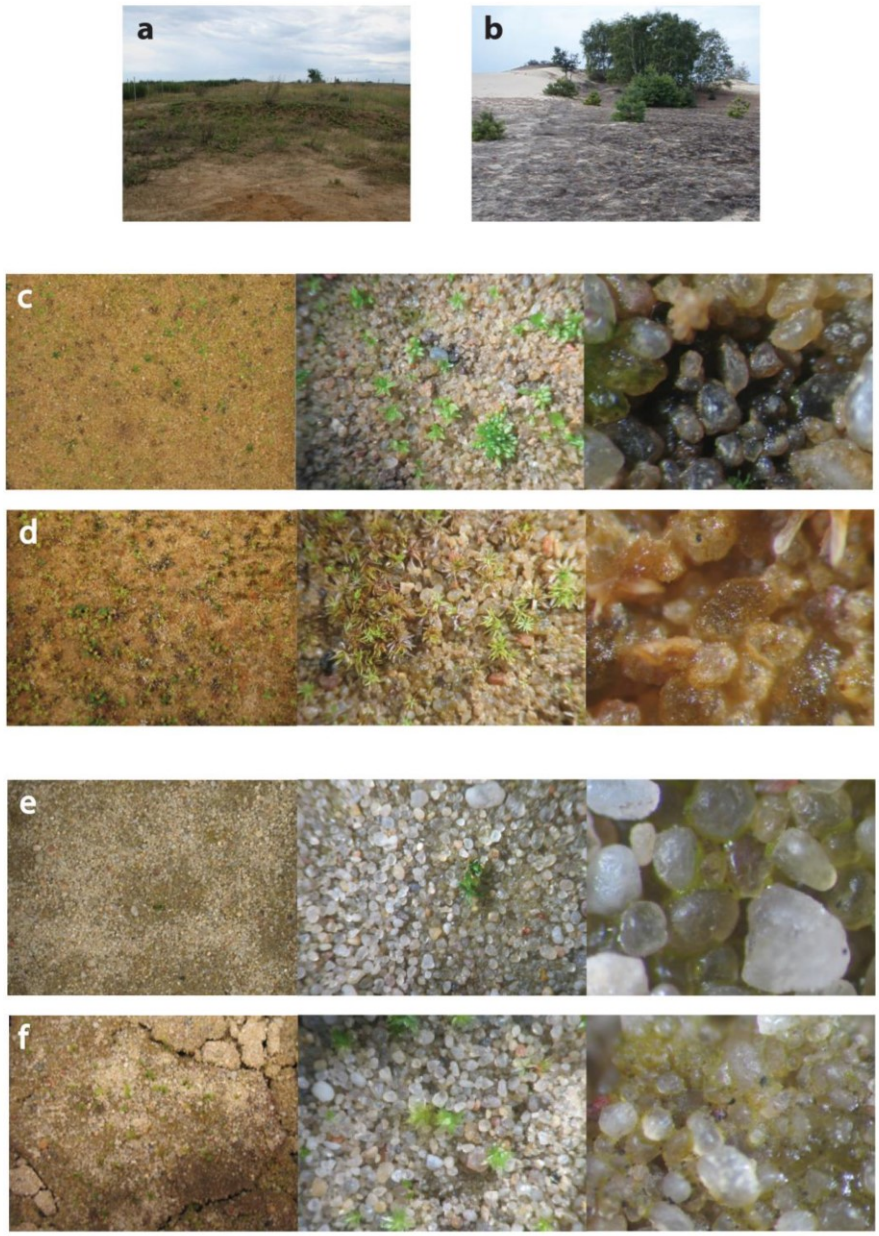


Fig. S4 Exemplary pictures of the Chicken Creek catchment (a) and the inland dune near Lieberose (b), as well as biocrusts after four (c+e) and ten months (d+f) of development on soil from Chicken Creek (c+d) or Lieberose (e+f).

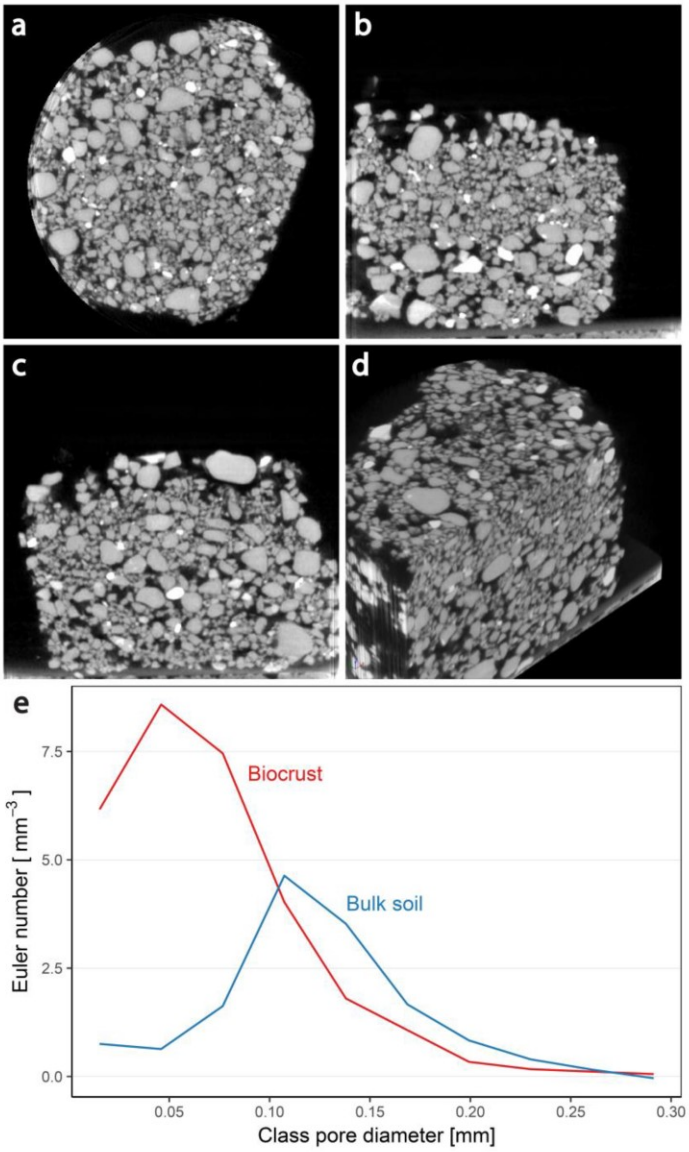


Fig. S5 Up (a-d): Computed tomography (CT) example images of biocrusts and the underlying soil originating from the Chicken Creek catchment and incubated for ten months. The pictures show an undisturbed sample from the top (a) and the sides (b-d). The side pictures present both the biocrusts and the underlying soil in the ratio of 50:50. Down (e): Euler number for pore size classes of 15 to 291 µm in biocrusts and the underlying soil calculated using the above CT images.

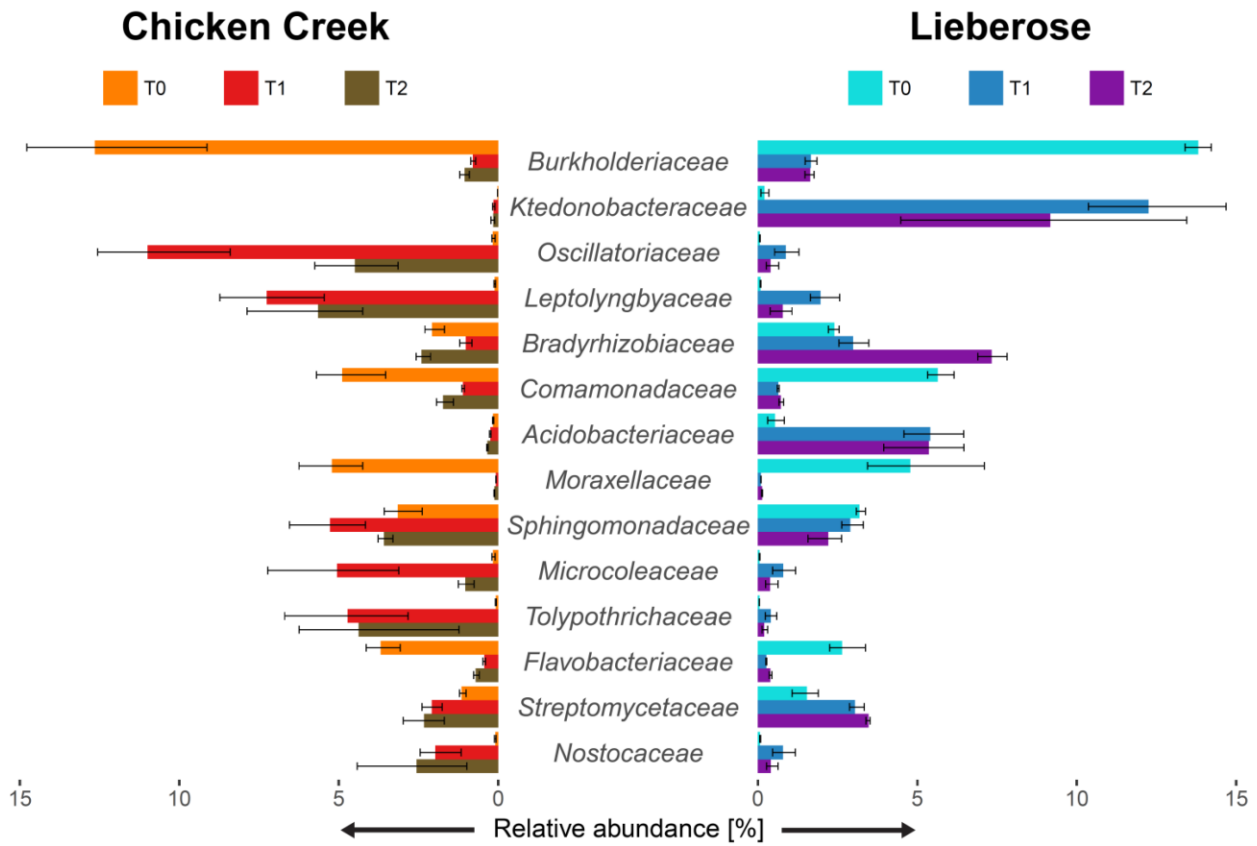


Fig. S6 Relative abundances of bacterial families dominating in metagenomes from two different soil substrates, and biocrusts cultivated in a microcosm experiment. Bulk soil was taken from the artificial catchment Chicken, and the inland dune near Lieberose. Samples were taken at the beginning of the experiment (bulk soils at T0), after four (biocrusts at T1) and ten months (biocrusts at T2) of incubation.

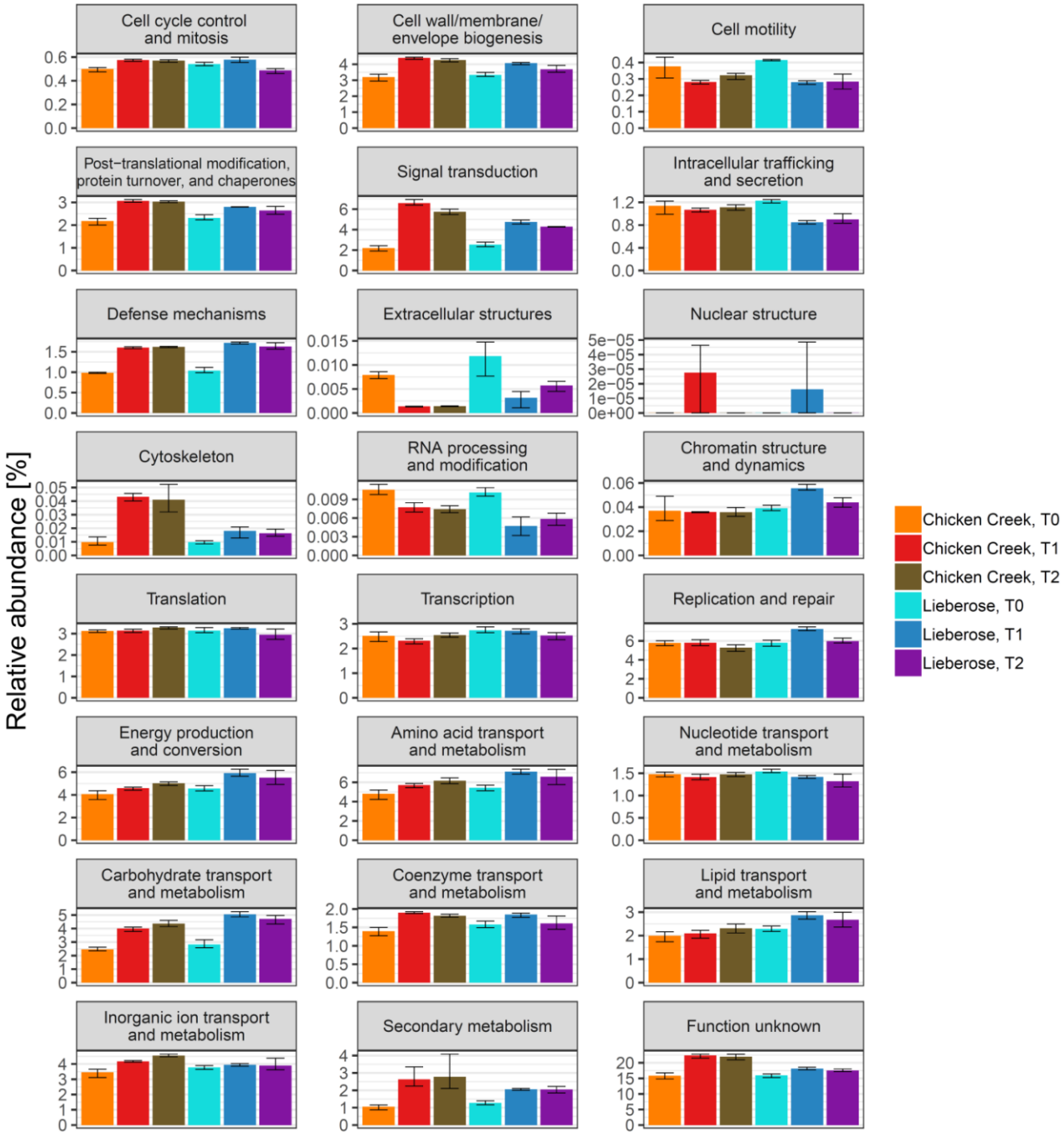


Fig. S7 Functional classification of bacterial reads according to COG functional categories.

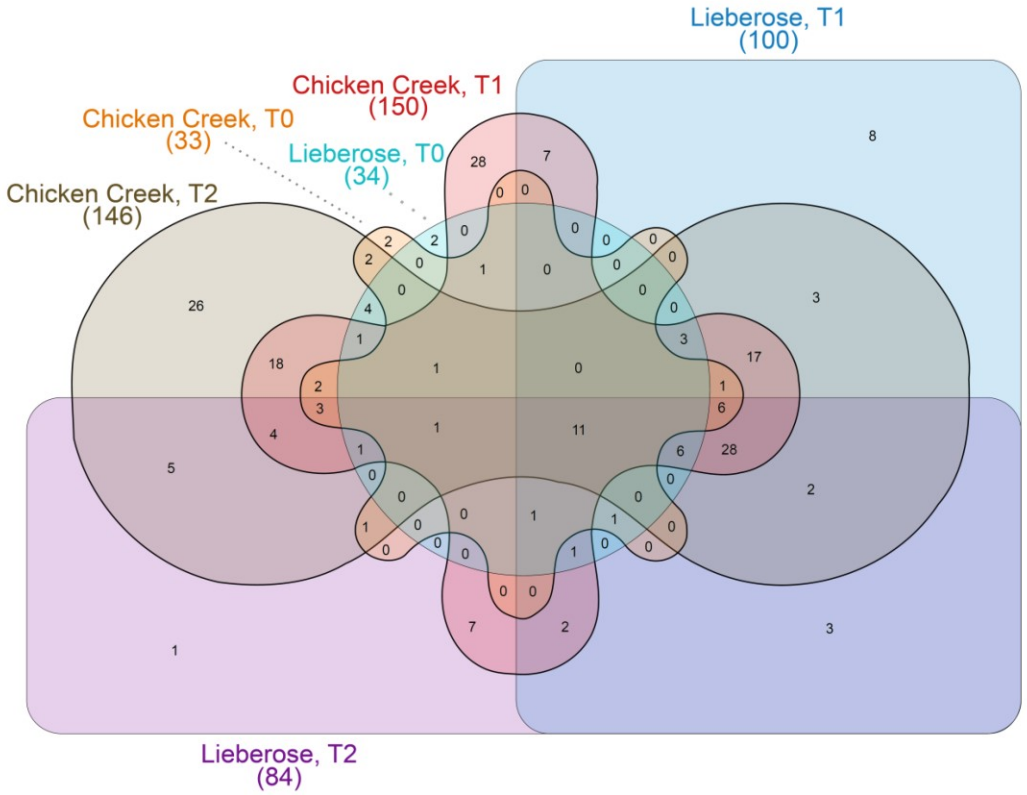


Fig. S8 VENN diagram showing bacterial families harboring genes related to EPS and LPS production in bulk soils from the artificial catchment Chicken Creek and the inland dune near Lieberose, as well as biocrusts cultivated on those soils in a microcosm experiment. Samples were taken at the beginning of the experiment (bulk soils at T0), after four (biocrusts at T1) and ten months (biocrusts at T2) of incubation. The numbers in brackets signify the total number of families harboring the investigated genes in the respective samples.

Table S1 Details of the raw and filtered sequencing data. Shown are the numbers of obtained reads, total length of the reads and average read length per sample before and after quality filtering. “C” and “L” at the beginning of the sample names stand for either “Chicken Creek” or “Lieberose”, respectively. The first following number stands for the sampling time point (T0, T1 and T2). The second following number distinguishes the replicates.

Chicken Creek

Raw data	C01	C02	C03	C11	C12	C13	C21	C22	C23
Number of reads	2630800	2871724	2785872	4106010	4033420	3166138	2837036	3716336	3291680
Total length of reads	791870800	864388924	838547472	1235909010	1214059420	953007538	853472671	1118016294	990254422
Average length of reads	301.00	301.00	301.00	301.00	301.00	301.00	300.83	300.84	300.84
Filtered data									
Number of reads	2518316	2762666	2685458	4045700	3983230	3135532	2817572	3691586	3271870
Total length of reads	303018026	355791397	361150059	923584003	969347744	773402666	681490167	891949733	813882126
Average length of reads	120.33	128.79	134.48	228.29	243.36	246.66	241.87	241.62	248.75

Lieberose

Raw data	L01	L02	L03	L11	L12	L13	L21	L22	L23
Number of reads	3290960	5614172	2931520	5106124	2146520	2698756	3202794	3782848	2751444
Total length of reads	990578960	1689865772	882114162	1536943324	645770263	811952501	963564101	1138011663	827744444
Average length of reads	301.00	301.00	300.91	301.00	300.85	300.86	300.85	300.83	300.84
Filtered data									
Number of reads	3124506	5325616	2797624	5060332	2129044	2676256	3187060	3761132	2737140
Total length of reads	402539051	704914302	400885847	1264941528	491768729	544078957	691111132	889116332	677362685
Average length of reads	128.83	132.36	143.30	249.97	230.98	203.30	216.85	236.40	247.47

Table S2 Influence of soil substrate (location), incubation time (time) and their interaction, on the relative abundances of the dominant bacterial families, as determined by a robust 2-way ANOVA. Effect sizes (ω^2) and significance levels were calculated based on triplicates (n = 3). Significance levels are represented by the amount of stars: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$.

ANOVA (omega squared and significance levels)						
	Location		Time		Location x Time	
Burkholderiaceae	0.01	ns	0.94	***	0.00	ns
Ktedonobacteraceae	0.43	***	0.22	***	0.21	***
Oscillatoriaceae	0.34	***	0.34	***	0.25	***
Leptolyngbyaceae	0.32	***	0.39	***	0.15	***
Bradyrhizobiaceae	0.35	***	0.41	***	0.22	***
Comamonadaceae	0.00	ns	0.91	***	0.02	ns
Acidobacteriaceae	0.51	***	0.22	***	0.20	***
Moraxellaceae	0.00	ns	0.88	***	0.00	ns
Sphingomonadaceae	0.31	***	0.18	ns	0.17	ns
Microcoleaceae	0.19	***	0.39	***	0.21	***
Tolypothrichaceae	0.29	***	0.14	***	0.10	***
Flavobacteriaceae	0.03	**	0.89	***	0.01	ns
Streptomycetaceae	0.23	***	0.58	***	0.02	ns
Nostocaceae	0.21	***	0.25	***	0.09	*

Table S3 Bacterial families whose relative abundances and potential to produce EPSs or LPSs were significantly affected by soil substrate (location), incubation time (time) or their interaction. Significant differences were determined by a robust 2-way ANOVA. ($n = 3$, $p < 0.05$).

Factor	Families whose abundance was affected	Families whose potential to produce EPS or LPS was affected
Location	*Actinopolysporaceae, Anaplasmataceae, Aphanizomenonaceae, Catenulisporaceae, Gordoniaceae, Micrococcaceae, Nakamurellaceae, Nocardiaceae, Ruminococcaceae, Segniliparaceae, Sphingomonadaceae, Synergistaceae, Ventosimonadaceae*	*Halanaerobiaceae, Hydrogenophilaceae, Rhodobacteraceae*
Time	*Acholeplasmataceae, Acidiferrobacteraceae, Alcaligenaceae, Alcanivoracaceae, Algiphilaceae, Alteromonadaceae, Anaerolineaceae, Anaeromyxobacteraceae, Aphanothecaceae, Aurantimonadaceae, Bacteroidaceae, Bartonellaceae, Blattabacteriaceae, Bogoriellaceae, Brevinemataceae, Brucellaceae, Burkholderiaceae, Campylobacteraceae, Candidatus.Brocadiaceae, Candidatus.Desulfofervidaceae, Candidatus.Paracaedibacteraceae, Caulobacteraceae, Cellulomonadaceae, Cellvibrionaceae, Chamaesiphonaceae, Chitinispirillaceae, Chlamydiaceae, Chlorobiaceae, Chromatiaceae, Chromobacteriaceae, Chroococcaceae, Chrysiogenaceae, Clostridiales Family XII. Incertae.Sedis, Comamonadaceae, Corynebacteriaceae, Coxiellaceae, Cyanobacteriaceae, Cyanothecaceae, Deferribacteraceae, Dermocarpellaceae, Desulfobacteraceae, Desulfobulbaceae, Desulfohalobiaceae, Desulfomicrobiaceae, Desulfonatronaceae, Desulfurobacteriaceae, Desulfuromonadaceae, Dictyoglomaceae, Enterococcaceae, Ferrovaceae, Fibrobacteraceae, Fimbriimonadaceae, Gaiellaceae, Gemmataceae, Geobacteraceae, Gloeobacteraceae, Gomontiellaceae, Hahellaceae, Halobacteriovoraceae, Holophagaceae, Hydrococcaceae, Hyellaceae, Hyphomonadaceae, Idiomarinaceae, Immundisolibacteraceae, Kangiellaceae, Kiloniellaceae, Kineosporiaceae, Kiritimatiellaceae, Kordiimonadaceae, Lachnospiraceae, Leptospiraceae, Leptotrichiaceae, Leuconostocaceae, Listeriaceae, Marinifilaceae, Mariprofundaceae, Melioribacteraceae, Methylacidiphilaceae, Methylophilaceae, Microbacteriaceae, Microcystaceae, Micromonosporaceae, Moraxellaceae, Nitriliruptoraceae, Nitrospinaceae, Oceanospirillaceae, Odoribacteraceae, Orbaceae, Oxalobacteraceae, Parvularculaceae, Pasteurellaceae, Pelagibacteraceae, Peptostreptococcaceae, Planctomycetaceae, Porphyromonadaceae, Porticoccaceae, Prevotellaceae, Prochloraceae, Pseudoalteromonadaceae, Pseudomonadaceae, Rhizobiaceae, Rhodobiaceae, Rhodocyclaceae, Rhodospirillaceae, Rickettsiaceae, Rikenellaceae, Rubrobacteraceae, Shewanellaceae, Sphingobacteriaceae, Spirulinaceae, Spongiibacteraceae, Staphylococcaceae, Streptococcaceae, Sutterellaceae, Synechococcaceae, Syntrophorhabdaceae, Thermoflexaceae, Trueperaceae, Vibrionaceae, Vulgatibacteraceae, Wenzhouxiangellaceae, Williamsiaceae, Xanthomonadaceae, Xenococcaceae*	*Bacillaceae, Caulobacteraceae, Gemmataceae, Gemmatimonadaceae, Hapalosiphonaceae, Ktedonobacteraceae, Methylobacteriaceae, Moraxellaceae, Nostocaceae, Paenibacillaceae, Rhodospirillaceae, Sphingobacteriaceae, Synechococcaceae, Xanthobacteraceae*
Location + time	*Acidothermaceae, Actinospicaceae, Aerococcaceae, Akkermansiaceae, Aquificaceae, Bacteriovoracaceae, Balneolaceae, Bifidobacteriaceae, Borreliaceae, Budviciaceae, Caldicoprobacteraceae, Carnobacteriaceae, Chitinophagaceae, Chroococcidiopsidaceae, Conexibacteraceae, Coriobacteriaceae, Crocinitomicaceae, Cryomorphaceae, Cryptosporangiaceae, Demequinaceae, Dermatophilaceae, Desulfarculaceae, Desulfurellaceae, Eubacteriaceae, Ferrimonadaceae, Flavobacteriaceae, Frankiaceae, Gracilibacteraceae, Halomonadaceae, Helicobacteraceae, Hydrogenothermaceae, Kosmotogaceae, Lentimicrobiaceae, Limnochordaceae, Marinilabiliaceae, Methylobacteriaceae, Methylothermaceae, Natranaerobiaceae, Oscillospiraceae, Patulibacteraceae, Prolixibacteraceae, Propionibacteriaceae, Proteinivoraceae, Pseudanabaenaceae, Pseudonocardiaceae, Puniceicoccaceae, Rubritaleaceae, Salinisphaeraceae, Saprospiraceae, Schleiferiaceae, Solirubrobacteraceae, Streptomycetaceae, Streptosporangiaceae, Syntrophobacteraceae, Syntrophomonadaceae, Thermaceae, Thermithiobacillaceae, Thermoleophilaceae, Thermomicrobiaceae, Thermomonosporaceae, Thermotogaceae, Waddliaceae, Woeseiaceae*	-

Location x time	*Acaryochloridaceae, Acetobacteraceae, Acidaminococcaceae, Acidimicrobiaceae, Acidithiobacillaceae, Acidobacteriaceae, Aeromonadaceae, Alicyclobacillaceae, Archangiaceae, Ardenticatenaceae, Atopobiaceae, Bacillaceae, Bacillales Family X. Incertae Sedis, Bdellovibrionaceae, Beijerinckiaceae, Bradyrhizobiaceae, Caldilineaceae, Cardiobacteriaceae, Catalimonadaceae, Chloroflexaceae, Chlorogloeopsidaceae, Christensenellaceae, Chthoniobacteraceae, Chthonomonadaceae, Clostridiaceae, Clostridiales Family XIII. Incertae Sedis, Clostridiales Family XVII. Incertae Sedis, Coleofasciculaceae, Competibacteraceae, Criblamydiaceae, Cyclobacteriaceae, Cytophagaceae, Dehalococcoidaceae, Desulfovibrionaceae, Ectothiorhodospiraceae, Eggerthellaceae, Enterobacteriaceae, Erwiniaceae, Erysipelotrichaceae, Erythrobacteraceae, Fervidobacteriaceae, Flammeovirgaceae, Fortieaceae, Francisellaceae, Fusobacteriaceae, Gallionellaceae, Gemmatimonadaceae, Geodermatophilaceae, Glycomycetaceae, Godleyaceae, Hafniaceae, Halanaerobiaceae, Halieaceae, Halobacteroidaceae, Halothiobacillaceae, Hapalosiphonaceae, Heliobacteriaceae, Hydrogenophilaceae, Hymenobacteraceae, Hyphomicrobiaceae, Ignavibacteriaceae, Isosphaeraceae, Kofleriaceae, Ktedonobacteraceae, Labilitrichaceae, Lactobacillaceae, Legionellaceae, Lentisphaeraceae, Leptolyngbyaceae, Merismopediaceae, Methylococcaceae, Methylocystaceae, Microcoleaceae, Micropepsaceae, Microthrixaceae, Morganellaceae, Mycobacteriaceae, Mycoplasmataceae, Myxococcaceae, Nannocystaceae, Neisseriaceae, Nitrosomonadaceae, Nitrospiraceae, Nostocaceae, Opitutaceae, Oscillatoriaceae, Oscillochloridaceae, Paenibacillaceae, Parachlamydiaceae, Pectobacteriaceae, Peptococcaceae, Persicobacteraceae, Phycisphaeraceae, Phyllobacteriaceae, Piscirickettsiaceae, Planococcaceae, Polyangiaceae, Rhodanobacteraceae, Rhodothermaceae, Rivulariaceae, Roseiflexaceae, Sandaracinaceae, Scytonemataceae, Selenomonadaceae, Sinobacteraceae, Solibacteraceae, Sphaerobacteraceae, Spirochaetaceae, Spiroplasmataceae, Sporichthyaceae, Sporolactobacillaceae, Sporomusaceae, Succinivibrionaceae, Sulfuricellaceae, Symbiobacteriaceae, Symphyonemataceae, Thermoactinomycetaceae, Thermoanaerobacteraceae, Thermoanaerobacterales Family III. Incertae Sedis, Thermoanaerobacterales Family IV. Incertae Sedis, Thermodesulfobacteriaceae, Thermodesulfobiaceae, Thermogemmatisporaceae, Tissierellaceae, Tolypothrichaceae, Veillonellaceae, Verrucomicrobia subdivision 3, Verrucomicrobia subdivision 6, Verrucomicrobiaceae, Xanthobacteraceae*	*Acetobacteraceae, Acidobacteriaceae, Archangiaceae, Beijerinckiaceae, Burkholderiaceae, Chamaesiphonaceae, Chloroflexaceae, Coleofasciculaceae, Cytophagaceae, Enterobacteriaceae, Erythrobacteraceae, Hyphomicrobiaceae, Isosphaeraceae, Leptolyngbyaceae, Microcoleaceae, Myxococcaceae, Opitutaceae, Oscillatoriaceae, Phyllobacteriaceae, Scytonemataceae, Sinobacteraceae, Sphingomonadaceae, Tolypothrichaceae*

Table S4 Influence of soil substrate (location), incubation time (time) and their interaction, on the relative abundances of bacterial reads assigned to COG functional categories, as determined by a robust 2-way ANOVA. Effect sizes (ω^2) and significance levels were calculated based on triplicates (n = 3). The "nuclear structure" category was not included in the statistical analysis due to the low number of reads. Significance levels are represented by the amount of stars: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$.

ANOVA (omega squared and significance levels)						
	Location		Time		Location x Time	
Cell cycle control and mitosis	0.01	ns	0.36	***	0.40	***
Cell wall/membrane/envelope biogenesis	0.07	**	0.74	***	0.08	ns
Cell motility	0.00	ns	0.66	***	0.03	ns
Post-translational modification, protein turnover, and chaperones	0.06	*	0.75	***	0.10	ns
Signal transduction	0.09	***	0.80	***	0.09	***
Intracellular trafficking, secretion, and vesicular transport	0.14	**	0.41	***	0.22	ns
Defense mechanisms	0.01	ns	0.96	***	0.00	ns
Extracellular structures	0.16	***	0.65	***	0.00	ns
Cytoskeleton	0.32	***	0.41	***	0.15	***
RNA processing and modification	0.13	*	0.65	***	0.03	ns
Chromatin structure and dynamics	0.33	***	0.11	**	0.14	ns
Translation, ribosomal structure and biogenesis	0.02	ns	0.00	ns	0.32	ns
Transcription	0.29	*	0.01	ns	0.16	ns
Replication, recombination and repair	0.29	***	0.33	***	0.17	*
Energy production and conversion	0.30	*	0.39	***	0.05	ns
Amino acid transport and metabolism	0.22	**	0.48	***	0.03	ns
Nucleotide transport and metabolism	0.00	ns	0.20	**	0.17	ns
Carbohydrate transport and metabolism	0.09	***	0.83	***	0.02	ns
Coenzyme transport and metabolism	0.00	ns	0.62	***	0.14	ns
Lipid transport and metabolism	0.44	**	0.18	*	0.05	ns
Inorganic ion transport and metabolism	0.04	ns	0.40	*	0.22	ns
Secondary metabolites biosynthesis, transport, and catabolism	0.03	ns	0.48	***	0.02	ns
Function unknown	0.27	***	0.51	***	0.15	***

Table S5 Influence of soil substrate (location), incubation time (time) and their interaction, on the relative abundances of bacterial reads assigned to genes encoding for proteins involved in EPS and LPS synthesis and excretion, as determined by a robust 2-way ANOVA. Effect sizes (ω^2) and significance levels were calculated based on triplicates (n = 3). Significance levels are represented by the amount of stars: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$.

ANOVA (omega squared and significance levels)						
	Location		Time		Location x Time	
wza	0.05	*	0.81	***	0.04	*
wcaB	0.11	ns	0.66	***	0.06	ns
wcaF	0.20	***	0.59	***	0.07	***
wcaK/amsJ	0.00	ns	0.48	ns	0.00	ns
kpsE	0.18	*	0.53	***	0.01	ns
algJ	0.07	ns	0.09	ns	0.00	ns
sacB	0.13	ns	0.00	ns	0.01	ns
wzt	0.23	**	0.33	***	0.33	***
lptC	0.05	ns	0.57	***	0.00	ns
lptF	0.21	ns	0.19	ns	0.12	ns
lptG	0.17	*	0.32	*	0.13	ns

License agreement for Publication 1

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS
Nov 09, 2020

This Agreement between Ms. Barbara Cania ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4901370153171
License date	Sep 03, 2020
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Microbial Ecology
Licensed Content Title	Biological Soil Crusts from Different Soil Substrates Harbor Distinct Bacterial Groups with the Potential to Produce Exopolysaccharides and Lipopolysaccharides
Licensed Content Author	Barbara Cania et al
Licensed Content Date	Aug 1, 2019
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	full article/chapter
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	yes
Title	Bacterial polysaccharides as drivers of soil aggregation
Institution name	Technische Universität München
Expected presentation date	Sep 2020
Requestor Location	Ms. Barbara Cania Moorkamp 2 Kiel, 24106 Germany Attn: Ms. Barbara Cania
Total	0.00 EUR

B Publication 2

Cania *et al. Environmental Microbiome* (2019) 14:1
https://doi.org/10.1186/s40793-019-0341-7

Environmental Microbiome

RESEARCH ARTICLE Open Access

A long-term field experiment demonstrates the influence of tillage on the bacterial potential to produce soil structure-stabilizing agents such as exopolysaccharides and lipopolysaccharides

Barbara Cania[1], Gisle Vestergaard[1,4], Maike Krauss[2], Andreas Fliessbach[2], Michael Schloter[1,3] and Stefanie Schulz[1*]

Abstract

Background: Stable soil aggregates are essential for optimal crop growth and preventing soil erosion. However, tillage is often used in agriculture to loosen the soil, which disrupts the integrity of these aggregates. Soil aggregation can be enhanced by bacteria through their ability to produce exopolysaccharides and lipopolysaccharides. These compounds stabilize soil aggregates by "gluing" soil particles together. However, it has yet to be shown how tillage influences the bacterial potential to produce aggregate-stabilizing agents. Therefore, we sampled conventional and reduced tillage treatments at 0–10 cm, 10–20 cm and 20–50 cm from a long-term field trial in Frick, Switzerland. We compared the stable aggregate fraction of the soil and the bacterial potential to produce exopolysaccharides (EPS) and lipopolysaccharides (LPS) under different tillage regimes by employing a shotgun metagenomic approach. We established a method which combines hidden Markov model searches with blasts against sequences derived from the Kyoto Encyclopedia of Genes and Genomes database to analyze genes specific for the biosynthesis of these compounds.

Results: Our data revealed that the stable aggregate fraction as well as the bacterial potential to produce EPS and LPS were comparable under both tillage regimes. The highest potential to produce these compounds was found in the upper soil layer, which was disturbed by tillage, but had higher content of organic carbon compared to the layer below the tillage horizon. Additionally, key players of EPS and LPS production differed at different sampling depths. Some families with high potential to produce EPS and LPS, such as *Chitinophagaceae* and *Bradyrhizobiaceae*, were more abundant in the upper soil layers, while others, e.g. *Nitrospiraceae* and *Planctomycetaceae*, preferred the lowest sampled soil depth. Each family had the potential to form a limited number of different aggregate-stabilizing agents.

Conclusions: Our results indicate that conventional tillage and reduced tillage equally promote the bacterial potential to produce EPS and LPS in the tillage horizon. However, as major bacterial groups triggering EPS and LPS formation were not the same, it is likely that gene expression pattern differ in the different treatments due to various pathways of gene induction and transcription in different bacterial species.

Keywords: Tillage, Soil aggregates, Exopolysaccharides, Lipopolysaccharides, Soil microbiome, Metagenomics, *wza*, *lptF*, *lptG*

* Correspondence: stefanie.schulz@helmholtz-muenchen.de
[1]Research Unit Comparative Microbiome Analysis, Helmholtz Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH), Ingolstädter Landstraße 1, 85764 Neuherberg, Germany
Full list of author information is available at the end of the article

Background

Globally, 33% of land resources have been classified as moderately to highly degraded [1]. The main causes of soil degradation are poor agricultural management practices, such as conventional tillage (CT), which lead to erosion, loss of soil organic carbon and nutrient imbalance [2]. It turns out that a combination of reduced tillage (RT) and organic farming (OF) is a good compromise to diminish the aforementioned problems [3–5]. However, RT is still not commonly used by organic farmers due to increased weed pressure, topsoil compaction and restricted N availability, which may compromise yield [6, 7]. As even one-time ploughing may counteract the benefits of RT, these practices need to be developed further under long-term OF [8, 9].

One of the advantages of RT over CT practices is the better preservation of soil aggregates [10, 11]. The presence of stable aggregates defines good soil structure, which improves crop growth and prevents erosion [12, 13]. The stability of aggregates strongly depends on their size. Microaggregates (< 250 μm) form slower than macroaggregates (> 250 μm), but they are also more stable, even under unfavorable soil management systems [14, 15]. Aggregate formation results from complex interactions between soil fauna, microorganisms, roots, inorganic binding agents and different environmental variables. Fungi have been considered as the most important microorganisms involved in the formation of macroaggregates due to their hyphal structure [15, 16]. In contrast, bacteria are of higher importance for soil aggregation at the microscale, as they are capable of synthesizing exopolysaccharides (EPS) and lipopolysaccharides (LPS), which act as "glue" for soil particles [15, 16]. Bacteria use these compounds for cell attachment to mineral surfaces, which fosters the formation of composite building units and microaggregates [15–17]. While EPS are a very diverse group of high-molecular-weight polymers composed of sugar residues, LPS share a common structure. The number of possible EPS structures is almost infinite [18]. Most EPS are initially synthesized intracellularly and then secreted to the external environment, which requires the contribution of at least three gene families: I) genes encoding for enzymes involved in biosynthesis of nucleotide sugars, II) genes encoding for glycosyltransferases, which catalyze transfer of the nucleotide sugars from activated donor molecules to specific acceptors in the plasma membrane, and III) genes encoding for proteins involved in EPS assembly and export [19]. Alternatively, EPS can be synthesized extracellularly by different synthase proteins [20]. Most enzymes involved in the EPS biosynthesis are strain-specific and can catalyze multiple metabolic processes.

LPS are glycolipids that are comprised of a lipid moiety (lipid A) and a polysaccharide (composed of O-antigen, outer core and inner core), both with variable structures [21]. These parts are synthesized independently inside a cell, and then ligated together at the inner membrane, forming a mature LPS. The mature molecule is transported to the cell surface by several proteins that form an LPS export complex. As in EPS biosynthesis, very few of these proteins are conserved and catalyze LPS production only [22].

The gluing properties of both types of polysaccharides could be crucial in agricultural soils, as it was demonstrated that even slight changes in the sugar composition drastically changed the physical properties of the polysaccharide [18]. Consequently, tillage might not only change the bacterial community composition in soil [23], but also the composition of EPS/LPS, and thus affect aggregate stability and de novo formation after disturbance.

The synthesis of EPS and LPS requires both high levels of energy and easily accessible carbon. Especially under CT, reduced soil organic carbon stocks have been frequently observed [4, 23, 24]. Therefore, we hypothesized that under long-term CT, abundance of EPS and LPS forming bacteria would be reduced compared to RT. To investigate this, high-throughput shotgun sequencing was used to obtain metagenomic information on microbiomes of three soil layers (0–10 cm, 10–20 cm and 20–50 cm) under RT and CT management from a long-term organic field trial in Frick (Switzerland). To analyze genes specific for EPS and LPS production, we used an approach which combined hidden Markov model (HMM) searches with blasts against sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. As we investigated bacterial potentials samples were taken in spring where an influence of plants and fertilization could be excluded.

Materials and methods

Site description and soil sampling

Soil samples were taken from a long-term trial in Frick, Switzerland (47°30′N, 8°01′E, 350 m a.s.l.), established in 2002 by the Research Institute of Organic Agriculture (FiBL). The site was under conventional management until 1995, when it changed to organic standards in accordance with the European Union Regulation (EEC) No. 2092/91. The mean annual precipitation and temperature are 1000 mm and 8.9 °C, respectively. The soil is a Stagnic Eutric Cambisol with a pH of 7.1 and composed of 22% sand, 33% silt and 45% clay. The factorial design includes the factors tillage, fertilisation and biodynamic preparations and has been described in detail by Berner et al. [3].

In this study, only the two tillage treatments were compared: conventional tillage (CT) with a mouldboard plough operating at 15–18 cm depth, and reduced tillage (RT) with a chisel and a skim plough (5–10 cm) used to

loosen the soil. In both systems, seedbed preparation was performed with a rotary harrow running at a depth of 5 cm. The usage of standard farming equipment was made possible by the plot size (12 m × 12 m). The plots were arranged in a strip-split-plot design.

Samples were taken from three out of four replicated plots per tillage system in the slurry fertilized plots without biodynamic preparations in March 2015 in a green manure ley, before tillage and subsequent maize cropping started. In 2014, winter wheat was harvested in July, followed by the seeding of a green manure mixture (Orgamix DS, *Trifolium incarnatum, Vicia villosa, Avena sativa*) in August, which was harvested in April 2015. All plots were fertilized with slurry during the wheat growing season in 2014 (the exact dates and fertilization details are summarized in Additional file 1). Soil samples were taken using a soil auger with a diameter of 2.5 cm. Approximately 10 cores per plot were sampled to a soil depth of 50 cm. Each soil core was divided into three layers: 0–10 cm, 10–20 cm and 20–50 cm. Samples from the same layer of each plot were homogenized, resulting in 18 samples (3 depths × 2 tillage treatments × 3 plot replicates). The samples were directly cooled in the field and either processed immediately (biochemical analyses) or stored in − 20 °C until processing (DNA extraction and sequencing).

Physical, chemical and major biological properties of soils

We determined the stable aggregate fraction (SAF) of the soils by a wet sieving technique, where 5 g of moist soil was immersed in water using a sieving apparatus according to Murer et al. [25]. After 5 min of moving the sieves up and down in the water phase, the remainder on the sieve consisting of aggregates and particles > 0.25 mm was dried at 105 °C. The aggregates were then destroyed by adding a 0.1 M $Na_4P_2O_7$ solution, leaving only particles > 0.25 mm (sand and organic debris) on the sieve that were dried again. Apart from using moist soil without further fractionation, the method follows the details as given by Murer et al. [25].

Soil organic carbon (SOC) concentration was determined by wet oxidation of 1 g of air-dried and ground soil in 20 ml of concentrated H_2SO_4 and 25 ml of 2 M $K_2Cr_2O_7$. The determination of dissolved organic carbon (DOC) and microbial biomass carbon (Cmic) was accomplished by means of a chloroform fumigation extraction method (CFE) using 20 g of moist soil, sieved on a 5 mm sieve. 0.5 M K_2SO_4 solution was added at a weight to volume (*w*/*v*) ratio of 1:4. Subsequently, measurements were performed using a TOC/TNb analyzer (Analytik Jena AG, Germany). DOC was determined from the non-fumigated samples, and Cmic was calculated as a difference between the fumigated and the non-fumigated samples. The assessment of SOC and Cmic was described in detail by Krauss et al. [23].

DNA extraction and sequencing

Total nucleic acids were directly extracted from 0.5 g of frozen soil according to the phenol-chloroform based DNA/RNA coextraction protocol described by Lueders et al. [26]. Beat beating was performed by means of CKMix tubes and a Precellys24 homogenizer (Bertin Technologies, France). Extracted DNA was checked for purity using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, USA). The quantity was also verified by means of a Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies, USA). Extracted DNA was then stored in − 20 °C until further processing.

One microgram of DNA from each sample was sheared using an E220 Focused-ultrasonicator (Covaris, USA), following the manufacturer's guideline for the target size of 500 bp (conditions: peak incident power – 175 W, duty factor – 5%, cycles per burst – 200, treatment time – 35 s, temperature – 7 °C, water level – 6, sample volume – 50 μl, intensifier – yes). Libraries were prepared with 50–100 ng of the sheared DNA, using a NEBNext Ultra DNA Library Prep Kit for Illumina, and NEBNext Multiplex Oligos for Illumina (New England Biolabs, UK) as barcodes. According to the manufacturer's manual, the NEBNext Adaptor from Illumina was diluted 10-fold to prevent the occurrence of dimers. Size selection was performed with Agencourt AMPure XP beads (Beckman Coulter, USA), using the volumes selecting for libraries with 400–500 bp inserts. The AMPure XP beads were also used for cleanup of PCR amplification and a following additional cleanup step to eliminate the residual primer dimers (1:0.6 DNA to bead ratio).

Library size was estimated using High Sensitivity DNA Analysis Kits together with a 2100 Bioanalyzer (Agilent, USA). DNA concentration was subsequently assessed by means of a Quant-iT PicoGreen dsDNA Assay Kit. Libraries were then diluted to a concentration of 4 nM each and pooled equimolar. 10 pM of the mixture was spiked with 30% PhiX, used as a quality and calibration control [27], and sequenced on a MiSeq sequencer using a MiSeq Reagent Kit v3 for 600 cycle (Illumina, USA).

Data filtering and taxonomic analysis

Raw sequencing data attained from the MiSeq was filtered according to Vestergaard et al. [28] by removing remnant adaptor sequences and trimming the reads. This was accomplished by using AdapterRemoval [29] set to: 5′/3′ terminal minimum Phred quality = 15, minimum read length = 50. PhiX contamination was removed using DeconSeq [30]. For taxonomic annotation, filtered reads were blasted against the National Center for Biotechnology Information Non-Redundant (NCBI-NR)

protein sequences database (October 2015) using Diamond (version 0.5.2.32) with sensitive parameters [31]. Based on the top 25 blast results (i.e. hits with the lowest e-value), a unique taxon ID was assigned to each filtered read with the MEtaGenome Analyzer software (MEGAN, version 5.10.6) [32]. During the MEGAN analysis, the following parameters were applied: MinScore = 50.0, MaxExpected = 0.01, TopPercent = 10.0, MinSupport = 1, MinComplexity = 0. Additionally, 16S rRNA gene sequences were identified using SortMeRNA (version 2.0) [33]. Taxonomy was assigned to those reads using QIIME (version 1.9.1) [34] based on the SILVA database (release 123).

Functional analysis

Protein sequences associated with EPS and LPS biosynthesis and excretion were downloaded directly from the online Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database (October 2016). They were examined for the presence of function-specific conserved domains using CD-search [35]. KEGG Orthology (KO) entries which contained such domains were then used to construct specific databases by means of Diamond. Hidden Markov models (HMMs) of corresponding conserved domains were obtained from the TIGRFAMs database (version 15) [36] and the Pfam database (version 30) [37]. FragGeneScan (version 1.19) [38] was used on the filtered sequencing reads to predict open-reading frames, which were subsequently scanned with HMMER (version 3) (hmmer.org). Reads matching the downloaded HMMs (E-value threshold = 10^{-5}), were blasted against the self-built KO databases. A KO ID was assigned to those reads for which the top 25 blast results were consistent. The specificity of this approach was verified by using blastx against the Non-redundant protein sequences (nr) database. Out of 81 examined KO numbers (67 for EPS and 14 for LPS), 14 gave sufficiently specific results. The results were considered sufficiently specific if 25 randomly selected reads (or all if less reads were assigned) per a KO number were assigned to the function of interest. Analysis of EPS and LPS biosynthesis and excretion was performed using separate databases. Open-reading frames of the assigned reads were searched against the full Pfam and TIGRFAMs databases. This resulted in 81.8% of the reads matching the downloaded HMMs. All examined KO numbers are listed in Additional file 2 and the HMMs and KO numbers used for the analysis are summarized in Table 1.

Statistical analysis and data visualization

All statistical analyses were conducted using R version 3 [39]. Metagenomic datasets were analyzed based on relative abundances of reads. These were obtained by dividing the number of reads assigned to a gene or organism by the total number of filtered reads per sample, and multiplying by 100. Effects of tillage, depth and their possible interaction were detected by multilevel models. For this purpose, the lme function from the nlme package was used [40]. The influence was considered significant when the p-value was below 5% ($P < 0.05$). Differences between sampled depths were identified by setting the following contrasts: 0–20 cm vs 20–50 cm and 0–10 cm vs 10–20 cm. For data derived from the metagenomic datasets, the Benjamini-Hochberg procedure was performed prior to analyzing contrasts. The Shannon-Wiener index was calculated using the alpha.div function of the R asbio package to measure diversity

Table 1 Proteins related to exo- and lipopolysaccharide production with corresponding KO numbers, HMM IDs and genes

Protein	KO number	HMM ID	Gene
polysaccharide export outer membrane protein Wza	K01991	PF02563	*wza*
alginate export outer membrane protein AlgE	K16081	PF13372	*algE*
alginate biosynthesis acetyltransferase AlgJ	K19295	PF16822	*algJ*
colanic acid biosynthesis acetyltransferase WcaB	K03819	TIGR04016	*wcaB*
colanic acid biosynthesis acetyltransferase WcaF	K03818	TIGR04008	*wcaF*
colanic acid/amylovoran biosynthesis pyruvyl transferase WcaK/AmsJ	K16710	TIGR04006	*wcaK/amsJ*
capsular polysaccharide export system permease KpsE	K10107	TIGR01010	*kpsE*
exopolysaccharide biosynthesis transmembrane protein EpsG	K19419	PF14897	*epsG*
exopolysaccharide biosynthesis tyrosine kinase modulator EpsA	K19420	TIGR01006	*epsA*
levansucrase SacB	K00692	PF02435	*sacB*
lipopolysaccharide transport system ATP-binding protein Wzt	K09691	PF14524	*wzt*
LptBFGC lipopolysaccharide export complex permease LptF	K07091	TIGR04407	*lptF*
LptBFGC lipopolysaccharide export complex permease LptG	K11720	TIGR04408, PF03739	*lptG*
LptBFGC lipopolysaccharide export complex inner membrane protein LptC	K11719	TIGR04409, PF06835	*lptC*

within the samples [41]. To visualize the level of dissimilarity between the samples, non-metric multidimensional scaling (NMDS) ordination plots were created based on the Bray-Curtis distance metrics, using the metaMDS function in the R vegan package [42]. The core microbiomes were identified by means of InteractiVenn [43]. For the purpose of calculating these cores, a family was recognized as present in a treatment only if it was detected in at least two out of three replicates.

Results

Soil properties

The stable aggregate fraction (SAF) of the soil, soil organic carbon (SOC), dissolved organic carbon (DOC) and microbial biomass carbon (Cmic) data is summarized in Table 2. Aggregate stability was highest in the 20–50 cm depth, and did not differ significantly between the upper depths. It was also not significantly influenced by tillage. SOC stocks decreased with depth, and were higher in the 0–20 cm depth under RT compared to CT. DOC concentrations were highest in the 0–10 cm depth under RT, and showed little difference between the other samples. Microbial biomass decreased with depth, and was more stratified under RT. In the 0–10 cm depth, Cmic values were higher under RT.

Shotgun sequencing characteristics

Shotgun sequencing of the 18 libraries, prepared from two tillage treatments – conventional (CT) and reduced (RT) – sampled at three depths (0–10 cm, 10–20 cm and 20–50 cm) from three independent plots treated as replicates, generated 11.8 gigabases of data in total. This corresponded to 39.307.875 filtered reads with an average length after trimming of 297 bp. Details of the sequencing run are summarized in Additional file 3.

Taxonomic analysis of the general bacterial community

When all filtered reads were blasted against the NCBI-NR database, 55.8% were assigned to Bacteria, 1.2% to Archaea, 1.3% to Fungi and 41.7% to others. Further analysis focused on bacteria and was conducted at the level of family, at which 21.1% of filtered reads were assigned. In total, bacteria comprised 296 families.

The non-metric multidimensional scaling (NMDS) ordination plot (Additional file 4A) showed a difference between the composition of bacterial families originating from the deepest sampled soil layer (20–50 cm) and the upper soil layers (0–10 cm and 10–20 cm), but revealed no clear separation of the tillage treatments. This was confirmed by means of a multilevel model. Abundances of 103 families were influenced by depth, while none was affected by tillage, and one by the interaction of both factors. The full list of impacted families can be taken from Additional file 5.

In-depth analysis of the effects of tillage, depth and their interaction on the general community structure was performed on dominant families whose abundance exceeded 0.5% (Fig. 1a). The most abundant family, *Anaerolineaceae*, together with *Nitrospiraceae*, were found mainly in 20–50 cm. *Chitinophagaceae*, *Bradyrhizobiaceae*, *Polyangiaceae* and *Cytophagaceae* had higher abundance in 0–20 cm. *Planctomycetaceae*, *Acidobacteriaceae*, *Verrucomicrobia subdivision 3*, *Flavobacteriaceae* and *Solibacteraceae* were not significantly influenced by either depth or tillage.

The results from the taxonomic analysis encompassing the entire metagenomic datasets were supported by SILVA's taxonomic annotations of the 16S rRNA gene. Of all filtered reads, 0.21% were assigned to the 16S rRNA gene. With both approaches, the bacterial communities showed similar distribution patterns and one third of the dominant families remained the same, with *Anaerolineaceae* staying the most abundant, regardless of the assignment method used (Additional file 6).

Relative abundances of genes catalyzing EPS and LPS synthesis and excretion

An approach combining hidden Markov model (HMM) searches with blasts against sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database was used to target genes specific for the biosynthesis and excretion of alginate, colanic acid, levan and other EPS, as well as LPS (Table 1). Sufficient coverage of the

Table 2 Stable aggregate fraction of the soil, carbon stocks and microbial biomass

Tillage	Conventional			Reduced		
Depth	0–10 cm	10–20 cm	20–50 cm	0–10 cm	10–20 cm	20–50 cm
SAF (%)*	56.19 ± 8.98	48.58 ± 3.05	65.21 ± 9.62	50.35 ± 8.30	51.52 ± 6.73	69.02 ± 2.26
SOC (%)#	2.30 ± 0.41	2.15 ± 0.33	1.25 ± 0.37	2.92 ± 0.28	2.31 ± 0.26	1.23 ± 0.22
DOC (mg kg^{-1})#	62.44 ± 9.79	53.61 ± 15.84	58.41 ± 10.23	99.49 ± 19.84	62.32 ± 7.62	51.97 ± 8.45
Cmic (mg kg^{-1})#	981.81 ± 158.92	849.18 ± 106.23	352.12 ± 121.46	1306.73 ± 122.07	932.26 ± 67.23	374.75 ± 58.22

SAF Stable aggregate fraction, *SOC* Soil organic carbon, *DOC* Dissolved organic carbon, and *Cmic* microbial biomass carbon values of soils under two tillage systems, sampled at three different depths. Average values and standard deviations (±) are calculated based on triplicates ($n = 3$). Detected influence ($p < 0.05$) is symbolized for depth (*) or interaction of tillage and depth (#), respectively

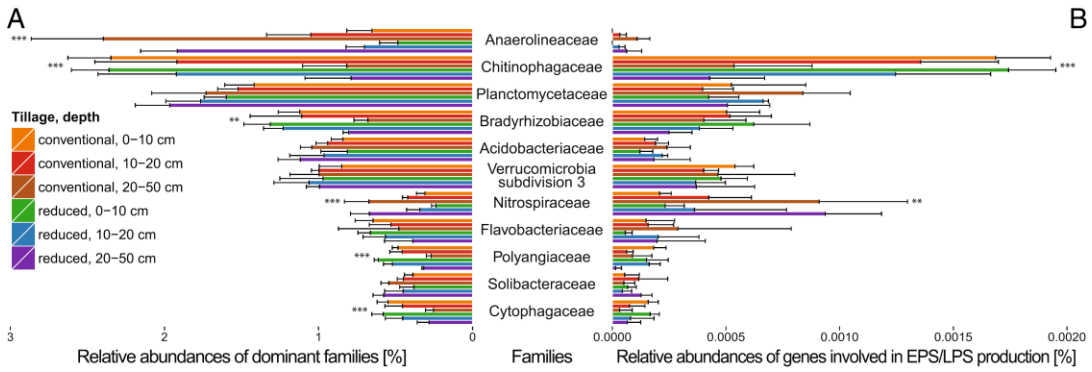


Fig. 1 Dominant families and their potential to produce EPS and LPS. Comparison of (A) abundances of dominant bacterial families whose relative abundances exceeded 0.5%, with (B) their potential to produce EPS and LPS. Taxonomic assignment was performed against the National Center for Biotechnology Information Non-Redundant (NCBI-NR) protein sequences database. Functional genes were assigned using hidden Markov models (HMMs) obtained from the TIGRFAMs and Pfam databases, and then sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database. Significant differences in the amount of annotated reads were determined by a multilevel model (n = 3). Detected influence is symbolized for depth (*) or interaction of tillage and depth (#), respectively. Significance levels are represented by the amount of symbols: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$. Error bars show standard deviation

diversity of the analyzed genes was confirmed by performing explanatory rarefaction analysis (Additional file 7).

In total, the investigated genes comprised 0.018% of all filtered reads (Fig. 2). Dominant genes, with a relative abundance above 0.005% of total reads, were *wza*, *lptF* and *lptG*, which encode for an outer membrane protein responsible for EPS excretion, and permeases of the LptBFGC LPS export system, respectively. Moderately abundant genes (> 0.001%) were *wcaB*, wcaF (encoding for a colanic acid biosynthesis acetyltransferases) and *wzt* (a gene which encodes for an ATP-binding protein of the LPS O-antigen transport system). Genes *algE*, *algJ*, *wcaK*/*amsJ*, *epsA*, *epsG*, *sacB* and *lptC* were the least abundant, with just a few reads annotated. Multilevel model analysis revealed depth as the main factor affecting the distribution pattern of the investigated

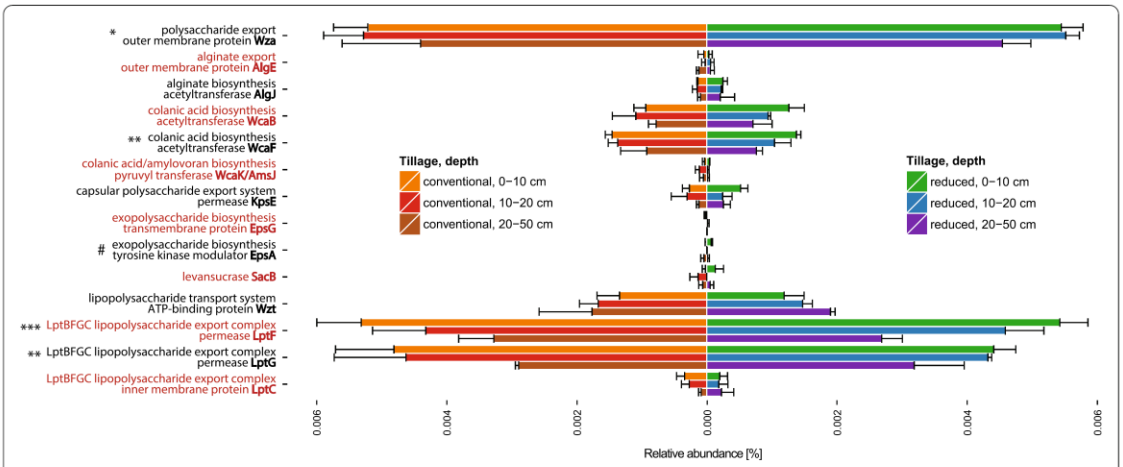


Fig. 2 Genes encoding for proteins involved in biosynthesis of EPS and LPS. Relative abundances of genes encoding for proteins involved in biosynthesis of EPS and LPS in soils under two tillage managements, sampled at three different depths. Functional genes were assigned using hidden Markov models (HMMs) obtained from the TIGRFAMs and Pfam databases, and then sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database. Significant differences in the amount of annotated reads were determined by a multilevel model (n = 3). Detected influence is symbolized for depth (*) or interaction of tillage and depth (#), respectively. Significance levels are represented by the amount of symbols: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$. Error bars show standard deviation

genes. Specifically, the relative abundance of *wcaF* and *lptFG* decreased with depth by half. In addition, the *epsA* gene (encoding for an EPS biosynthesis tyrosine kinase modulator) was influenced by interaction of depth and tillage. This gene was more abundant in 0–10 cm under RT, compared to 20–50 cm under CT, and no reads were detected in 10–20 cm under both CT and RT. The majority of the analyzed genes, namely *wza*, *algEJ*, *wcaB*, *wcaK/amsJ*, *kpsE*, *epsG*, *sacB*, *wzt* and *lptC*, were not significantly affected by either tillage or depth.

Investigation of potential EPS/LPS producers

One hundred thirty-eight bacterial families harbored the investigated genes, including all dominant families (Fig. 1b). The highest numbers of sequences related to EPS and LPS synthesis and excretion (> 0.001%) were assigned to *Chitinophagaceae*, *Nitrospiraceae* and *Planctomycetaceae*. *Anaerolineaceae*, despite their high abundance, harbored a very low number of copies of the investigated genes (< 0.0002%).

The NMDS (Additional file 4B) plot once again revealed depth as the main factor affecting the distribution of the investigated genes among bacterial families. However, the influence of depth was much less pronounced than in case of the general bacterial community (Additional file 4A). This was confirmed with a multilevel model. The overall relative abundances of the investigated genes were impacted by depth in four families affiliated with EPS/LPS synthesis and excretion, while tillage had no influence, and interaction had an effect on one family only. The full list of influenced families can be taken from Additional file 5.

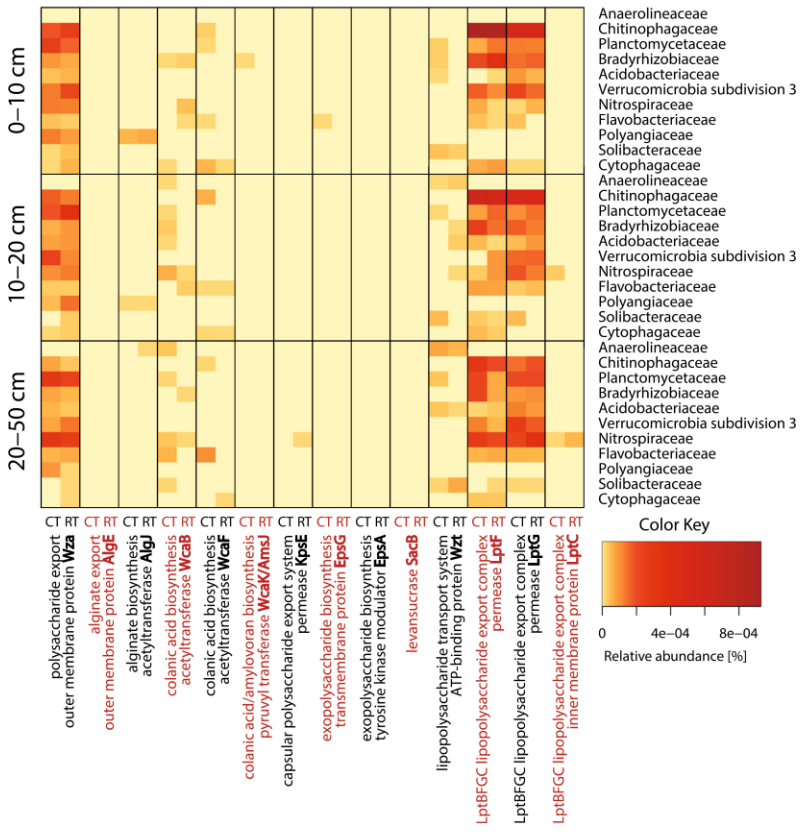


Fig. 3 Genes encoding for proteins involved in EPS and LPS biosynthesis found in the dominant families. Heatmap representation of the mean relative numbers of genes encoding for proteins involved in EPS and LPS synthesis and excretion found in the dominant bacterial families whose abundance exceeded 0.5% in the samples from Frick taken at three depths. "CT" and "RT" on the x axis stand for "conventional tillage" and "reduced tillage", respectively. Functional genes were assigned using hidden Markov models (HMMs) obtained from the TIGRFAMs and Pfam databases, and then sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database. Taxonomic assignment was performed against the National Center for Biotechnology Information Non-Redundant (NCBI-NR) protein sequences database

Taxonomic affiliation of the individual genes encoding for proteins involved in EPS and LPS biosynthesis was analyzed using a heatmap (Fig. 3). The most abundant genes, *wza* and *lptFG*, were harbored by most of the dominant families. *Anaerolineaceae* had neither *wza* nor *lptFG*, but harbored the *wzt* gene, which is part of the same LPS synthesis pathway. Moreover, *Polyangiaceae* carried the *wza* gene, but showed no potential to produce LPS. The other investigated genes were not so widely distributed among the dominant families. In particular, *algE*, *epsA* and *sacB*, encoding respectively for alginate export outer membrane protein, exopolysaccharide biosynthesis tyrosine kinase modulator and levansucrase, were not detected in any of the dominant families. As shown by means of a multilevel model, the gene copy numbers of *wcaF*, *epsA*, *sacB*, *wzt* and *lptC* were influenced by interaction of tillage and depth in *Chitinophagaceae*, *Bacillaceae*, *Micrococcaceae*, *Candidatus Brocadiaceae* and *Sulfuricellaceae*, whilst the abundances of *algJ* and *lptC* changed with depth in *Polyangiaceae* and *Sphingomonadaceae*, respectively.

Since *wza* and *lptFG* were dominating among the investigated genes, their taxonomic affiliation was analyzed in more detail. These genes were present in a total of 50 families associated with 11 phyla. The core microbiomes harboring the respective genes under both tillage treatments were identified at each sampled depth (Fig. 4). At each depth, on average ten families carried the respective genes under both tillage managements, while five were unique for either CT or RT. Overall, the three genes harbored by the core families accounted for 22.7% of all reads assigned to all the investigated genes, while 1.8 and 2.1% were unique for CT and RT, respectively. The diversity of families carrying *wza* and *lptG* significantly decreased with depth (Additional file 8). Depth triggered a decrease of *wza* and *lptFG* in *Chitinophagaceae*. The relative number of *wza* gene copies decreased with depth also in *Flammeovirgaceae* and *Labilitrichaceae*. Furthermore, depth caused a decrease of *lptG* in *Bdellovibrionaceae*, but *lptF* increased in *Nitrospiraceae*. Finally, the interaction of depth and tillage affected *lptG* in *Pseudomonadaceae*. This gene was more abundant in 0–10 cm under RT, compared to 20–50 cm under CT, and no reads were detected in the other samples.

Discussion

Different factors could affect the stable aggregate fraction in soil

EPS and LPS are of great importance for agricultural soils, as they reduce soil erodibility by improving soil structure [44]. However, tillage disrupts soil aggregates and alters soil physical and chemical properties. These include bulk density, pore structure, water availability and soil organic carbon [45]. Thus, changes in bacterial communities are likely to occur. This has been reported by multiple studies [46–50]. Especially CT disturbs bacterial habitats and dilutes nutrient pools by mixing topsoil with subsoil. In our study, soil organic carbon (SOC), dissolved organic carbon (DOC) and microbial biomass carbon (Cmic) had higher values in the tillage horizon under RT compared to CT. This corresponds to the data found in the literature [23, 51–53]. The increase of Cmic suggests that the absolute number of bacteria capable of synthesizing EPS and LPS should be higher under RT. Thus, we assumed that the higher DOC concentrations promotes bacteria which are able to produce EPS and LPS, and that the stable aggregate fraction (SAF) of the soil is higher under RT. Surprisingly, at our sampling site, SAF was comparable between the two tillage systems and increased significantly only below the tillage horizon. However, this might be caused by soil physical properties. Specifically, the clay content (45%) was very high at our site. Meta-analysis performed by Cooper et al. [7] suggests that the differences between tillage systems could be more pronounced in soils with a lower clay content (< 40%). Building good soil structure is more challenging in light, sandy soils, as they lack the fine particles necessary to form stable soil aggregates [54]. Conversely, soil biology has a strong influence on SAF. This includes the activity of bacteria, fungi, earthworms and plants. On one hand, the effect of plants and earthworms is rather indirect and includes for example cast formation by earthworms or increasing microbial activity by the release of organic substances to the soil via the rhizosphere of plants [15, 16]. On the other hand, bacteria and fungi directly promote aggregate formation by the excretion of gluing agents such as EPS, LPS and fungal glycoproteins, or by physical binding of soil particles by fungal mycelium [15]. Similar to the general increase of Cmic in the topsoil under RT, Kuntz et al. [55] also observed higher fungal abundances in that soil layer.

While it is obvious that ploughing physically disturbs fungal hyphae and consequently aggregates connected to them, the effect on the bacterial potential to promote aggregate formation can be much more subtle. Especially EPS composition and regulation of the respective genes is species-specific, thus a shift in the bacterial community strongly influences their potential to promote aggregate formation. To detect changes in the bacterial potential to produce EPS and LPS, we applied a metagenomic approach. As many proteins or their functional domains from genes encoding for EPS and LPS biosynthesis pathways are associated with other cellular

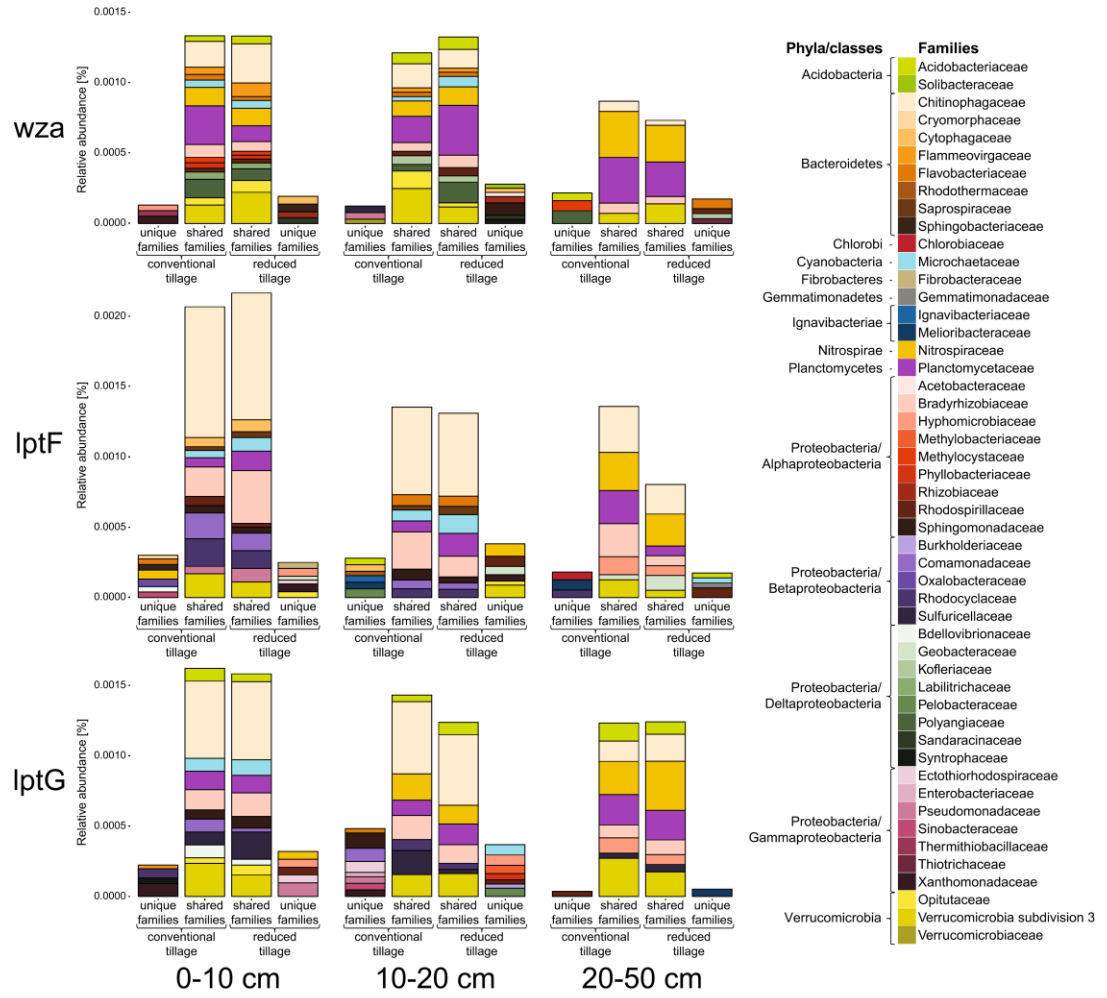


Fig. 4 Taxonomic assignment of the *wza*, *lptF* and *lptG* genes. Mean relative abundances of the *wza*, *lptF* and *lptG* genes, encoding respectively for a polysaccharide outer membrane exporter and permeases of the LptBFGC LPS export complex, in the samples taken at three depths from plots under conventional and reduced tillage managements. Displayed are the distributions of gene copies among bacterial families listed on the right side of the graph. Only families found in at least two out of three replicates are presented. The color code is arranged according to the phylogenetic affiliation of the respective families. Compared are gene abundances in families harboring the respective genes uniquely under conventional or reduced tillage and families harboring the genes under both tillage managements. Functional genes were assigned using hidden Markov models (HMMs) obtained from the TIGRFAMs and Pfam databases, and then sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database. Taxonomic assignment was performed against the National Center for Biotechnology Information Non-Redundant (NCBI-NR) protein sequences database

activities as well [56], we used a pipeline combining hidden Markov model (HMM) searches with blasts against sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database to target selected genes specific for our functions of interest. Although the contribution of fungi to aggregate formation is well accepted, our analysis exclusively focused on bacteria due to the well-described biases of the existing databases towards bacteria [57]. This is also visible in our dataset, where 55.8% of sequences were assigned to bacteria, while only 1.3% could be assigned to fungi. Moreover, fungal genes require long reads for accurate annotation due to many intronic sequences [58].

Key genes encoding for selected EPS and LPS biosynthesis pathways

Identified as key components of the analyzed EPS and LPS synthesis and excretion pathways were *wza* and *lptFG*, which encode for an outer membrane protein Wza and permeases of the LptBFGC LPS export complex (LptF and LptG), respectively. Wza acts as a translocation channel across the outer membrane for a variety of exopolysaccharides in a wide range of taxa. It is also characterized by the presence of a very well conserved polysaccharide export sequence domain (pfam 02563) [20]. Similarly, LptF and LptG are essential for transport of mature LPS to the outer membrane. These two proteins are highly conserved among Gram-negative bacteria, unlike another component of the LptBFGC transport complex, LptC [59, 60]. Benedet et al. [61] recently reported the isolation of mutants lacking LptC and suggested its supportive role in the LPS translocation. In our study, the respective gene, *lptC*, had just a few reads annotated and was detected in only one of the dominant families, *Nitrospiraceae*, even though all of them belong to phyla known to produce LPS [62]. The family *Nitrospiraceae* is essential for nitrification, and thus its high abundance in agricultural soils is expected [63].

Less abundant genes were *wcaBF* and *wzt*, encoding for the colanic acid biosynthesis acetyltransferases WcaB and WcaF, and an ATP-binding protein Wzt of the LPS O-antigen transport system, respectively. Relatively high abundances of the genes from the colanic acid biosynthesis pathway are not surprising, as colanic acid is one of the most common exopolysaccharides. However, it is also one of the exopolysaccharides secreted by Wza. Therefore, *wcaBF* were less abundant than *wza* due to their lower universality [19]. In contrast, *wzt* is involved in translocating the O-antigen to the outer leaflet of the inner membrane where it gets ligated to the other parts of LPS [64]. Thus, the lower abundance of *wzt* compared to *lptFG* can be explained by the fact that the O-antigen is not an essential component of LPS [22].

Finally, a very low number of reads was assigned to the other investigated genes, which catalyze the biosynthesis and export of alginate, colanic acid, levan and other extracellular and capsular polysaccharides. EPS biosynthesis pathways are generally poorly conserved and often species or strain-specific, so this result was expected [19, 20, 65]. Also the low abundances of *algEJ*, encoding for alginate export outer membrane protein AlgE and alginate biosynthesis acetyltransferase AlgJ, were understandable because alginate is produced by various bacteria from the genera *Pseudomonas* and *Azotobacter* [66]. These genera belong to the family *Pseudomonadaceae*, which was not dominant in our metagenomes. *Pseudomonadaceae* contains many plant growth promoting endophytes and rhizobacteria [67]. Its low abundance could be related to the poorly established vegetation at the time of sampling (March) and the fact that bulk soil samples were investigated instead of rhizosphere samples.

EPS and LPS biosynthesis is important in agricultural soils, yet could be easily hindered

In our metagenomes, almost half (46.6%) of all bacterial families harbored genes affiliated with EPS and LPS biosynthesis. These included all dominant families whose abundance exceeded 0.5%. The ability to form EPS or LPS seems to be, therefore, an important trait for bacteria living in agricultural soils.

Despite the fact that all families dominating in our metagenomes harbored genes encoding for EPS or LPS biosynthesis, none of them accommodated genes from more than one of the investigated EPS biosynthesis pathways. This is not surprising, as few bacteria are known to produce more than one type of EPS [68]. Nonetheless, some of the analyzed genes were not represented in any of the dominant families. This could have several reasons, including: (i) that those genes might be harbored by low abundant families only, which were below the detection limit of our approach, or (ii) no genome of a representative taxon was sequenced so far. This is very likely, as databases for sequencing analysis are still biased towards fast-growing bacteria, while soil contains many slow-growing bacteria, which are difficult to isolate and culture.

Bacterial potential to produce EPS and LPS is affected by the interplay of tillage type and tillage depth

Tillage not only disrupts soil aggregates, but also alters soil physical and chemical properties. Especially CT disturbs microbial habitats and reduces available nutrients by mixing topsoil with subsoil. Since bacterial polysaccharides contribute to soil aggregation, which was reported to be lower under CT [10, 11], and EPS and LPS production requires a lot of carbon, we hypothesized that CT weakens the bacterial potential to produce soil structure-stabilizing agents. Contradictory to this hypothesis, but similar to the SAF results in our study, the tillage system influenced bacterial community composition and the potential to synthesize EPS and LPS only in the context of the depth factor. Direct effects of tillage were visible only when closely analyzing the taxonomic affiliation of the key genes of the investigated EPS and LPS biosynthesis pathways (*wza*, *lptFG*). At each soil depth, unique families harboring the respective genes under either CT or RT were described. These findings are in accordance with the theory about functional

redundancy, which states that different taxa are able to perform the same functions under changed conditions [69]. However, EPS and LPS produced by different bacteria may differ in quality, and can have different gluing properties [70, 71]. Therefore, the differences in aggregate preservation observed in other studies could be related to the differences in the properties of EPS and LPS produced under CT and RT.

The fact that the differences between tillage systems at our sampling sites were not more pronounced in comparison to other studies, is surprising, but not unprecedented. In fact, our results are in agreement with other functional analyses of agricultural soils. The work of de Vries et al. [58], who also compared CT and RT using metagenomics, and Grafe et al. [72], who compared different fertilization regimes, found little significant effects on bacterial community structure and functionality. Both studies implied that under long-term management, bacterial communities are very stable and hardly differ between treatments. In fact, it is more likely that regulation takes place on the RNA level, as tillage alters soil conditions, and thus might influence metabolic activity of soil structure-stabilizing bacteria. Ultimately, the yield of EPS and LPS could be increased or decreased by multiple factors, e.g. carbon sources or oxygen availability [18]. Thus, a metatranscriptomic analysis of the soil samples should be the next step. Ideally, omics data and SAF measurements should be correlated with the content of bacterial polysaccharides in soil. Redmile-Gordon et al. [73] made efforts to evaluate the suitability of different extracellular polymeric substances extraction methods for this medium. However, the existing methodologies are still biased and do not allow for distinction between polysaccharides of different origins (bacterial, fungal, plant, etc.). Therefore, further research needs to address these issues in order to establish a standardized protocol.

RT and CT promote the potential to produce EPS and LPS in bacteria

To our knowledge, previous metagenomic comparisons of tillage systems encompassed only surface soil samples [46–50]. However, other studies on tillage included analyses of chemical and physical properties of soil also at deeper levels [74–76]. The studies revealed that tillage can differentially impact bacterial habitats of different soil layers. Gadermaier et al. [4] demonstrated, that also at our sampling site the effects of tillage on soil organic carbon, microbial biomass and soil nutrients, varied with the depth of sampling. Therefore, we expected that bacterial communities at different depths would be differently affected by tillage, prompting the inclusion of a depth factor in our metagenomic analysis.

It is well-known that bacterial communities change with depth in undisturbed soils [77–79]. We showed that the composition of bacterial families in the upper soil depths (in the tillage horizon) differs from the one in lower soil depths (below the tillage horizon) also in tilled soils. This happens because specific conditions of different soil depths select for the best-adapted microorganisms. That is to say, deeper soil layers are generally more oxygen-depleted and nutrient-poor than upper soil layers. In the deepest soil layer which we sampled, the dominant family was *Anaerolineaceae*. Unsurprisingly, its members are strictly anaerobic oligotrophs [80, 81]. Including *Anaerolineaceae*, 34.8% of bacterial families detected in our metagenomes were significantly influenced by depth. Furthermore, we showed that in tilled soils, depth has not only a big influence on bacterial community composition, but also on relative abundances of genes involved in EPS and LPS synthesis and excretion. The relative abundances of *wza*, *wcaF* and *lptFG* were higher in the upper soil layers. Additionally, *epsA* was influenced by the interaction of tillage and depth, but its low abundance undermines the significance of this finding. Moreover, the diversity of bacterial families which harbored *wza* and *lptG*, two out of three most abundant genes of the analyzed biosynthesis pathways, decreased with depth. These effects should be even more pronounced due to the stratification of Cmic, which was significantly higher under both CT and RT, in the 0–20 cm layers as opposed to the 20–50 cm layers. Although we expected higher potential to produce EPS and LPS in the deeper, undisturbed soil layers, these observations suggest that EPS and LPS synthesis plays a bigger role in the surface soil layers, which are regularly disturbed by tillage. This could be explained by better aeration and availability of nutrients in the tillage horizon, as these parameters are known to be important for EPS and LPS production [18, 82, 83]. Otherwise, Galant et al. [84] postulated that disturbances increase the diversity and productivity of bacteria performing important ecological functions, which also coincides with our results. In our study, the disturbance caused by tillage could select for bacteria which are capable of synthesizing protective compounds, such as EPS and LPS.

Finally, it is difficult to separate depth and tillage effects, as the depth effects might be also induced by tillage. The stratification of soil chemical and physical properties in our study was artificially induced by tillage [4]. In particular, soil organic carbon (SOC) steadily decreased with depth under both CT and RT. By introducing such changes in soil properties along the soil profile, tillage indirectly caused the shifts in bacterial communities allocated as the effects of depth. Those shifts could be driven primarily by the disturbance caused by tillage. Specifically, tillage could stir the established bacterial communities in the tillage horizon, making it possible for new taxa to emerge. At the same time, a

long-term competition in the undisturbed soil layers below the tillage horizon would enable only the best-adapted bacteria to thrive. This type of competition-driven dominance of selected taxa is well-known in ecological communities [85, 86]. Moreover, it has recently been demonstrated that periodic disturbances have an impact on bacterial communities by promoting the cohabitation of ecologically different bacteria [84]. In conclusion, as similar effects of depth were detected under both CT and RT, the impact of tillage in general, might be more selective than the subtle differences between these two systems.

Conclusions

Although a typical stratification of soil carbon and microbial biomass was observed under RT in our study, no difference in the stable aggregate fraction of the soil or the potential to produce EPS and LPS was observed between RT and CT systems. While the potential to produce EPS and LPS was enhanced in the tillage horizon, tillage affected the taxonomic affiliation of genes encoding for proteins involved in the biosynthesis of specific EPS and LPS. These compounds can have different properties depending on the bacterial producers. Thus, the regulation of EPS and LPS formation can take place at two levels: (i) even small changes in the bacterial community composition could disturb the overall capacity of EPS and LPS to stabilize soil structure, or (ii) regulation takes place on the level of gene expression. Consequently, future studies need to figure out under which conditions the potential to produce EPS and LPS is recalled. However, the fast turnover of mRNA would require another sampling strategy which accounts for that dynamic, such as high resolved samplings throughout the season and the day, as beside tillage, also carbon input by plants and fertilization might influence expression of the respective genes. Moreover, soil at the sampling site was already well-structured due to its high clay content. We expect a stronger effect of tillage in sandy soils, which lack the fine particles necessary to form stable soil aggregates.

Additional file

Additional file 1: Agricultural practices applied in the experimental field in Frick prior to sampling. (DOC 28 kb)

Additional file 2: KO numbers related to EPS or LPS production found in the online Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database (October 2016). (DOC 26 kb)

Additional file 3: Details of the sequencing run. Shown are the numbers of obtained reads, total length of the reads and average read length per sample before and after quality filtering. "C" and "R" at the beginning of sample names stand for either "conventional tillage" or "reduced tillage", respectively. The following "A", "B" and "C" stand for the sampling depth (A – 0-10 cm, B – 10-20 cm and C – 20-50 cm). (DOC 43 kb)

Additional file 4: NMDS ordination plots depicting taxonomic profiles of bacteria at the family level in conventional and reduced tillage-treated soils sampled at three different depths. Ellipses drawn around triplicates represent a 95% confidence level. Shown is A) overall community, and B) affiliation of genes related to EPS and LPS synthesis. Each point in the plot represents a different sample ($n = 18$), and the location of the points is based on Bray-Curtis distances. Taxonomic assignment was performed against the National Center for Biotechnology Information Non-Redundant (NCBI-NR) protein sequences database. Functional genes were assigned using hidden Markov models (HMMs) obtained from the TIGRFAMs and Pfam databases, and then sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database. (PDF 152 kb)

Additional file 5: Bacterial families whose relative abundances and potential to produce EPS or LPS were significantly affected depth or interaction of tillage and depth. Significant differences between the treatments were determined by a multilevel model ($n = 3$, $p < 0.05$). (DOC 30 kb)

Additional file 6: Comparison of the 35 most abundant bacterial families according to taxonomic annotations based on the NCBI-NR and SILVA databases. (PDF 248 kb)

Additional file 7: Rarefaction curves of metagenomic datasets derived from conventional and reduced tillage-treated soils sampled at three different depths. Depicted is the number of assigned genes involved in EPS and LPS production as a function of sequencing depth. The genes were assigned using hidden Markov models (HMMs) obtained from the TIGRFAMs and Pfam databases, and then sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database. "C" and "R" at the beginning of sample names stand for either "conventional tillage" or "reduced tillage", respectively. The following "A", "B" and "C" stand for the sampling depth (A – 0-10 cm, B – 10-20 cm and C – 20-50 cm). (JPEG 1320 kb)

Additional file 8: Boxplot depicting Shannon-Weiner index values which describe the diversity of bacterial families harboring genes *wza*, *lptF* and *lptG* at three depths. Significant influence of depth, but not tillage, was detected when applying a multilevel model analysis (n = 3). Therefore, tillage treatments were pooled for this plot. The influence of depth is symbolized with "*". Significance levels are represented by the amount of symbols: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$. (PDF 117 kb)

Abbreviations

Cmic: Microbial biomass carbon; CT: Conventional tillage; DOC: Dissolved organic carbon; EPS: Exopolysaccharides; LPS: Lipopolysaccharides; RT: Reduced tillage; SAF: Stable aggregate fraction; SOC: Soil organic carbon

Acknowledgments

We thank Susanne Kublik and Anne Schöler for their help in performing the sequencing run. In addition, we would like to thank Lisa Guan for English proofreading.

Funding

The authors gratefully acknowledge the funding provided by the German Federal Office for Agriculture and Food (BLE) and the financial support for the Frick trial provided by the Swiss Federal Office for Agriculture (FOAG), the Software AG-Stiftung, the Stiftung zur Pflege von Mensch, Mitwelt und Erde and the Edith Maryon Stiftung. Moreover, this study was associated to the project Fertility Building Management Measures in Organic Cropping Systems (FertilCrop), which received funding from the CORE Organic Plus Funding Bodies, being partners of the FP7 ERA-Net project CORE Organic Plus.

Availability of data and materials

The raw sequencing data used in this study is available at the sequencing read archive (SRA) under the accession number PRJNA387672.

Authors' contributions

BC designed the experiment, carried out the laboratory work, analyzed the data and wrote the manuscript. GV contributed to the data analysis. MK and AF were responsible for the field work and and chemical analyses. AF, MS

and SS contributed to the design of the experiment. All authors edited the manuscript and approved the final draft.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

Author details
[1]Research Unit Comparative Microbiome Analysis, Helmholtz Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH), Ingolstädter Landstraße 1, 85764 Neuherberg, Germany. [2]Department of Soil Sciences, Research Institute of Organic Agriculture (FiBL), Ackerstrasse 113, 5070 Frick, Switzerland. [3]Soil Science, Technical University of Munich, Emil-Ramann-Straße 2, 85354 Freising, Germany. [4]Section for Bioinformatics, Department of Health Technology, Technical University of Denmark, 2800 Lyngby, Denmark.

Received: 5 March 2019 Accepted: 7 March 2019
Published online: 28 March 2019

References

1. FAO, ITPS. Status of the world's soil resources (SWSR) – main report. Rome: Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils; 2015.
2. Montgomery DR. Soil erosion and agricultural sustainability. P Natl Acad Sci USA. 2007;104:13268–72.
3. Berner A, Hildermann I, Fliessbach A, Pfiffner L, Niggli U, Mäder P. Crop yield and soil fertility response to reduced tillage under organic management. Soil Till Res. 2008;101:89–96.
4. Gadermaier F, Berner A, Fließbach A, Friedel JK, Mäder P. Impact of reduced tillage on soil organic carbon and nutrient budgets under organic farming. Renew Agr Food Syst. 2012;27:68–80.
5. Mäder P, Fliessbach A, Dubois D, Gunst L, Fried P, Niggli U. Soil fertility and biodiversity in organic farming. Science. 2002;296:1694–7.
6. Peigné J, Ball BC, Roger-Estrade J, David C. Is conservation tillage suitable for organic farming? A review. Soil Use Manage. 2007;23:129–44.
7. Cooper J, Baranski M, Stewart G, Nobel-de Lange M, Bàrberi P, Fließbach A, Peigné J, Berner A, Brock C, Casagrande M. Shallow non-inversion tillage in organic farming maintains crop yields and increases soil C stocks: a meta-analysis. Agron Sustain Dev. 2016;36:1–20.
8. Grandy AS, Robertson GP, Thelen KD. Do productivity and environmental trade-offs justify periodically cultivating no-till cropping systems? Agron J. 2006;98:1377–83.
9. Stockfisch N, Forstreuter T, Ehlers W. Ploughing effects on soil organic matter after twenty years of conservation tillage in Lower Saxony, Germany. Soil Tillage Res. 1999;52:91–101.
10. Bartlova J, Badalikova B, Pospíšilová L, Pokorný E, Šarapatka B. Water stability of soil aggregates in different systems of tillage. Soil Water Res. 2015;10:147–54.
11. Singh P, Heikkinen J, Mitra S, Regina K. Impact of tillage and residue management practices on soil aggregates and soil organic carbon. In: Raju NJ, editor. Geostatistical and geospatial approaches for the characterization of natural resources in the environment. Cham: Springer; 2016. p. 381–4.
12. Horn R, Taubner H, Wuttke M, Baumgartl T. Soil physical properties related to soil structure. Soil Till Res. 1994;30:187–216.
13. Lynch JM, Bragg E. Microorganisms and soil aggregate stability. In: Stewart BA, editor. Advances in soil science. New York: Springer; 1985. p. 133–71.
14. Snyder VA, A VM. Structure. In: Daniel H, editor. Encyclopedia of soils in the environment. Volume 4. New York: Academic Press; 2005. p. 54–68.
15. Totsche KU, Amelung W, Gerzabek MH, Guggenberger G, Klumpp E, Knief C, Lehndorff E, Mikutta R, Peth S, Prechtel A. Microaggregates in soils. J Plant Nutr Soil Sc. 2018;181:104–36.
16. Six J, Bossuyt H, Degryze S, Denef K. A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. Soil Till Res. 2004;79:7–31.
17. Abu-Lail NI, Camesano TA. Role of lipopolysaccharides in the adhesion, retention, and transport of Escherichia coli JM109. Environ Sci Technol. 2003;37:2173–83.
18. Suresh Kumar A, Mody K, Jha B. Bacterial exopolysaccharides–a perception. J Basic Microb. 2007;47:103–17.
19. Schmid J, Sieber V, Rehm B. Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. Front Microbiol. 2015;6:496.
20. Pereira SB, Mota R, Santos CL, De Philippis R, Tamagnini P. Assembly and export of extracellular polymeric substances (EPS) in cyanobacteria: a phylogenomic approach. Adv Bot Res. 2013;65:235–79.
21. Hunt F. Patterns of LPS synthesis in gram negative bacteria. J Theor Biol. 1985;115:213–9.
22. Wang X, Quinn PJ. Lipopolysaccharide: biosynthetic pathway and structure modification. Prog Lipid Res. 2010;49:97–107.
23. Krauss M, Ruser R, Müller T, Hansen S, Mäder P, Gattinger A. Impact of reduced tillage on greenhouse gas emissions and soil carbon stocks in an organic grass-clover ley-winter wheat cropping sequence. Agric Ecosyst Environ. 2017;239:324–33.
24. Quintero M, Comerford NB. Effects of conservation tillage on total and aggregated soil organic carbon in the Andes. Open J Soil Sci. 2013;3:361–73.
25. Murer EJ, Baumgarten A, Eder G, Gerzabek MH, Kandeler E, Rampazzo N. An improved sieving machine for estimation of soil aggregate stability (SAS). In: Brussaard L, Kooistra MJ, editors. Soil Structure/Soil Biota Interrelationships. Amsterdam: Elsevier; 1993. p. 539–47.
26. Lueders T, Manefield M, Friedrich MW. Enhanced sensitivity of DNA-and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. Environ Microbiol. 2004;6:73–8.
27. Mukherjee S, Huntemann M, Ivanova N, Kyrpides NC, Pati A. Large-scale contamination of microbial isolate genomes by Illumina PhiX control. Stand Genomic Sci. 2015;10:18.
28. Vestergaard G, Schulz S, Schöler A, Schloter M. Making big data smart—how to use metagenomics to understand soil quality. Biol Fert Soils. 2017;53:479–84.
29. Schubert M, Lindgreen S, Orlando L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res Notes. 2016;9:88.
30. Schmieder R, Edwards R. Fast identification and removal of sequence contamination from genomic and metagenomic datasets. PLoS One. 2011;6:e17288.
31. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12:59–60.
32. Huson DH, Mitra S, Ruscheweyh H-J, Weber N, Schuster SC. Integrative analysis of environmental sequences using MEGAN4. Genome Res. 2011;21:1552–60.
33. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 2012;28:3211–7.
34. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7:335–6.
35. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI. CDD: NCBI's conserved domain database. Nucleic Acids Res. 2015;43:D222–6.
36. Haft DH, Selengut JD, Richter RA, Harkins D, Basu MK, Beck E. TIGRFAMs and genome properties in 2013. Nucleic Acids Res. 2013;41:D387–95.
37. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 2016;44:D279–85.
38. Rho M, Tang H, Ye Y. FragGeneScan: predicting genes in short and error-prone reads. Nucleic Acids Res. 2010;38:e191.
39. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2016.
40. Pinheiro J, Bates D, DebRoy S, Sarkar D. nlme: Linear and nonlinear mixed effects models. 2018; R package version 3.1–131.1.
41. Aho K. asbio: a collection of statistical tools for biologists. 2015; R package version 1.1–5.
42. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Steven MHH, Wagner H. vegan: community ecology package. 2015; R package version 2.3–0.

43. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics. 2015;16:169.
44. Barthes B, Roose E. Aggregate stability as an indicator of soil susceptibility to runoff and erosion; validation at several levels. Catena. 2002;47:133–49.
45. Mathew RP, Feng Y, Githinji L, Ankumah R, Balkcom KS. Impact of no-tillage and conventional tillage systems on soil microbial communities. Appl Environ Soil Sci. 2012;2012:548620.
46. Carbonetto B, Rascovan N, Álvarez R, Mentaberry A, Vázquez MP. Structure, composition and metagenomic profile of soil microbiomes associated to agricultural land use and tillage systems in argentine pampas. PLoS One. 2014;9:e99949.
47. Miura T, Niswati A, Swibawa I, Haryani S, Gunito H, Arai M, Yamada K, Shimano S, Kaneko N, Fujie K. Shifts in the composition and potential functions of soil microbial communities responding to a no-tillage practice and bagasse mulching on a sugarcane plantation. Biol Fert Soils. 2016;52:307–22.
48. Navarro-Noya YE, Gómez-Acata S, Montoya-Ciriaco N, Rojas-Valdez A, Suárez-Arriaga MC, Valenzuela-Encinas C, Jiménez-Bueno N, Verhulst N, Govaerts B, Dendooven L. Relative impacts of tillage, residue management and crop-rotation on soil bacterial communities in a semi-arid agroecosystem. Soil Biol Biochem. 2013;65:86–95.
49. Souza RC, Cantão ME, Vasconcelos ATR, Nogueira MA, Hungria M. Soil metagenomics reveals differences under conventional and no-tillage with crop rotation or succession. Appl Soil Ecol. 2013;72:49–61.
50. Souza RC, Hungria M, Cantão ME, Vasconcelos ATR, Nogueira MA, Vicente VA. Metagenomic analysis reveals microbial functional redundancies and specificities in a soil under different tillage and crop-management regimes. Appl Soil Ecol. 2015;86:106–12.
51. Kandeler E, Tscherko D, Spiegel H. Long-term monitoring of microbial biomass, N mineralisation and enzyme activities of a Chernozem under different tillage management. Biol Fert Soils. 1999;28:343–51.
52. Staley T. Soil microbial biomass alterations during the maize silage growing season relative to tillage method. Soil Sci Soc of Am J. 1999;63:1845–7.
53. Balota EL, Colozzi-Filho A, Andrade DS, Dick RP. Microbial biomass in soils under different tillage and crop rotation systems. Biol Fert Soils. 2003;38:15–20.
54. Bronick CJ, Lal R. Soil structure and management: a review. Geoderma. 2005;124:3–22.
55. Kuntz M, Berner A, Gattinger A, Scholberg J, Mäder P, Pfiffner L. Influence of reduced tillage on earthworm and microbial communities under organic arable farming. Pedobiologia. 2013;56:251–60.
56. Pereira SB, Mota R, Vieira CP, Vieira J, Tamagnini P. Phylum-wide analysis of genes/proteins related to the last steps of assembly and export of extracellular polymeric substances (EPS) in cyanobacteria. Sci Rep-UK. 2015;5:14835.
57. Wooley JC, Godzik A, Friedberg I. A primer on metagenomics. PLoS Comput Biol. 2010;6:1–13.
58. De Vries M, Schöler A, Ertl J, Xu Z, Schloter M. Metagenomic analyses reveal no differences in genes involved in cellulose degradation under different tillage treatments. FEMS Microbiol Ecol. 2015;91:fiv069.
59. Putker F, Bos MP, Tommassen J. Transport of lipopolysaccharide to the gram-negative bacterial cell surface. FEMS Microbiol Rev. 2015;39:985–1002.
60. Ruiz N, Gronenberg LS, Kahne D, Silhavy TJ. Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of Escherichia coli. P Natl Acad Sci USA. 2008;105:5537–42.
61. Benedet M, Falchi FA, Puccio S, Di Benedetto C, Peano C, Polissi A, Dehò G. The lack of the essential LptC protein in the trans-envelope lipopolysaccharide transport machine is circumvented by suppressor mutations in LptF, an inner membrane component of the Escherichia coli transporter. PLoS One. 2016;11:e0161354.
62. Antunes LCS, Poppleton D, Klingl A, Criscuolo A, Dupuy B, Brochier-Armanet C, Beloin C, Gribaldo S. Phylogenomic analysis supports the ancestral presence of LPS-outer membranes in the Firmicutes. Elife. 2016;5:e14589.
63. Daims H. The family Nitrospiraceae. In: Rosenberg E, editor. The prokaryotes. Berlin, Heidelberg: Springer; 2014. p. 733–49.
64. Narita S. ABC transporters involved in the biogenesis of the outer membrane in gram-negative bacteria. Biosci Biotechnol Biochem. 2011;75:1044–54.
65. Skorupska A, Janczarek M, Marczak M, Mazur A, Król J. Rhizobial exopolysaccharides: genetic control and symbiotic functions. Microb Cell Factories. 2006;5:7.
66. Hay ID, Wang Y, Moradali MF, Rehman ZU, Rehm BHA. Genetics and regulation of bacterial alginate production. Environ Microbiol. 2014;16:2997–3011.
67. Preston GM. Plant perceptions of plant growth-promoting Pseudomonas. Philos T Roy Soc A. 2004;359:907–18.
68. Kwon K-J, Park K-J, Kim J-D, Kong J-Y, Kong I-S. Isolation of two different polysaccharides from halophilic Zoogloea sp. Biotechnol Lett. 1994;16:783–8.
69. Allison SD, Martiny JBH. Resistance, resilience, and redundancy in microbial communities. P Natl Acad Sci USA. 2008;105:11512–9.
70. Erridge C, Bennett-Guerrero E, Poxton IR. Structure and function of lipopolysaccharides. Microbes Infect. 2002;4:837–51.
71. Sutherland IW. Microbial polysaccharides from gram-negative bacteria. Int Dairy J. 2001;11:663–74.
72. Grafe M, Goers M, Sv T, Baum C, Zimmer D, Leinweber P, Vestergaard G, Kublik S, Schloter M, Schulz S. Bacterial potentials for uptake, solubilization and mineralization of extracellular phosphorus in agricultural soils are highly stable under different fertilization regimes. Envir Microbiol Rep. 2018;10:320–7.
73. Redmile-Gordon M, Brookes P, Evershed R, Goulding K, Hirsch P. Measuring the soil-microbial interface: extraction of extracellular polymeric substances (EPS) from soil biofilms. Soil Biol Biochem. 2014;72:163–71.
74. Cheng-Fang L, Dan-Na Z, Zhi-Kui K, Zhi-Sheng Z, Jin-Ping W, Ming-Li C, Cou-Gui C. Effects of tillage and nitrogen fertilizers on CH4 and CO2 emissions and soil organic carbon in paddy fields of Central China. PLoS One. 2012;7:e34642.
75. Jabro J, Stevens W, Iversen W, Evans R. Tillage depth effects on soil physical properties, sugarbeet yield, and sugarbeet quality. Commun Soil Sci Plan. 2010;41:908–16.
76. Angers DA, Bolinder MA, Carter MR, Gregorich EG, Drury CF, Liang BC, Voroney RP, Simard RR, Donald RG, Beyaert RP. Impact of tillage practices on organic carbon and nitrogen storage in cool, humid soils of eastern Canada. Soil Till Res. 1997;41:191–201.
77. Uksa M, Schloter M, Endesfelder D, Kublik S, Engel M, Kautz T, Köpke U, Fischer D. Prokaryotes in subsoil—evidence for a strong spatial separation of different phyla by analysing co-occurrence networks. Front Microbiol. 2015;6:1269.
78. Fierer N, Schimel JP, Holden PA. Variations in microbial community composition through two soil depth profiles. Soil Biol Biochem. 2003;35:167–76.
79. Griffiths RI, Whiteley AS, O'donnell AG, Bailey MJ. Influence of depth and sampling time on bacterial community structure in an upland grassland soil. FEMS Microbiol Ecol. 2003;43:35–43.
80. Hui N, Jumpponen A, Francini G, Kotze DJ, Liu X, Romantschuk M, Strömmer R, Setälä H. Soil microbial communities are shaped by vegetation type and park age in cities under cold climate. Environ Microbiol. 2017;19:1281–95.
81. Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H, Kamagata Y. Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. Nov., sp. nov. and Leptolinea tardivitalis gen. Nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the bacterial phylum Chloroflexi. Int J Syst Evol Micr. 2006;56:1331–40.
82. Exley RM, Shaw J, Mowe E, Sun Y-h, West NP, Williamson M, Botto M, Smith H, Tang CM. Available carbon source influences the resistance of Neisseria meningitidis against complement. J Exp Med. 2005;201:1637–45.
83. McDonald IJ, Adams GA. Influence of cultural conditions on the lipopolysaccharide composition of Neisseria sicca. Microbiology+. 1971;65:201–7.
84. Galand PE, Lucas S, Fagervold SK, Peru E, Pruski AM, Vétion G, Dupuy C, Guizien K. Disturbance increases microbial community diversity and production in marine sediments. Front Microbiol. 2016;7:1950.
85. Huston M, Lawton JH. Biological diversity. The coexistence of species on changing landscapes. Nature. 1995;373:32.
86. Rashit E, Bazin M. Environmental fluctuations, productivity, and species diversity: an experimental study. Microb Ecol. 1987;14:101–12.

Additional file 1: Agricultural practices applied in the experimental field in Frick prior to sampling.

Date	Procedure
20.10.2013	Winter wheat seeding
19.03.2014	Slurry application, 69 kg total N/ha
09.04.2014	Slurry application, 57 kg total N/ha
17.07.2014	Winter wheat harvest
25.08.2014	Green manure seeding
17./18.03.2015	Soil sampling

Additional file 2: KO numbers related to EPS or LPS production found in the online Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database (October 2016).

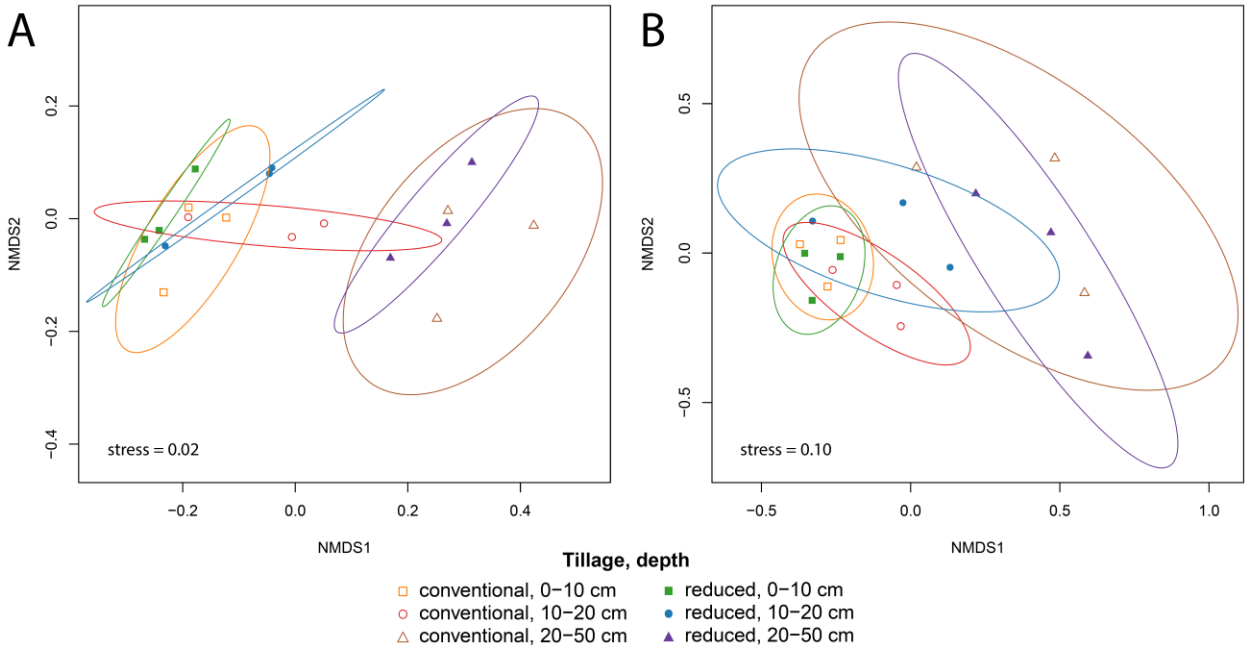
EPS (67)	K00689, K00692, K00694, K00752, K00903, K01991, K03207, K03208, K03606, K03818, K03819, K09688, K09689, K10107, K13620, K13650, K13654, K13683, K13684, K16081, K16552, K16553, K16554, K16555, K16556, K16557, K16558, K16560, K16561, K16562, K16563, K16564, K16565, K16566, K16567, K16568, K16692, K16696, K16700, K16701, K16702, K16703, K16708, K16709, K16710, K16711, K16712, K16713, K19292, K19293, K19294, K19295, K19296, K19418, K19419, K19420, K19421, K19422, K19424, K19425, K19426, K19427, K19428, K19429, K19430, K19431, K19667
LPS (14)	K05399, K05790, K06861, K07091, K07271, K08280, K08992, K09690, K09691, K09774, K11719, K11720, K16695, K19804

Additional file 3: Details of the sequencing run. Shown are the numbers of obtained reads, total length of the reads and average read length per sample before and after quality filtering. “C” and “R” at the beginning of sample names stand for either “conventional tillage” or “reduced tillage”, respectively. The following “A”, “B” and “C” stand for the sampling depth (A – 0-10 cm, B – 10-20 cm and C – 20-50 cm).

Raw data	CA1	CA2	CA3	CB1	CB2	CB3	CC1	CC2	CC3
Number of reads	2549636	2708128	2387144	1956536	2095558	2372948	1675552	1761742	2084406
Total length of reads	767440436	815146528	718530344	588917336	630762958	714257348	504341152	530284342	627406206
Average length of reads	301	301	301	301	301	301	301	301	301
Filtered data									
Number of reads	2549495	2707738	2387077	1956447	2095401	2372859	1675291	1761533	2083272
Total length of reads	757415945	805059148	709958547	581913254	623067219	705306628	497711661	522694430	611096918
Average length of reads	297.08	297.32	297.42	297.43	297.35	297.24	297.09	296.73	293.34

Raw data	RA1	RA2	RA3	RB1	RB2	RB3	RC1	RC2	RC3
Number of reads	2355348	2178762	2670446	2030052	2318146	2442574	1659310	2252744	1812660
Total length of reads	708959748	655807362	803804246	611045652	697761946	735214774	499452310	678075944	545610660
Average length of reads	301	301	301	301	301	301	301	301	301
Filtered data									
Number of reads	2355296	2178624	2670386	2030004	2318025	2442433	1659110	2252382	1812502
Total length of reads	700755406	648213723	792113860	603953567	689473640	726293001	493120371	668547107	538679590
Average length of reads	297.52	297.53	296.63	297.51	297.44	297.36	297.22	296.82	297.20

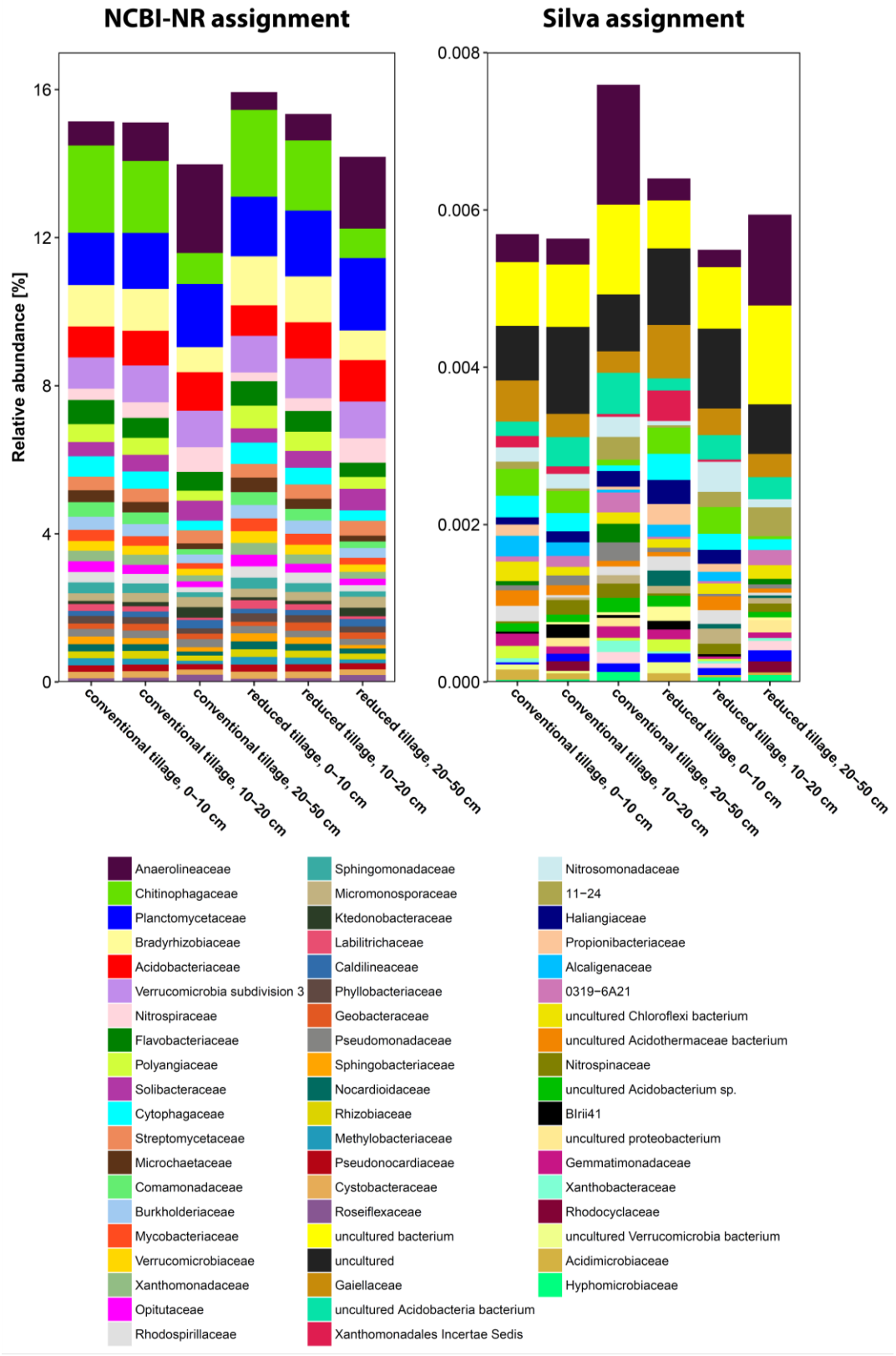
Additional file 4: NMDS ordination plots depicting taxonomic profiles of bacteria at the family level in conventional and reduced tillage-treated soils sampled at three different depths. Ellipses drawn around triplicates represent a 95% confidence level. Shown is A) overall community, and B) affiliation of genes related to EPS and LPS synthesis. Each point in the plot represents a different sample (n = 18), and the location of the points is based on Bray-Curtis distances. Taxonomic assignment was performed against the National Center for Biotechnology Information Non-Redundant (NCBI-NR) protein sequences database. Functional genes were assigned using hidden Markov models (HMMs) obtained from the TIGRFAMs and Pfam databases, and then sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database.



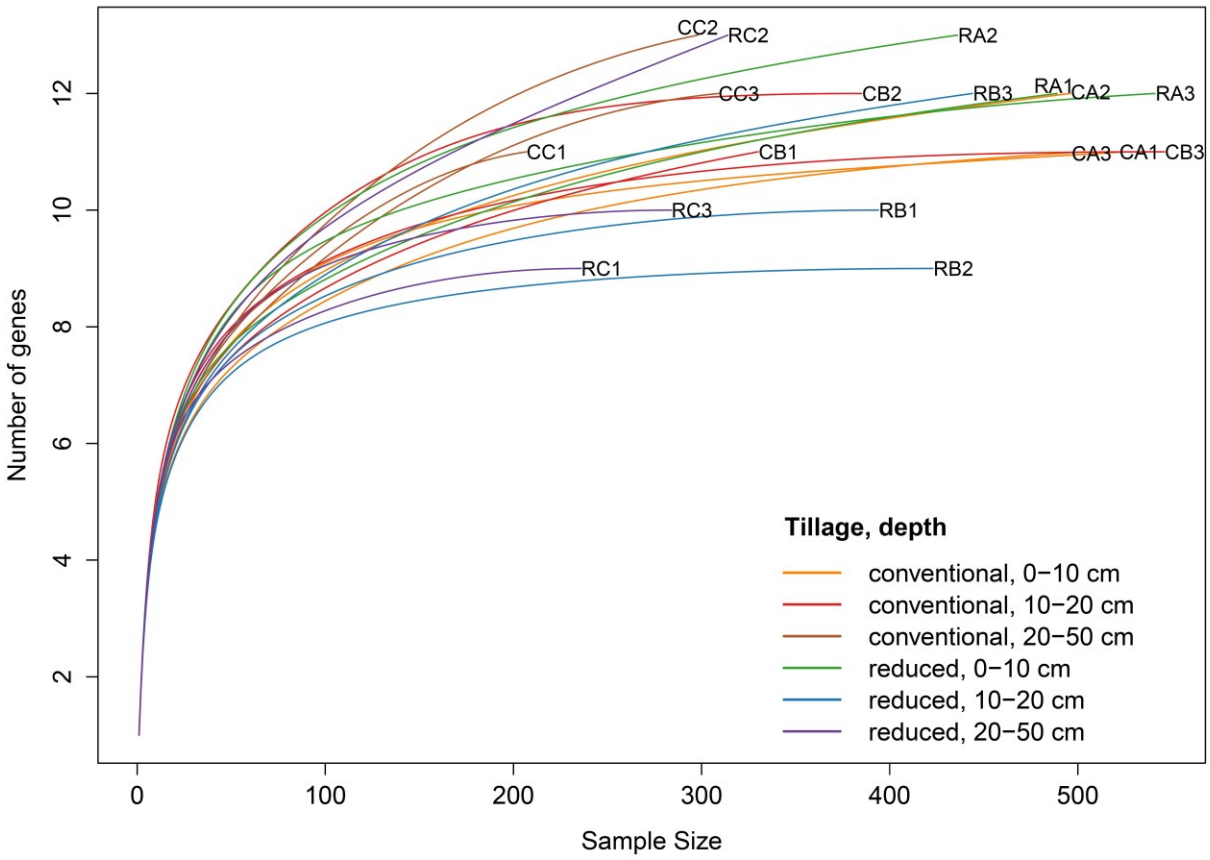
Additional file 5: Bacterial families whose relative abundances and potential to produce EPS or LPS were significantly affected depth or interaction of tillage and depth. Significant differences between the treatments were determined by a multilevel model ($n = 3$, $p < 0.05$).

Factor	Families whose abundance was affected	Families whose potential to produce EPS/LPS was affected
Depth	*Acetobacteraceae, Acidimicrobiaceae, Alcaligenaceae, Alicyclobacillaceae, Anaerolineaceae, Aquificaceae, Ardenticatenaceae, Bacillaceae, Bdellovibrionaceae, Bogoriellaceae, Bradyrhizobiaceae, Burkholderiaceae, Caldilineaceae, Caldisericaceae, Candidatus Actinomarinaceae, Candidatus.Brocadiaceae, Carnobacteriaceae, Caulobacteraceae, Cellulomonadaceae, Chitinophagaceae, Chlorobiaceae, Chloroflexaceae, Chromobacteriaceae, Comamonadaceae, Conexibacteraceae, Corynebacteriaceae, Cyclobacteriaceae, Cytophagaceae, Dehalococcoidaceae, Deinococcaceae, Desulfobacteraceae, Desulfohalobiaceae, Ectothiorhodospiraceae, Erythrobacteraceae, Flammeovirgaceae, Geodermatophilaceae, Gordoniaceae, Herpetosiphonaceae, Hydrogenophilaceae, Hyphomicrobiaceae, Hyphomonadaceae, Ignavibacteriaceae, Intrasporangiaceae, Kofleriaceae, Ktedonobacteraceae, Labilitrichaceae, Magnetococcaceae, Marinilabiliaceae, Melioribacteraceae, Methylobacteriaceae, Microbacteriaceae, Microchaetaceae, Micrococcaceae, Mycobacteriaceae, Nakamurellaceae, Nannocystaceae, Nitrosomonadaceae, Nitrospiraceae, Nocardioidaceae, Nostocaceae, Opitutaceae, Oscillochloridaceae, Oxalobacteraceae, Paenibacillaceae, Patulibacteraceae, Peptococcaceae, Phycisphaeraceae, Phyllobacteriaceae, Polyangiaceae, Porphyromonadaceae, Propionibacteriaceae, Rhodobacteraceae, Rhodobiaceae, Rhodospirillaceae, Rivulariaceae, Roseiflexaceae, Rubrobacteraceae, Sandaracinaceae, Saprospiraceae, Schleiferiaceae, Scytonemataceae, Sinobacteraceae, Solirubrobacteraceae, Sphingobacteriaceae, Sphingomonadaceae, Spirochaetaceae, Sulfuricellaceae, Sutterellaceae, Syntrophaceae, Syntrophobacteraceae, Syntrophomonadaceae, Syntrophorhabdaceae, Thermaceae, Thermoanaerobacteraceae, Thermogemmatisporaceae, Thiotrichaceae, Trueperaceae, Verrucomicrobiaceae, Vibrionaceae, Vulgatibacteraceae, Waddliaceae, Xanthobacteraceae, Xanthomonadaceae*	*Chitinophagaceae, Flammeovirgaceae, Labilitrichaceae, Nitrospiraceae*
Interaction of tillage and depth	*Jonesiaceae*	*Sulfuricellaceae*

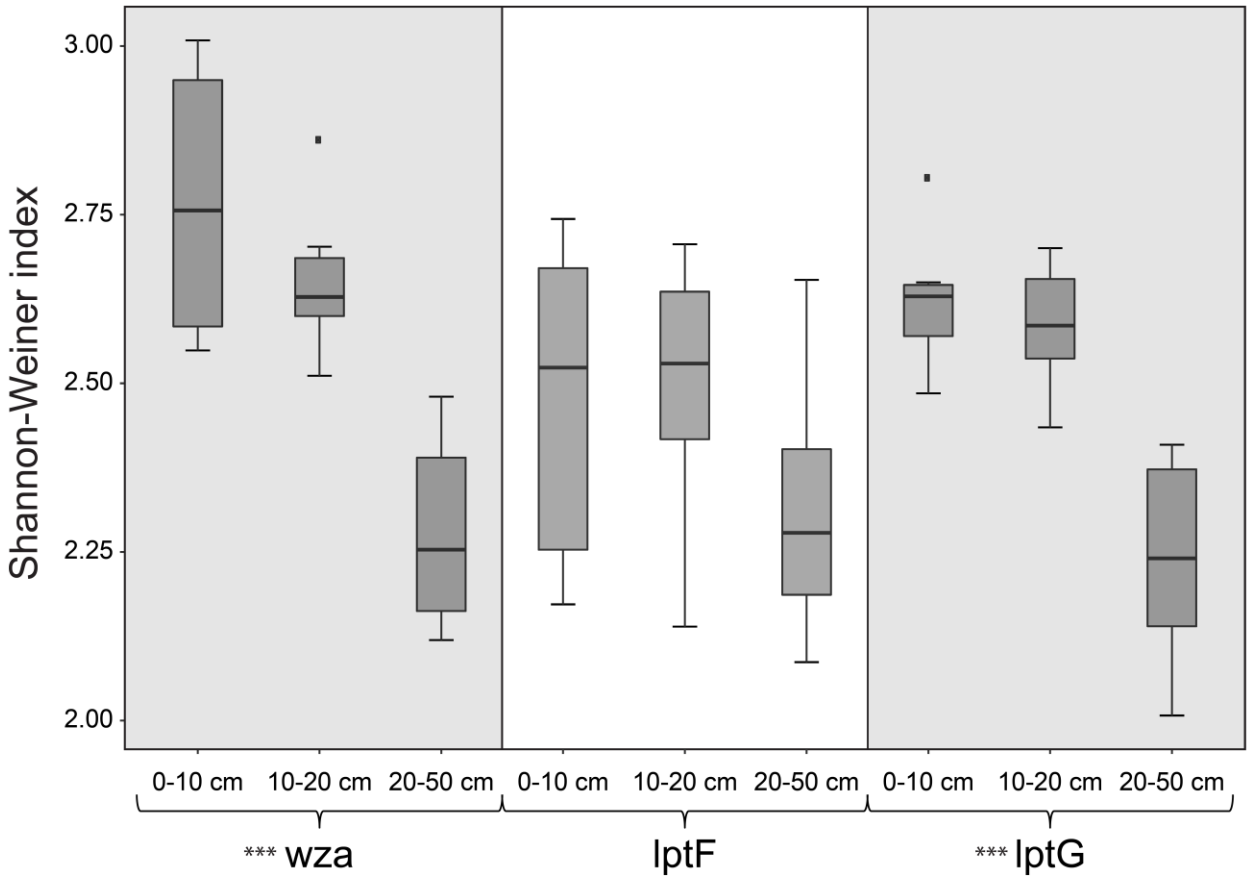
Additional file 6: Comparison of the 35 most abundant bacterial families according to taxonomic annotations based on the NCBI-NR and SILVA databases.



Additional file 7: Rarefaction curves of metagenomic datasets derived from conventional and reduced tillage-treated soils sampled at three different depths. Depicted is the number of assigned genes involved in EPS and LPS production as a function of sequencing depth. The genes were assigned using hidden Markov models (HMMs) obtained from the TIGRFAMs and Pfam databases, and then sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database. “C” and “R” at the beginning of sample names stand for either “conventional tillage” or “reduced tillage”, respectively. The following “A”, “B” and “C” stand for the sampling depth (A – 0-10 cm, B – 10-20 cm and C – 20-50 cm).



Additional file 8: Boxplot depicting Shannon-Weiner index values which describe the diversity of bacterial families harboring genes wza, lptF and lptG at three depths. Significant influence of depth, but not tillage, was detected when applying a multilevel model analysis (n = 3). Therefore, tillage treatments were pooled for this plot. The influence of depth is symbolized with "*". Significance levels are represented by the amount of symbols: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$.



C Publication 3

frontiers in Microbiology

ORIGINAL RESEARCH
published: 07 April 2020
doi: 10.3389/fmicb.2020.00568

Site-Specific Conditions Change the Response of Bacterial Producers of Soil Structure-Stabilizing Agents Such as Exopolysaccharides and Lipopolysaccharides to Tillage Intensity

*Barbara Cania[1], Gisle Vestergaard[1,2], Marjetka Suhadolc[3], Rok Mihelič[3], Maike Krauss[4], Andreas Fliessbach[4], Paul Mäder[4], Anna Szumełda[5], Michael Schloter[1,6] and Stefanie Schulz[1]**

[1] Research Unit Comparative Microbiome Analysis, Helmholtz Zentrum München, Research Center for Environmental Health (GmbH), Neuherberg, Germany, [2] Department of Health Technology, Section for Bioinformatics, Technical University of Denmark, Lyngby, Denmark, [3] Chair of Soil and Environmental Science, Department of Agronomy, University of Ljubljana, Ljubljana, Slovenia, [4] Department of Soil Sciences, Research Institute of Organic Agriculture (FiBL), Frick, Switzerland, [5] Stanisław Karłowski Foundation, Silnowo, Poland, [6] Chair of Soil Science, Technical University of Munich, Freising, Germany

OPEN ACCESS

Edited by:
Manuel Delgado Baquerizo, University of Alicante, Spain

Reviewed by:
Youzhi Feng, Institute of Soil Science (CAS), China
Katharina Kujala, University of Oulu, Finland

***Correspondence:**
Stefanie Schulz
stefanie.schulz@helmholtz-muenchen.de

Specialty section:
This article was submitted to Terrestrial Microbiology, a section of the journal Frontiers in Microbiology

Received: 26 November 2019
Accepted: 16 March 2020
Published: 07 April 2020

Citation:
Cania B, Vestergaard G, Suhadolc M, Mihelič R, Krauss M, Fliessbach A, Mäder P, Szumełda A, Schloter M and Schulz S (2020) Site-Specific Conditions Change the Response of Bacterial Producers of Soil Structure-Stabilizing Agents Such as Exopolysaccharides and Lipopolysaccharides to Tillage Intensity. Front. Microbiol. 11:568. doi: 10.3389/fmicb.2020.00568

Agro-ecosystems experience huge losses of land every year due to soil erosion induced by poor agricultural practices such as intensive tillage. Erosion can be minimized by the presence of stable soil aggregates, the formation of which can be promoted by bacteria. Some of these microorganisms have the ability to produce exopolysaccharides and lipopolysaccharides that "glue" soil particles together. However, little is known about the influence of tillage intensity on the bacterial potential to produce these polysaccharides, even though more stable soil aggregates are usually observed under less intense tillage. As the effects of tillage intensity on soil aggregate stability may vary between sites, we hypothesized that the response of polysaccharide-producing bacteria to tillage intensity is also determined by site-specific conditions. To investigate this, we performed a high-throughput shotgun sequencing of DNA extracted from conventionally and reduced tilled soils from three tillage system field trials characterized by different soil parameters. While we confirmed that the impact of tillage intensity on soil aggregates is site-specific, we could connect improved aggregate stability with increased absolute abundance of genes involved in the production of exopolysaccharides and lipopolysaccharides. The potential to produce polysaccharides was generally promoted under reduced tillage due to the increased microbial biomass. We also found that the response of most potential producers of polysaccharides to tillage was site-specific, e.g., *Oxalobacteraceae* had higher potential to produce polysaccharides under reduced tillage at one site, and showed the opposite response at another site. However, the response of some potential producers of polysaccharides to tillage did not depend on site characteristics, but rather on their taxonomic affiliation, i.e., all members of *Actinobacteria* that responded

to tillage intensity had higher potential for exopolysaccharide and lipopolysaccharide production specifically under reduced tillage. This could be especially crucial for aggregate stability, as polysaccharides produced by different taxa have different "gluing" efficiency. Overall, our data indicate that tillage intensity could affect aggregate stability by both influencing the absolute abundance of genes involved in the production of exopolysaccharides and lipopolysaccharides, as well as by inducing shifts in the community of potential polysaccharide producers. The effects of tillage intensity depend mostly on site-specific conditions.

Keywords: tillage, soil aggregation, exopolysaccharides, lipopolysaccharides, soil microbiome, metagenomics, *wza*

INTRODUCTION

Worldwide, 75 billion tons of soil are lost every year by erosion of arable lands (ELD Initiative, 2015). Soil erosion mostly occurs due to the degradation of soil structure, which is defined as the arrangement of soil particles into aggregates (Le Bissonnais et al., 1993). A continuous disruption of soil aggregates by agricultural practices such as conventional tillage (CT) may lead to increased soil compaction as well as loss of organic matter and soil biodiversity. As a result, water retention is impaired, anoxic conditions may appear and nutrient cycling slows down. Such soils are more susceptible to erosion induced by water or wind (Holland, 2004). To protect the aggregated soil structure and prevent soil loss, soil conservation techniques such as reduced tillage (RT) are increasingly encouraged (FAO, 2017).

Although many researchers (Jacobs et al., 2009; Mikha et al., 2013; Bartlova et al., 2015; Sheehy et al., 2015; Singh et al., 2016) observed that soil aggregates are better preserved under RT compared with CT, others (Asgari, 2014; Cania et al., 2019a) reported no effect of tillage intensity on soil aggregation. According to Cooper et al. (2016), the strength of the positive effects of RT strongly depends on soil texture. Clay particles have much stronger aggregating properties compared with silt and sand (Weil and Brady, 2017). Therefore, the positive effects of RT should be more emphasized in soils with lower clay content, where maintaining high aggregation is more challenging (Cooper et al., 2016). This is in agreement with studies using clayey soils that revealed no differences in aggregate stability when RT and CT were compared (Asgari, 2014; Cania et al., 2019a). However, soil aggregation is a complex process that is driven by both abiotic and biotic factors (Six et al., 2004), and still little is known about how tillage intensity influences soil biota and their capabilities to trigger aggregate formation. Babin et al. (2019) proposed that the effects of tillage intensity on soil biota and aggregate formation driven by them can be better studied in soils with lower clay content due to their lower buffering capacity. In accordance with this hypothesis, studies on sandy and silty soils showed that improved soil aggregation under less intense tillage corresponded to increased fungal biomass and glomalin production (Beare et al., 1997; Wright et al., 1999; Cookson et al., 2008; Dai et al., 2015; Lu et al., 2018). However, while many researchers investigated the influence of tillage intensity on the aggregating capabilities of fungi, less attention was given to bacteria. In fact, only in a recent study (Cania et al., 2019a), we compared the effects of CT and RT on the bacterial potential for soil aggregation. Here, we investigated the bacterial potential for the production of exopolysaccharides (EPSs) and lipopolysaccharidesof (LPSs), which act as binding agents for soil particles (Six et al., 2004; Costa et al., 2018; Totsche et al., 2018). EPSs are high-molecular-weight sugars exuded by a wide range of taxa (Suresh Kumar et al., 2007), whereas LPSs are complex glycolipids attached to the outer membrane of most Gram-negative bacterial cells (Whitfield and Trent, 2014). Bacteria use these compounds for attachment to soil particles, which promotes the formation of soil aggregates (Jacques, 1996; Sutherland, 2001). According to Lehmann et al. (2017), the bacterial production of adhesive polysaccharides is one of the most crucial biotic mechanisms of soil aggregation. We (Cania et al., 2019a) could show that while the relative abundance of bacteria capable to form EPSs and LPSs was comparable between CT and RT, the community composition of the potential producers of these compounds differed. As the aggregating efficiency of adhesive polysaccharides produced by different taxa varies greatly (Costa et al., 2018), tillage impact on the community composition of EPS and LPS producers could be critical for the stability of agricultural soils.

However, disentangling the link between soil aggregate stability, tillage and soil texture requires the analysis of long-term experiments, as tillage effects build up over time (Stockfisch et al., 1999; Grandy et al., 2006). Therefore, it was our aim to investigate to which extend the response of polysaccharide-producing bacteria to tillage intensity is driven by the differences in soil texture at different long-term experimental sites, and how this is connected to the stability of soil aggregates. As we were interested in the long-term tillage impact, we focused on parameters that change over a long time period such as aggregate stability and the structure and genetic potential of soil bacterial communities, as opposed to transient parameters such as gene expression and the contents of EPSs and LPSs in soil (Redmile-Gordon et al., 2020). We expected to see more differences in the relative abundance and community composition of potential EPS and LPS producers between CT and

Abbreviations: Cmic, microbial biomass carbon; Corg, organic carbon; CT, conventional tillage; DOC, dissolved organic carbon; DON, dissolved organic nitrogen; EPS, exopolysaccharide; LPS, lipopolysaccharide; Nmic, microbial biomass nitrogen; RT, reduced tillage; SAF, stable aggregate fraction.

RT in silty and sandy soils compared with clayey soils. We also assumed that the differences in the potential to produce bacterial polysaccharides would be reflected by changes in aggregate stability. To address our research questions, we performed a high-throughput shotgun sequencing of DNA extracted from conventionally and reduced tilled soils from three long-term field trials characterized by different soil textures (clayey, loamy, sandy). We used a targeted bioinformatics pipeline to analyze bacterial communities potentially involved in the production of EPSs and LPSs.

MATERIALS AND METHODS

Sites Description and Sampling

Soil samples were taken from three field trials that differed in soil properties but were comparable in terms of soil management, particularly tillage application (CT and RT). The basic information on the trials, such as the locations, soil types and climatic conditions, is summarized in **Table 1**.

At the Frick trial (Switzerland), CT has been based on ploughing with a moldboard plough operating at 15–18 cm depth, while for RT, soil loosening has been performed at a depth of 5–10 cm with a chisel and a skim plough, with occasional non-inversion loosening to 15–20 cm. In both systems, the seedbed preparation has been done using a rotary harrow running at 5 cm depth. The last tillage operations before sampling occurred in September 2016 right after harvesting spelt and before grass-clover was sown. In 2017, five grass-clover harvests were done in April, June, July, September and November. The plots were fertilized with slurry in 2016 in August, and in 2017 in April, June and August at the rates of 45, 40, 25 and 30 m^3 ha^{-1}, respectively, before sowing and after the first, second and third grass-clover harvest.

At the Moškanjci trial (Slovenia), for CT, a moldboard plough operating at 20 cm depth has been used, followed by soil bed preparation with a rotary hoe. For RT, a special machine – 4-row disc harrow with individually suspended discs and a system for varying the working angle (and thus the tilling intensity) – has been applied in one or two passes to till the soil and prepare the seedbed. The depth of RT was up to 10 cm. In both systems, the main soil tillage was done in October 2016, just before winter rye was sown. After the winter rye was harvested in July 2017, a mixture of cover crops was sown for green manure. The fertilization with slurry was applied in March 2017 in the amount of 20 m^3 ha^{-1}.

At the Juchowo trial (Poland), CT has been performed by ploughing up to 30 cm deep with an Ecomat plough, while for RT, soil loosening up to 10 cm deep has been done using a cultivator with goosefeet sweeps. In August 2015, the last tilling occurred after harvesting spelt and before lupine was sown. Slurry was applied in March 2015 at the rate of 16 m^3 ha^{-1}. Biodynamic preparations, consisting of subtle amounts of silica (horn silica 501) or fresh cow manure (horn manure 500) dissolved in a large volume of water, were sprayed at the rates of 12 L ha^{-1}, 35 L ha^{-1}, and 12 L ha^{-1} in March, April and May, respectively. Kieserit (MgO 25%, S 20%) was applied in May at the rate of 50 kg ha^{-1}.

Each trial consisted of four replicated plots per treatment, out of which three were sampled in spring, before any tilling and subsequent sowing started, in 2016 in Juchowo, and in 2018 in Frick and Moškanjci. Approximately 10 cores per plot were taken using soil augers with a diameter of 2.5 cm to a soil depth of 10 cm. Cores from the same plot were homogenized, resulting in 18 samples (3 trials x 2 tillage treatments x 3 plot replicates). After being directly cooled in the field, one part of each homogenized soil sample was stored at 4°C and used for physicochemical measurements, and the other was stored at −20°C for DNA extraction and sequencing.

TABLE 1 | Trial characteristics.

Trial	Frick	Moškanjci	Juchowo
Trial start	2002	1999	2010
Geographic coordinates	47°30′N, 8°01′E	46°03′N, 15°04′E	53°40′N, 16°30′E
Elevation [m a.s.l.]	350	225	160
Soil type	Stagnic Eutric Cambisol	Skeletic Eutric Cambisol	Haplic Arenosol
Soil texture	clayey	loamy	sandy
Climate type	temperate	continental	continental
Mean annual temperature [°C]	8.9	10.6	8.5
Mean annual precipitation [mm]	1000	913	750

Physicochemical Measurements

Stable aggregate fraction (SAF) of the soils was determined using a wet sieving method described by Murer et al. (1993). Field-moist soil samples, 4 g of each, were placed onto 0.25 mm sieves of a sieving apparatus, and immersed in water. After 5 min of wet sieving, the aggregates remaining on the sieves were dried at 105°C and weighed. The aggregates were then destroyed by covering them with 0.1 M sodium pyrophosphate ($Na_4P_2O_7$) solution for 2 h. The remaining particles >0.25 mm (sand and organic debris) were dried and weighed again. The SAF was calculated as percentage of stable aggregates in a moist sample, after applying a correction for sand particles. Aside from using moist soil without previous fractionation, the technique followed the details according to Murer et al. (1993). Soil texture was determined using a combined sieving and sedimentation approach (ISO 11277 2009). The determination of soil organic carbon (Corg) was based on the Walkley-Black wet digestion procedure (Walkley and Black, 1934). The measurement of dissolved organic carbon (DOC) and nitrogen (DON), as well as Cmic and nitrogen (Nmic) was accomplished by means of the chloroform fumigation-extraction (CFE) method (Brookes et al., 1985; Vance et al., 1987). DOC and DON were measured in unfumigated samples, while Cmic and Nmic were calculated as a difference in extractable carbon and nitrogen between fumigated and unfumigated soils. To calculate the microbial biomass, the difference was then multiplied with the empirical factors kEC and kEN to achieve Cmic (kEC = 0.45) and Nmic (kEN = 0.54) (Joergensen, 1996; Joergensen and Mueller, 1996).

The measurement of pH was performed in a 1:2.5 (v/v) suspension of soil in demineralized water (pH in H_2O) after standing overnight (ISO 10390 2005).

DNA Extraction, Library Preparation and Sequencing

Metagenomic DNA was extracted from 0.5 g of frozen soil according to the phenol-chloroform based DNA/RNA co-extraction protocol described by Lueders et al. (2004). Sample lysis was performed using CKMix tubes and a Precellys24 homogenizer (Bertin Technologies, France). Extracted DNA was quantified by means of a Qubit 4 Fluorometer and a Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, United States). The purity was also assessed by measuring the A260 nm/A280 nm and A260 nm/A230 nm ratios using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, United States). After extraction, DNA was stored at −20°C until further processing.

1 μg of DNA per sample was sheared using an E220 Focused-ultrasonicator (Covaris, United States), targeting 500 bp fragments as described in the protocol of the producer. Metagenomic libraries were constructed with 100 ng of the sheared DNA by means of a NEBNext Ultra II DNA Library Prep Kit for Illumina and NEBNext Multiplex Oligos for Illumina (New England Biolabs, United Kingdom). Following the manufacturer's guideline, the provided adaptor was diluted 10-fold to prevent the occurrence of dimers. Size selection was carried out with Agencourt AMPure XP beads (Beckman Coulter, United States), selecting for libraries with 500–700 bp inserts. The beads were also used for purification of PCR reactions and an additional final purification step to eliminate residual primer dimers (1:0.6 DNA to bead ratio). PCR amplification was performed with 12 cycles.

Library size and concentration were evaluated using a Fragment Analyzer and a DNF-473 Standard Sensitivity NGS Fragment Analysis Kit (Advanced Analytical, United States). Libraries were pooled equimolarily (4 nM), and 17 pM of the mixture was spiked with 1% PhiX. Paired-end sequencing was performed on a MiSeq sequencer using a MiSeq Reagent Kit v3 for 600 cycles (Illumina, United States). As a negative control, a reagent-only sample was processed alongside the biological samples at the DNA extraction and library preparation steps, and included in the sequencing run. Raw sequencing data obtained from the MiSeq is available at the Sequence Read Archive (SRA) under the accession number PRJNA555481.

Bioinformatical Analysis

The raw sequencing data was filtered according to Vestergaard et al. (2017). Remnant adaptor sequences were removed by means of AdapterRemoval v2.1.7 (Schubert et al., 2016). Using the same program, terminal nucleotides with Phred quality scores lower than 15 were trimmed, and reads shorter than 50 bp were discarded. PhiX contamination was removed by means of DeconSeq v0.4.3 (Schmieder and Edwards, 2011).

Filtered reads were taxonomically classified by aligning against the National Center for Biotechnology Information Non-Redundant (NCBI-NR) protein sequences database (January 2017) using Kaiju v1.4.4 (Menzel et al., 2016) in Greedy mode with five allowed mismatches. 16S rRNA gene sequences were additionally identified using SortMeRNA v2.0 (Kopylova et al., 2012) with the SILVA SSU database (release 132).

Subsequent assignment of genes specific for EPS and LPS biosynthesis was performed only for reads classified by Kaiju as bacterial. Hidden Markov model (HMM) searching combined with blasting against protein sequences obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (October 2016) was carried out according to Cania et al. (2019a). HMMs were downloaded from the TIGRFAMs (version 15) (Haft et al., 2013) and Pfam (version 30) (Finn et al., 2016) databases. Open-reading frames were predicted using FragGeneScan v1.19 (Rho et al., 2010), and then scanned using HMMER v3.1b2 (Mistry et al., 2013). Hits with a maximum *E*-value of 1×10^{-5} were mapped to KEGG Orthology (KO) numbers. KO numbers were assigned to the reads for which the best 25 blast results were matching. Blasting was performed by means of Diamond v0.8.38 (Buchfink et al., 2015) with more-sensitive parameters. Genes *epsA* and *epsG* had only a few reads annotated to them (7 and 4, respectively), mostly in single replicates. Therefore, they were excluded from the analysis, as a higher sequencing depth would be required to study them. The list of HMMs and KO numbers used in this study is presented in **Table 2**.

Statistical Analysis and Data Visualization

Statistical analysis and data visualization were performed using R v3.4.4 (R Core Team, 2016). Sequencing data was analyzed based on relative abundance of reads obtained by dividing the number of reads assigned to a bacterial family or gene, by the total number of bacterial reads per sample. The resultant decimals were subsequently multiplied by 100 in order to convert them into percentages. Absolute gene abundance was additionally estimated according to Zhang et al. (2017) by multiplying the relative gene abundance in percentage with the Cmic value, and dividing by 100. Although Cmic includes both bacterial and fungal carbon, the low and stable values of the Cmic/Nmic ratio indicate that the estimated values are rather precise (Cheng et al., 2013). A robust 2-way independent analysis of variance (ANOVA) was used to detect global effects of trial, tillage intensity and their possible interaction in the whole dataset. Local effects of tillage intensity were detected separately in the datasets from each trial by means of a robust *t*-test. Both statistical tests were based on the median as M-estimator, and used 2000 bootstrap samples. The tests were ran by employing the pbad2way and pb2gen functions from the WRS package (Wilcox and Schönbrodt, 2014). The influence was regarded as significant if the *p*-value was smaller than 5 % ($P < 0.05$). The false discovery rate was controlled in the data derived from the metagenomic datasets with the Benjamini-Hochberg procedure. Effect sizes were calculated in the form of omega squared (ω^2) and Pearson's correlation coefficient r as described by Field et al. (2012). The dissimilarity

TABLE 2 | Proteins related to exo- and lipopolysaccharide production with corresponding KO numbers, HMM IDs and genes.

Protein	KO number	HMM ID	Gene
polysaccharide export outer membrane protein Wza	K01991	PF02563	*wza*
colanic acid biosynthesis acetyltransferase WcaB	K03819	TIGR04016	*wcaB*
colanic acid biosynthesis acetyltransferase WcaF	K03818	TIGR04008	*wcaF*
colanic acid/amylovoran biosynthesis pyruvyl transferase WcaK/AmsJ	K16710	TIGR04006	*wcaK/amsJ*
capsular polysaccharide export system permease KpsE	K10107	TIGR01010	*kpsE*
alginate export outer membrane protein AlgE	K16081	PF13372	*algE*
alginate biosynthesis acetyltransferase AlgJ	K19295	PF16822	*algJ*
levansucrase SacB	K00692	PF02435	*sacB*
lipopolysaccharide transport system ATP-binding protein Wzt	K09691	PF14524	*wzt*
LptBFGC lipopolysaccharide export complex inner membrane protein LptC	K11719	TIGR04409, PF06835	*lptC*
LptBFGC lipopolysaccharide export complex permease LptF	K07091	TIGR04407	*lptF*
LptBFGC lipopolysaccharide export complex permease LptG	K11720	TIGR04408, PF03739	*lptG*

between the samples was explored using principal component analysis (PCA) ordination created with the rda function in the vegan package (Oksanen et al., 2018). For the same purpose, Bray-Curtis dissimilarity matrices were calculated by means of the vegdist function, also from the vegan package. Negative eigenvalues were corrected using the Caillez procedure (Cailliez, 1983). The Bray-Curtis distances were used to create principal coordinate analysis (PCoA) ordinations by means of the pcoa function from the ape package (Paradis et al., 2004). The PCA ordination was generated for the soil properties data, whereas the PCoA ordinations were generated for the sequencing data. For both types of ordination plots, ellipses representing 95 % confidence level were drawn around triplicates with the ellipse package (Murdoch and Chow, 2018).

RESULTS

Soil Properties

Different soil texture in Frick (46.5% clay, 24.5% silt, and 29.0% sand), Moškanjci (17.6% clay, 43.4% silt, and 39.0% sand) and Juchowo (2.4% clay, 6.4% silt, and 91.2% sand) was the primary reason why these trials were included in the current study. However, soil samples of CT and RT from the trials differed also in several other measured physical, chemical and biological parameters (**Table 3**). The parameters measured as part of the current study included soil texture, SAF, Corg, DOC, DON, pH, Cmic, Nmic, and the Cmic/Nmic ratio. PCA (**Figure 1A**) of these parameters revealed that the samples were separated primarily based on their place of origin. However, in Moškanjci, soil properties differed also between tillage treatments. Statistical analysis confirmed these observations. Aside from the soil texture data, robust ANOVA detected significant differences in Corg, DOC, DON, Cmic and Nmic values between the trials. Corg was highest in Frick (2.66%), intermediate in Moškanjci (1.59%) and lowest in Juchowo (0.81%). DOC was highest in Frick (68.8 mg/kg), intermediate in Juchowo (42.2 mg/kg) and lowest in Moškanjci (22.5 mg/kg). DON was higher in Frick and Juchowo (12.0 and 10.8 mg/kg, respectively) compared with Moškanjci (3.1 mg/kg). Cmic and Nmic were highest in Frick (1341.9 and 191.3 mg/kg, respectively), intermediate in Moškanjci (342.2 and 53.0 mg/kg, respectively) and lowest in Juchowo (121.5 and 18.5 mg/kg, respectively). Moreover, robust ANOVA identified significant interaction effects between trial and tillage intensity on SAF and pH. Both parameters were highest in Frick (94.3% and 7.2, respectively), intermediate in Moškanjci (69.3% and 6.7, respectively) and lowest in Juchowo (5.5% and 6.3, respectively). Interestingly, SAF and pH were similar under CT and RT in Frick and Juchowo, but higher under RT compared with CT in Moškanjci. Furthermore, robust ANOVA detected significant differences between tillage intensities in Corg, Cmic, and Nmic, with higher values of all of them under RT compared with CT in all three trials. However, the Cmic/Nmic ratios were not significantly influenced by the type of tillage, and did not vary between the sites. The described data is summarized in **Table 3**.

Baseline Data of the Shotgun Sequencing

Shotgun sequencing of 18 libraries resulted in 14.91 Gbases of data in total (49.54 million reads with a length of 301 bp each). Following quality control, the metagenomic datasets consisted of 14.73 Gbases (49.53 million reads). The number of filtered reads per sample ranged from 1.6 to 3.7 million, with a mean of 2.8 million. The average length of reads after filtering varied between samples from 292 to 298 bp, with a mean of 297 bp. More details of the raw and filtered sequencing data can be found in **Supplementary Table S1**. No reads were obtained for the negative control, confirming a lack of contamination during DNA extraction and library preparation.

Taxonomic Profiling

When all filtered reads were aligned against the NCBI-NR database, 76.8% were assigned to *Bacteria*, 2.9% to *Eukaryota*, 0.6% to *Archaea*, 0.1% to *Viruses* and 19.6% were unclassified. The results of the taxonomic profiling using the NCBI-NR database were supported by SILVA's assignment of the 16S rRNA gene. Even though only 0.07% of all filtered reads were annotated to the 16S rRNA gene, bacterial communities showed similar distribution patterns regardless of the analytical approach (data

TABLE 3 | Soil texture (clay, silt and sand content), stable aggregate fraction (SAF), organic carbon (Corg), dissolved organic carbon (DOC) and nitrogen (DON), pH, microbial biomass carbon (Cmic) and nitrogen (Nmic) as well as the Cmic/Nmic ratio data of the soils under conventional (CT) and reduced (RT) tillage. Influence of trial, tillage and their interaction was determined by a robust 2-way ANOVA.

Soil properties																					
Trial	**Tillage**	**Clay[%]**		**Silt[%]**		**Sand[%]**		**SAF[%]**		**Corg[%]**		**DOC[mg/kg]**		**DON[mg/kg]**		**pH**		**Cmic[mg/kg]**		**Nmic[mg/kg]**		**Cmic/Nmic**
Frick	CT	45.95 ± 6.01		23.01 ± 3.36		31.03 ± 3.39		94.32 ± 3.38		2.31 ± 0.34		69.32 ± 18.00		10.47 ± 2.36		7.33 ± 0.21		1147.20 ± 5.37		164.75 ± 4.37		6.97 ± 0.22
	RT	47.02 ± 4.14		26.03 ± 0.98		26.95 ± 3.80		90.83 ± 0.75		3.00 ± 0.24		68.23 ± 16.78		13.50 ± 2.65		7.08 ± 0.26		1536.49 ± 233.29		217.93 ± 37.58		7.07 ± 0.33
Moškanjci	CT	19.17 ± 1.65		43.23 ± 0.91		37.60 ± 2.55		69.32 ± 3.02		1.37 ± 0.11		20.46 ± 0.28		2.75 ± 1.30		6.43 ± 0.04		248.40 ± 27.17		35.54 ± 0.82		6.99 ± 0.77
	RT	15.93 ± 0.21		43.60 ± 1.51		40.47 ± 1.69		84.27 ± 2.10		1.81 ± 0.09		24.50 ± 1.76		3.48 ± 1.37		6.95 ± 0.09		435.95 ± 81.56		70.53 ± 11.79		6.17 ± 0.36
Juchowo	CT	3.04 ± 1.09		7.04 ± 2.45		89.92 ± 1.46		5.49 ± 0.59		0.70 ± 0.22		40.52 ± 2.59		11.00 ± 0.20		6.30 ± 0.36		118.88 ± 11.88		16.76 ± 5.34		7.54 ± 2.11
	RT	1.79 ± 0.40		5.83 ± 1.84		92.38 ± 2.06		5.36 ± 1.39		0.91 ± 0.24		43.86 ± 5.42		10.67 ± 3.81		6.20 ± 0.17		124.06 ± 66.58		20.18 ± 10.78		6.09 ± 0.51
Robust ANOVA (omega squared and significance levels)																					
Trial		0.97	***	0.98	***	0.99	***	0.98	***	0.85	***	0.79	***	0.75	***	0.68	***	0.92	***	0.91	***	0.00	ns
Tillage		0.00	ns	0.00	ns	0.00	ns	0.00	ns	0.07	***	0.00	ns	0.00	ns	0.00	ns	0.03	*	0.04	**	0.08	ns
Trial x Tillage		0.00	ns	0.00	ns	0.00	ns	0.01	***	0.01	ns	0.00	ns	0.00	ns	0.10	**	0.02	ns	0.01	ns	0.00	ns

Effect sizes (ω^2) and significance levels were calculated based on triplicates (n = 3). Significance levels are represented by the amount of stars: 1–$p < 0.05$, 2–$p < 0.01$, 3–$p < 0.001$.

not shown). For this reason, as well as because the focus of this study was on EPS and LPS producers of bacterial origin, only bacterial reads identified by means of the NCBI-NR database were analyzed further.

64.2% of the bacterial reads could be assigned at the family level. In total, 385 bacterial families were detected in the datasets, out of which 379 were present under both tillage treatments in all three trials. The remaining six families were very low abundant and altogether accounted for 0.0002% of all bacterial reads. In comparison, the most dominant families – *Bradyrhizobiaceae*, *Streptomycetaceae*, and *Sphingomonadaceae* – comprised 4.6%, 3.1%, and 2.2% of the bacterial reads, respectively. In total, fourteen families each accounted for more than 1% of all bacterial reads, and thus were dominant in the metagenomic datasets (**Supplementary Figure S1**). These families were all members of the phyla that ranked in the top ten most abundant in our study, led by *Proteobacteria* and *Actinobacteria*, which represented 37.5% and 20.8% of the bacterial reads, respectively (**Supplementary Figure S2**). Even though almost all of the detected families were present in all samples, their relative abundance differed between the trials or tillage types. As revealed by robust ANOVA, *Bradyrhizobiaceae* were more abundant in Juchowo (6.4%) compared with Moškanjci (3.7%) and Frick (3.4%), whereas *Sphingomonadaceae* were most abundant in Moškanjci (2.6%), intermediate in Juchowo (2.3%) and least in Frick (1.9%). Aside from *Bradyrhizobiaceae* and *Sphingomonadaceae*, another ten dominant families were significantly affected by trial, and three of them were additionally influenced by tillage treatment. In total, 237 families were influenced only by trial, 1 – only by the type of tillage, 4 – by both factors, and 8 – by the interaction of both factors. Interestingly, all the families that were significantly affected by tillage intensity (*Mycobacteriaceae*, *Nocardioidaceae*, *Micromonosporaceae*, *Glycomycetaceae* and *Dermatophilaceae*) were members of *Actinobacteria*. All of them were more abundant under RT compared with CT. The full list of impacted families can be found in **Supplementary Table S2**; the significance levels and ω^2 values for the dominant families are listed in **Supplementary Table S3**. PCoA (**Figure 1B**) confirmed that the relative abundances of bacterial families were affected primarily by trial, whereas the influence of tillage intensity played only a minor role in shaping the bacterial communities, and was visible mainly in Moškanjci. The results of the PCoA highly resembled those of the PCA of soil properties (**Figure 1A**), indicating that local conditions strongly affect the composition of bacterial communities and/or vice versa.

Genes Catalyzing EPS and LPS Biosynthesis

Genes specific for the synthesis and export of colonic acid, alginate and other EPSs as well as LPSs (**Table 2**), were identified using an approach combining hidden Markov model (HMM) searches with blasts against sequences derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. In total, the investigated genes comprised 0.033% of all bacterial reads (**Figure 2**). A rarefaction analysis (**Supplementary Figure S3**)

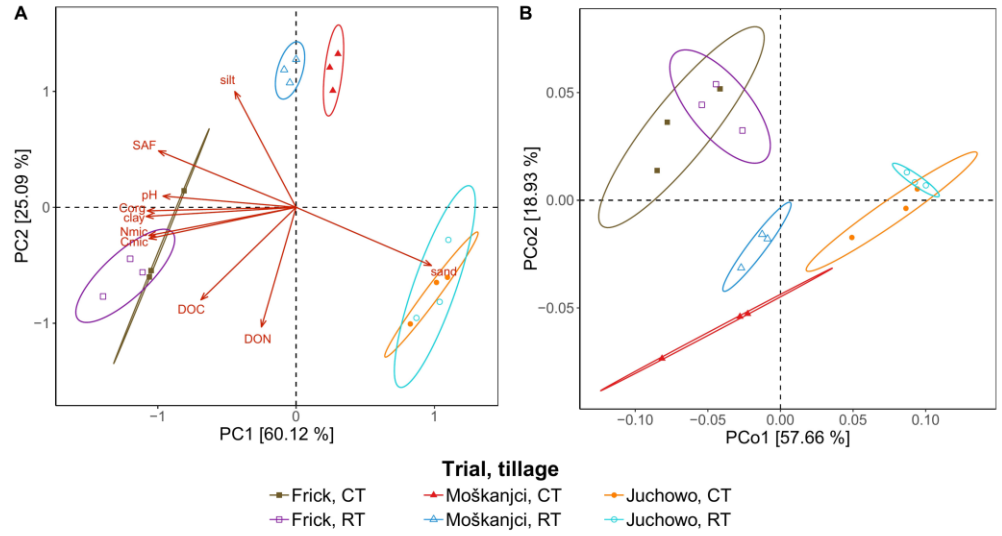


FIGURE 1 | PCA of soil parameters **(A)**, and PCoA based on Bray-Curtis distances depicting taxonomic profiles of bacteria at the family level **(B)**. Ellipses drawn around triplicates represent a 95% confidence level.

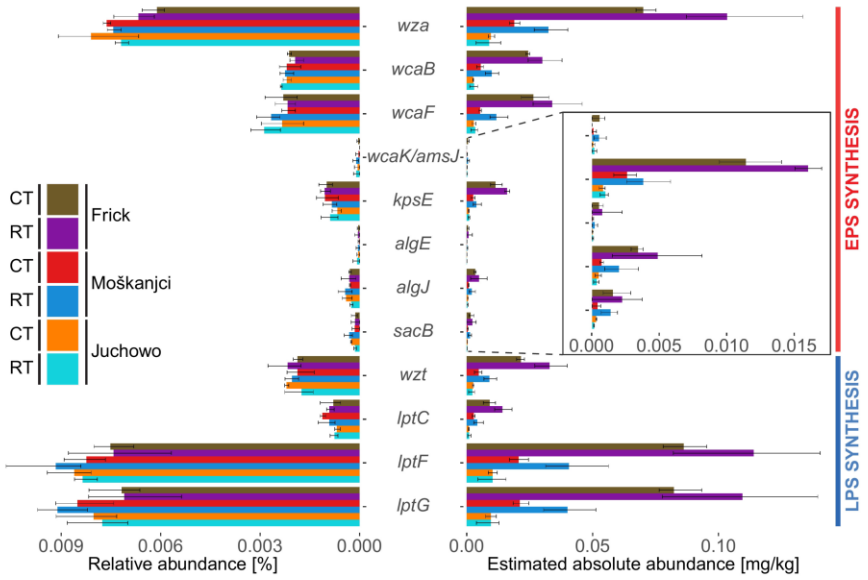


FIGURE 2 | Comparison of the relative and absolute (estimated based on Cmic values) abundances of genes related to the formation of EPSs and LPSs. Error bars show standard deviations.

confirmed that the sequencing depth was sufficient to capture the total diversity of these genes in the analyzed samples. Dominant genes, with a relative abundance between 0.007% and 0.008% of all bacterial reads, were *wza*, *lptF*, and *lptG*. Genes *wcaB*, *wcaF*, *kpsE*, *wzt*, and *lptC* were moderately abundant (0.001% – 0.002%). The least abundant (>0.001%) were genes *wcaK*/*amsJ*, *algE*, *algJ*, and *sacB*. Robust ANOVA revealed that only the relative abundance of *wza* differed significantly between samples. This gene was more abundant in Moškanjci and Juchowo than in Frick. The relative abundance of the other investigated genes was not significantly affected by either trial or tillage intensity. The significance levels

and ω^2 values for the investigated genes can be taken from **Supplementary Table S4**.

Compared with the relative abundance, the absolute abundance of genes related to the formation of EPSs and LPSs, estimated based on the Cmic values, showed higher variability (**Figure 2**). Robust ANOVA (**Supplementary Table S5**) indicated a significant impact of trial on most of the investigated genes. Specifically, *wza*, *wcaB*, *wcaF*, *kpsE*, *algJ*, *sacB*, *wzt*, *lptC*, *lptF* and *lptG* were most abundant in Frick, and least in Juchowo. Only *wcaK*/*amsJ* and *algE* were not significantly affected by trial. The estimated absolute abundance of none of the analyzed genes was significantly influenced by the type of tillage. However, robust ANOVA mostly describes differences in the complete dataset, and thus small but significant differences might be missed by this approach. Therefore, to identify site-specific impacts of the different forms of tillage on the estimated absolute abundance of the investigated genes, metagenomic datasets from the different trials were analyzed separately. Robust *t*-test (**Supplementary Table S5**) revealed three genes that responded to tillage intensity in Frick, eight in Moškanjci and one in Juchowo. In Frick, *wza*, *kpsE* and *wzt* were more abundant under RT compared with CT. Similarly, *wza*, *wcaB*, *wcaF*, *algJ*, *sacB*, *wzt*, *lptF* and *lptG* had higher abundances under RT compared with CT in Moškanjci. In Juchowo, *sacB* was more abundant under CT compared with RT.

Potential EPS/LPS Producers

The investigated genes were found in 260 bacterial families, including all dominant families (**Supplementary Figure S1**). Robust ANOVA detected that the potential of the bacterial community to produce EPSs and LPSs was most affected by trial alone and the interaction of tillage and trial rather than by tillage alone. The significance levels and ω^2 values for the dominant families are listed in **Supplementary Table S6**, and the full list of impacted families can be found in **Supplementary Table S2**. The relative abundance of reads related to EPS and LPS synthesis and export differed significantly between trials for five of the dominant families, and for further six among the less abundant members of the community of potential polysaccharide producers. Another six of the less abundant families were affected by the interaction of trial and tillage intensity, but none was influenced by tillage treatment alone.

While no general trends in the response to tillage were identified by robust ANOVA, robust *t*-test detected that the number of reads assigned to the genes of interest differed significantly between tillage types for 11 families in Frick, for 12 families in Moškanjci, and for 9 families in Juchowo. The statistical data for the families that responded to tillage intensity can be taken from **Supplementary Table S7**. Comparison of those families (**Figure 3**) revealed that their response to tillage type was trial-dependent. No family responded to tillage treatment in all three trials, but *Acetobacteraceae* showed higher potential for EPS and LPS formation under RT compared with CT in both Frick (0.00018% under RT vs. 0.00011% under CT) and Juchowo (0.00082% under RT vs. 0.00060% under CT). Surprisingly, some bacterial families responded differently when trials were compared. *Flavobacteriaceae* had higher potential for EPS and LPS biosynthesis under CT in Frick (0.00111% under CT vs. 0.00083% under RT), and showed the opposite behavior in Juchowo (0.00079% under RT vs. 0.00039% under CT). *Cytophagaceae* also had higher potential for EPS and LPS formation under CT in Frick (0.00043% under CT vs. 0.00007% under RT), but contrasting results were observed in Moškanjci (0.00032% under RT vs. 0.00026% under CT). Similarly, *Oxalobacteracea* showed higher potential for EPS and LPS biosynthesis in Moškanjci under CT (0.00090% under CT vs. 0.00029% under RT), but in Juchowo this bacterial family had higher potential for adhesive polysaccharide production under RT (0.00096% under RT vs. 0.00042% under CT). Out of the families affected by tillage intensity at each trial, more families (63.6% in Frick, 75.0% in Moškanjci, and 66.7% in Juchowo) showed higher potential for EPS and LPS formation under RT compared with CT. Amongst the families that responded to the type of tillage, members of *Actinobacteria* (*Geodermatophilaceae* and *Cellulomonadaceae* in Frick, *Mycobacteriaceae* and *Frankiaceae* in Moskanjci, as well as *Nocardioidaceae* in Juchowo) were found, and showed higher potential for EPS and LPS formation under RT compared with CT. Interestingly, families belonging to *Actinobacteria* were also the major responding groups to the type of tillage when total bacterial diversity was assessed (see above).

DISCUSSION

Link Between Soil Aggregate Stability, Tillage Intensity and Soil Texture

Preservation of soil aggregates is vital for soil sustainability in agro-ecosystems (Amezketa, 1999; Bronick and Lal, 2005). However, it is still not completely understood how soil aggregation is affected by agricultural practices such as different forms of tillage. Therefore, we investigated the effects of CT and RT on the bacterial potential to produce soil-aggregating agents such as EPSs and LPSs in agricultural trials with different soil properties. We showed that bacteria could be important drivers of aggregate formation and stabilization, but the effects of tillage intensity on soil aggregates strongly depend on site-specific properties.

The investigated sites differed in their aggregate stability, soil texture and organic carbon content. The most stable aggregates and the highest clay and organic carbon content were found in Frick, whereas least stable aggregates and the lowest clay and organic carbon content characterized the Juchowo site. Clay particles have the highest propensity compared with sand and silt particles to attract each other, form bonds and, in consequence, soil aggregates (Bronick and Lal, 2005; Weil and Brady, 2017; Totsche et al., 2018). Similarly, organic carbon is one of the most important binding agents responsible for the cementation of soil particles and aggregate formation and stabilization (Bronick and Lal, 2005). Therefore, the differences in clay and organic carbon content could largely explain the variability in the stable aggregate fraction between the three sites. However, we could also demonstrate a connection between the stable aggregate fraction and the genes involved in the biosynthesis of EPSs and LPSs.

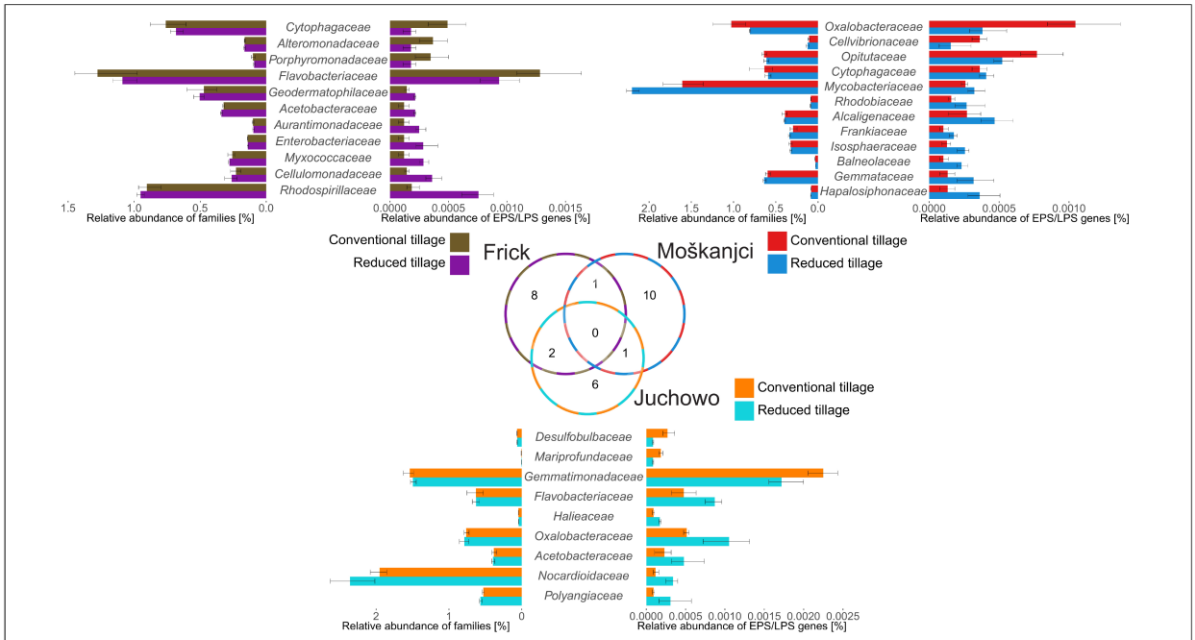


FIGURE 3 | Families whose relative abundance of genes involved in the production of EPSs and LPSs was affected by tillage intensity. The Venn diagram shows the number of families that responded to tillage at each trial. The tornado plots compare the relative abundances of the responsive families with the relative abundances of their sequences related to EPS and LPS formation. The responsive families are sorted according to the log2 fold change in their relative abundances of EPS and LPS production-related sequences between conventional and reduced tillage. Error bars show standard deviations.

Using Cmic values as a reference, we estimated that most of the investigated genes related to the formation of these adhesive polysaccharides (83.3%) were the most abundant in Frick, and the lowest abundance values were found in Juchowo. Furthermore, the estimated absolute abundances could explain the variability in the aggregate stability between tillage intensities as well, which was impossible using clay and organic carbon content alone. Similarly to other studies (Jacobs et al., 2009; Mikha et al., 2013; Bartlova et al., 2015; Sheehy et al., 2015; Singh et al., 2016), we observed more stable soil aggregates under less intense tillage in Moškanjci. However, in Frick and Juchowo, there was no significant difference in the stable aggregate fraction under CT and RT. While soil texture did not differ significantly between tillage intensities at any of the sites, and organic carbon content was higher under less intensive tillage at all three sites, the highest number of investigated genes (66.7%) affected by tillage intensity was found in Moškanjci (compared with 25.0% in Frick and 8.3% in Juchowo). Moreover, all of the genes affected by tillage treatment in Moškanjci had higher estimated absolute abundances under RT compared with CT. These results confirm the relationship between the bacterial genes involved in EPS and LPS formation and the stable aggregate fraction.

However, the question remains why the bacterial potential to form adhesive polysaccharides was affected most in Moškanjci compared with Frick and Juchowo. This could have been caused by the longest duration of the trial in Moškanjci (continuous for 19 years from 1999 until the time of sampling) compared with Frick (16 years) and Juchowo (6 years), as it was shown before that the effects of tillage on soil properties build up over time (Stockfisch et al., 1999; Grandy et al., 2006). The fact that tillage system had a significant influence on soil pH in Moškanjci, but not the other sites, supports this theory. Alternatively, the impact of tillage system on the bacterial potential to produce adhesive polysaccharides could have been regulated by site characteristics, such as soil texture. Smaller soil particles not only can be easier aggregated together by physicochemical forces, but also by organic agents such as EPSs and LPSs, which act like a glue (Bronick and Lal, 2005; Weil and Brady, 2017; Totsche et al., 2018). Therefore, the high clay content in Frick could help stabilize the soil and also protect bacteria living inside soil aggregates from being disturbed by both CT and RT, whereas the dominance of sand in Juchowo could hinder soil aggregation and make aggregate communities equally susceptible to being disturbed by CT and RT. Overall, the difference between the level of disturbance introduced by CT and RT could be emphasized only in soils with a balanced content of small and big particles, such as in Moškanjci. This theory stresses the importance of further research on the influence of soil texture and other parameters on the response of bacteria to tillage, as this area has been so far understudied.

We based our study on the metagenomic data on EPS and LPS production as well as aggregate stability measurements as the most appropriate parameters demonstrating the long-term influence of tillage on the aggregating capabilities of

bacteria. Observing the short-term effects of tillage would require measuring more dynamic parameters such as the content of bacterial polysaccharides in soil. However, while the existing methodologies of determining the polysaccharide content in this medium (Redmile-Gordon et al., 2014; Redmile-Gordon et al., 2020) do not allow for the distinction of polysaccharides from different sources, bacteria are not the only soil biota capable of producing these compounds. In fact, other organisms such as fungi or algae are also able to release polysaccharides with similar gluing properties, although these compounds are considered to be of lesser importance for soil aggregation compared with bacterial polysaccharides (Lehmann et al., 2017). Still, it would certainly be advantageous to analyze all organisms with the potential to produce adhesive polysaccharides using a single approach, especially since it would solve the issue of not being able to distinguish between different origins of polysaccharides found in soil. Unfortunately, the sequence databases are still biased toward bacteria, and therefore lacking gene sequences specific for the biosynthesis of extracellular polysaccharides. Moreover, the analysis of eukaryotes by means of short-read sequencing is difficult due to the high number of intronic sequences that these organisms possess (De Vries et al., 2015). Further work needs to be aimed at overcoming these obstacles if the full understanding of the complexities of soil aggregation is to be reached.

Importance of EPS/LPS Genes in Agricultural Soils

While absolute abundances describe the potential significance of taxa or genes for the ecosystem functioning, the proportional importance of these groups within communities can be better explained by relative abundances. The relative abundances of the EPS/LPS genes in our study were similar to the relative abundances of functional genes analyzed in other studies, responsible e.g., for cellulose degradation (De Vries et al., 2015) or phosphorus turnover (Grafe et al., 2018), which is in line with the fact that they are not housekeeping genes. However, EPSs and LPSs increase the ecological fitness of bacteria, and therefore it could be expected that the highest relative abundance of these genes would be found amongst the most dominant taxa. EPSs and LPSs provide protection from environmental stresses and predation, mediate surface attachment, function in cell-cell and cell-plant interactions, and act as carbon reserves (Kierek and Watnick, 2003; Suresh Kumar et al., 2007; Lindhout et al., 2009). Moreover, these compounds enable bacteria to create favorable hydrological niches in soil by improving soil structure (Benard et al., 2019). In our study, the highest relative abundance of the genes related to the production of adhesive polysaccharides were harbored by *Bradyrhizobiaceae* and *Sphingomonadaceae*. Both families are common in soils, and often live in a close association with plant roots (de Souza et al., 2014; Glaeser and Kämpfer, 2014). They are also well-known producers of adhesive polysaccharides (de Souza et al., 2014; Glaeser and Kämpfer, 2014). Therefore, their high potential to form these compounds was not surprising. Furthermore, these families were amongst the most dominant families in our metagenomic datasets, and the genes involved in EPS and LPS formation were found in all of the other dominant families as well. However, the investigated genes were found not only in the most abundant families, but, in fact, in most (67.5%) of the detected families. In comparison, the genes related to the production of adhesive polysaccharides were found in 9.5 - 64.7% of bacterial families from initial ecosystems and biological soil crusts (Cania et al., 2019b). This suggests that the potential to form EPSs and LPSs is indeed an important trait for bacteria living in agricultural soils, possibly because aside from natural events, these soils are regularly disturbed by anthropogenic influences such as tillage practices.

We found very little impact of tillage intensity on both the bacterial community composition and the genes related to EPS and LPS production. Only 1.3% of all bacterial families showed a consistent response to tillage at all three sites. These responders were all families of filamentous bacteria belonging to the phylum *Actinobacteria* (Rosenberg et al., 2014), and had higher relative abundance under RT compared with CT. This suggests that bacterial filaments could be disturbed by intensive tillage in a similar manner as fungal hyphae (Beare et al., 1997; Wright et al., 1999; Borie et al., 2006; Wang et al., 2010; Kihara et al., 2012; Lu et al., 2018). These results are in line with other metagenomic analyses of agricultural soils (De Vries et al., 2015; Grafe et al., 2018) implying that bacterial communities are overall very stable, in both composition and functionality, under long-term management, and show little differences between agricultural treatments.

The genes related to the production of EPSs and LPSs maintained stable relative abundances not only under different tillage intensities, but also between the sites. Amongst the investigated genes, the only one that showed significant changes in our datasets was *wza*. This gene had higher relative abundance in Moškanjci and Juchowo compared to Frick. Previous studies showed that *wza* is an important component of the EPS synthesis pathway (Cania et al., 2019a,b). This gene encodes for an outer membrane channel transporting a variety of EPSs through the outer membrane of many different bacterial taxa (Pereira et al., 2013). Its lower relative abundance in Frick could indicate that the conditions there are more favorable for bacteria compared to Moškanjci and Juchowo, and do not require as high potential for the formation of EPSs. For example, fine-textured clayey soils offer bacteria more protection from protozoan predation than more coarse-textured silty and sandy soils (Rutherford and Juma, 1992). Similarly, soils with higher clay content tend to be better at retaining water (Amooh and Bonsu, 2015). Therefore, the protection from engulfment by protozoa and from desiccation provided by polysaccharides could possibly hold less importance in soils with high clay content, such as in Frick, compared with silty and sandy soils, such as in Moškanjci and Juchowo, respectively.

While site had little impact on the relative abundances of the investigated genes, it had an important role in shaping the composition of bacterial communities. In fact, the relative abundances of most (64.7%) detected families differed between the sites. Such strong influence of site as a factor shaping bacterial communities was expected, as soil bacteria can be affected by soil and site characteristics such as pH, climate, nutrient availability,

and plant species (Fierer, 2017). These parameters varied between the investigated sites, and their impact could be seen in our metagenomic datasets, including the potential EPS and LPS producers such as the aforementioned *Bradyrhizobiaceae* and *Sphingomonadaceae*. *Bradyrhizobiaceae*, which are known for forming symbiotic associations with lupine (Reeve et al., 2013; de Souza et al., 2014), had the highest relative abundance in Juchowo, where this plant species was cultivated at the time of sampling, whereas *Sphingomonadaceae*, a bacterial family with more oligotrophic traits (Glaeser and Kämpfer, 2014), had the highest relative abundance in Moškanjci, where nutrient availability (dissolved organic carbon and nitrogen) was the lowest. The fact that these important potential producers of EPSs and LPSs differed in their relative abundance between the sites makes it surprising that site as a factor did not have a stronger influence on the relative abundance of the investigated genes. After all, the functioning of bacterial communities has been repeatedly shown to depend on their composition (Langenheder et al., 2006; Strickland et al., 2009; Reed and Martiny, 2013; Logue et al., 2016). However, some studies also found that the functional structure can be highly conserved among bacterial communities inhabiting similar environments despite their taxonomic variability (Louca et al., 2018). This is in agreement with the theory about functional redundancy, which states that a community maintains important functions even though its members may change (Allison and Martiny, 2008). Along the same lines, Fondi et al. (2016) postulated that bacterial gene pools are shaped by broad ecological niches, such as soil, sea water, inland water or host. As the bacterial potential for EPS and LPS production (i) could possibly increase the ecological fitness of bacteria living in agricultural soils, (ii) was found in most families detected in our metagenomic datasets, including all the dominant families, (iii) was maintained at a stable level by the bacterial communities at the studied sites despite different taxonomic structures, we propose that this potential is promoted mainly in agricultural soils. Whether this applies to soils in general as well as other environments requires further investigation.

Site-Specific Response to Tillage of Potential EPS/LPS Producers

Even though the relative abundances of genes involved in the formation of EPSs and LPSs were mostly stable in our study, the aggregating capabilities of bacterial polysaccharides produced under different tillage treatments at the studied sites might have differed. In fact, even closely related bacteria can produce different types of polysaccharides (Sutherland and Thomson, 1975). At the same time, it has been shown that the slightest structural changes can lead to different properties of a polysaccharide (Suresh Kumar et al., 2007). Therefore, the aggregating capacities of polysaccharides may differ depending on which bacteria produce them. In our study, we found several bacterial families whose relative abundance of genes related to the formation of EPSs and LPSs differed under CT and RT. Even though the response of these families could be predicted to some extent, as all the responders that belonged to *Actinobacteria* had higher potential to form adhesive polysaccharides under RT compared with CT, none of the identified responsive families responded to tillage intensity at all the studied sites. This indicates that the influence of tillage on polysaccharide-producing bacteria is site-specific. Furthermore, the response of a few of the families whose potential for EPS and LPS formation was affected by tillage was inconsistent when the sites were compared. For example, the relative abundance of genes involved in the production of polysaccharides harbored by *Oxalobacteracea* was higher under CT in Moškanjci, but under RT in Juchowo. *Oxalobacteracea* are commonly found in soils, and some members of this family are employed in agriculture as plant growth-promoting agents (Baldani et al., 2014). However, little is known about the polysaccharides that they synthesize (Hiraishi et al., 1997). This underlines the importance of metagenomic studies, which enable the investigation of the community dynamics of bacterial EPS and LPS producers under natural conditions. While untargeted isolation attempts of soil-aggregating bacteria from agricultural soils yield mostly easily culturalable taxa such as *Pseudomonas* and *Bacillus* (Caesar-Tonthat et al., 2014), metagenomics could help to identify the potential key players of soil aggregation, and design more targeted isolation-based approaches. In turn, metagenomic studies, which are limited by the availability of data obtained from isolates, would benefit from additional cultivation efforts. Such complementation of different methodological approaches is especially critical, as the direct measurement of bacterial polysaccharides in soils still requires more research before they will be reliable and give more information than aggregate stability data (Redmile-Gordon et al., 2014; Costa et al., 2018).

CONCLUSION

Our study shows that the bacterial potential to form EPSs and LPSs is a possible link between soil aggregate stability, tillage intensity and soil texture. Specifically, we found that improved aggregate stability was connected with increased absolute abundance of genes related to the production of adhesive polysaccharides, and that the positive effects of RT over CT were most pronounced in the soil with a balanced content of clay, silt and sand. We propose that this could be because predominantly clayey soils are stabilized by their high clay content by itself, whereas very sandy soils lack the particles that could be easily glued together by bacterial polysaccharides into stable soil aggregates. This needs to be further investigated under more controlled conditions, as field trials are characterized by many other parameters that could also influence bacterial responses. Additionally, our results show that although the potential to produce EPSs and LPSs seems to be an important trait for bacteria in agricultural soils, as they try to maintain its stable levels within their communities, tillage intensity could have an impact on the aggregating properties of bacterial polysaccharides by inducing shifts in the community of potential polysaccharide producers. As the observed effects of tillage intensity were site-specific, and were likely connected to the differences in soil texture, we propose that further research should focus on disentangling the complicities of bacterial responses to disturbances in soils with different textures.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the Sequence Read Archive PRJNA555481.

AUTHOR CONTRIBUTIONS

BC designed the experiment, carried out the laboratory work, analyzed the data, and wrote the manuscript. GV contributed to the data analysis. MaS, RM, MK, AF, PM, and AS were responsible for the field work and chemical analyses. AF, MiS, and SS contributed to the design of the experiment. All authors edited the manuscript and approved the final draft.

FUNDING

The authors gratefully acknowledge the funding provided by the German Federal Office for Agriculture and Food (BLE, funding number: 2811OE001), the financial support for the Moškanjci trial provided by the Slovenian Research Agency (L4-9315, P4-0085), and for the Frick trial provided by the Swiss Federal Office for Agriculture (FOAG), Software AG-Stiftung, Stiftung zur Pflege von Mensch, Mitwelt und Erde, and Edith Maryon Stiftung. Moreover, this study was associated to the project Fertility Building Management Measures in Organic Cropping Systems (FertilCrop), which received funding from the CORE Organic Plus Funding Bodies, being partners of the FP7 ERA-Net project CORE Organic Plus.

ACKNOWLEDGMENTS

The authors wish to thank Susanne Kublik for performing the sequencing run, and Abilash Chakravarthy Durai Raj for bioinformatical advice.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.00568/full#supplementary-material

REFERENCES

Allison, S. D., and Martiny, J. B. H. (2008). Resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11512–11519. doi: 10.1073/pnas.0801925105

Amezketa, E. (1999). Soil aggregate stability: a review. *J. Sustain. Agr.* 14, 83–151.

Amooh, M. K., and Bonsu, M. (2015). Effects of soil texture and organic matter on evaporative loss of soil moisture. *J. Glob. Agric. Ecol.* 3, 152–161.

Asgari, H. (2014). Effect of agronomic practices on the aggregate stability and organic carbon of soil (case study: the northern of Aq Qala). *Environ. Resour. Res.* 2, 95–106.

Babin, D., Deubel, A., Jacquiod, S., Sørensen, S. J., Geistlinger, J., Grosch, R., et al. (2019). Impact of long-term agricultural management practices on soil prokaryotic communities. *Soil Biol. Biochem.* 129, 17–28. doi: 10.1016/j.soilbio.2018.11.002

Baldani, J. I., Rouws, L., Cruz, L. M., Olivares, F. L., Schmid, M., and Hartmann, A. (2014). "The family Oxalobacteraceae," in *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*, ed. E. Rosenberg (Heidelberg: Springer), 919–974. doi: 10.1007/978-3-642-30197-1_291

Bartlova, J., Badalikova, B., Pospíšilová, L., Pokorný, E., and Šarapatka, B. (2015). Water stability of soil aggregates in different systems of tillage. *Soil Water Res.* 10, 147–154. doi: 10.17221/132/2014-swr

Beare, M., Hu, S., Coleman, D., and Hendrix, P. (1997). Influences of mycelial fungi on soil aggregation and organic matter storage in conventional and no-tillage soils. *Appl. Soil Ecol.* 5, 211–219. doi: 10.1016/s0929-1393(96)00142-4

Benard, P., Zarebanadkouki, M., Brax, M., Kaltenbach, R., Jerjen, I., Marone, F., et al. (2019). Microhydrological niches in soils: how mucilage and eps alter the biophysical properties of the rhizosphere and other biological hotspots. *Vadose Zone J.* 18, 1–10. doi: 10.2136/vzj2018.12.0211

Borie, F., Rubio, R., Rouanet, J., Morales, A., Borie, G., and Rojas, C. (2006). Effects of tillage systems on soil characteristics, glomalin and mycorrhizal propagules in a Chilean Ultisol. *Soil Tillage Res.* 88, 253–261. doi: 10.1016/j.still.2005.06.004

Bronick, C. J., and Lal, R. (2005). Soil structure and management: a review. *Geoderma* 124, 3–22.

Brookes, P., Landman, A., Pruden, G., and Jenkinson, D. (1985). Chloroform fumigation and the release of soil nitrogen: a rapid direct extraction method to measure microbial biomass nitrogen in soil. *Soil Biol. Biochem.* 17, 837–842. doi: 10.1016/0038-0717(85)90144-0

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176

Caesar-Tonthat, T., Stevens, W. B., Sainju, U. M., Caesar, A. J., West, M., and Gaskin, J. F. (2014). Soil-aggregating bacterial community as affected by irrigation, tillage, and cropping system in the northern great plains. *Soil Sci.* 179, 11–20. doi: 10.1097/ss.0000000000000036

Cailliez, F. (1983). The analytical solution of the additive constant problem. *Psychometrika* 48, 305–308. doi: 10.1109/83.777085

Cania, B., Vestergaard, G., Krauss, M., Fliessbach, A., Schloter, M., and Schulz, S. (2019a). A long-term field experiment demonstrates the influence of tillage on the bacterial potential to produce soil structure-stabilizing agents such as exopolysaccharides and lipopolysaccharides. *Environ. Microbiome* 1, 1–14.

Cania, B., Vestergaard, G., Kublik, S., Köhne, J. M., Fischer, T., Albert, A., et al. (2019b). Biological soil crusts from different soil substrates harbor distinct bacterial groups with the potential to produce exopolysaccharides and lipopolysaccharides. *Microb. Ecol.* 79, 326–341. doi: 10.1007/s00248-019-01415-6

Cheng, F., Peng, X., Zhao, P., Yuan, J., Zhong, C., Cheng, Y., et al. (2013). Soil microbial biomass, basal respiration and enzyme activity of main forest types in the Qinling Mountains. *PLoS One* 8:e67353. doi: 10.1371/journal.pone.0067353

Cookson, W. R., Murphy, D. V., and Roper, M. M. (2008). Characterizing the relationships between soil organic matter components and microbial function and composition along a tillage disturbance gradient. *Soil Biol. Biochem.* 40, 763–777. doi: 10.1016/j.soilbio.2007.10.011

Cooper, J., Baranski, M., Stewart, G., Nobel-De Lange, M., Bàrberi, P., Fließbach, A., et al. (2016). Shallow non-inversion tillage in organic farming maintains crop yields and increases soil C stocks: a meta-analysis. *Agron. Sustain. Dev.* 36, 1–20.

Costa, O. Y., Raaijmakers, J. M., and Kuramae, E. E. (2018). Microbial extracellular polymeric substances: ecological function and impact on soil aggregation. *Front. Microbiol.* 9:1636. doi: 10.3389/fmicb.2018.01636

Dai, J., Hu, J., Zhu, A., Bai, J., Wang, J., and Lin, X. (2015). No tillage enhances arbuscular mycorrhizal fungal population, glomalin-related soil protein content, and organic carbon accumulation in soil macroaggregates. *J. Soil Sediment.* 15, 1055–1062. doi: 10.1007/s11368-015-1091-9

de Souza, J. A. M., Alves, L. M. C., De Mello Varani, A., and De Macedo Lemos, E. G. (2014). "The family Bradyrhizobiaceae," in *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*, eds E. Rosenberg, E. F. DeLong,

S. Lory, E. Stackebrandt, and F. Thompson (Heidelberg: Springer), 135–154. doi: 10.1007/978-3-642-30197-1_253

De Vries, M., Schöler, A., Ertl, J., Xu, Z., and Schloter, M. (2015). Metagenomic analyses reveal no differences in genes involved in cellulose degradation under different tillage treatments. *FEMS Microbiol. Ecol.* 91:fiv069. doi: 10.1093/femsec/fiv069

ELD Initiative (2015). *The Value of Land: Prosperous Lands and Positive Rewards Through Sustainable Land Management*. Available online at: www.eld-initiative.org (accessed June 18, 2019).

FAO, (2017). *Voluntary Guidelines for Sustainable Soil Management*.Rome: Food and Agriculture Organization of the United Nations.

Field, A., Miles, J., and Field, Z. (2012). *Discovering Statistics Using R*.Thousand Oaks, CA: Sage Publications.

Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat. Rev. Microbiol.* 15, 579–590. doi: 10.1038/nrmicro.2017.87

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279-D285. doi: 10.1093/nar/gkv1344

Fondi, M., Karkman, A., Tamminen, M. V., Bosi, E., Virta, M., Fani, R., et al. (2016). "Every gene is everywhere but the environment selects": global geolocalization of gene sharing in environmental samples through network analysis. *Genome Biol. Evol.* 8, 1388–1400. doi: 10.1093/gbe/evw077

Glaeser, S. P., and Kämpfer, P. (2014). "The family Sphingomonadaceae," in *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*, eds E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Heidelberg: Springer), 641–707. doi: 10.1007/978-3-642-30197-1_302

Grafe, M., Goers, M., Tucher, S. V., Baum, C., Zimmer, D., Leinweber, P., et al. (2018). Bacterial potentials for uptake, solubilization and mineralization of extracellular phosphorus in agricultural soils are highly stable under different fertilization regimes. *Environ. Microbiol. Rep.* 10, 320–327. doi: 10.1111/1758-2229.12651

Grandy, A. S., Robertson, G. P., and Thelen, K. D. (2006). Do productivity and environmental trade-offs justify periodically cultivating no-till cropping systems? *Agron. J.* 98, 1377–1383. doi: 10.2134/agronj2006.0137

Haft, D. H., Selengut, J. D., Richter, R. A., Harkins, D., Basu, M. K., and Beck, E. (2013). TIGRFAMs and genome properties in 2013. *Nucleic Acids Res.* 41, D387–D395. doi: 10.1093/nar/gks1234

Hiraishi, A., Shin, Y. K., and Sugiyama, J. (1997). Proposal to reclassify zoogloea ramigera IAM 12670 (PR Dugan 115) as *Duganella zoogloeoides* gen. nov., sp. nov. *Int. J. Syst. Evol. Microbiol.* 47, 1249–1252. doi: 10.1099/00207713-47-4-1249

Holland, J. (2004). The environmental consequences of adopting conservation tillage in Europe: reviewing the evidence. *Agric. Ecosyst. Environ.* 103, 1–25. doi: 10.1016/j.agee.2003.12.018

Jacobs, A., Rauber, R., and Ludwig, B. (2009). Impact of reduced tillage on carbon and nitrogen storage of two Haplic Luvisols after 40 years. *Soil Tillage Res.* 102, 158–164. doi: 10.1016/j.still.2008.08.012

Jacques, M. (1996). Role of lipo-oligosaccharides and lipopolysaccharides in bacterial adherence. *Trends Microbiol.* 4, 408–409.

Joergensen, R. G. (1996). The fumigation-extraction method to estimate soil microbial biomass: calibration of the kEC value. *Soil Biol. Biochem.* 28, 25–31. doi: 10.1016/0038-0717(95)00102-6

Joergensen, R. G., and Mueller, T. (1996). The fumigation-extraction method to estimate soil microbial biomass: calibration of the kEN value. *Soil Biol. Biochem.* 28, 33–37. doi: 10.1016/0038-0717(95)00101-8

Kierek, K., and Watnick, P. I. (2003). The *Vibrio cholerae* O139 O-antigen polysaccharide is essential for Ca2+-dependent biofilm development in sea water. *Proc. Natl. Acad. Sci. U.S.A.* 100, 14357–14362. doi: 10.1073/pnas.2334614100

Kihara, J., Martius, C., Bationo, A., Thuita, M., Lesueur, D., Herrmann, L., et al. (2012). Soil aggregation and total diversity of bacteria and fungi in various tillage systems of sub-humid and semi-arid Kenya. *Appl. Soil Ecol.* 58, 12–20. doi: 10.1016/j.apsoil.2012.03.004

Kopylova, E., Noé, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28, 3211–3217. doi: 10.1093/bioinformatics/bts611

Langenheder, S., Lindström, E. S., and Tranvik, L. J. (2006). Structure and function of bacterial communities emerging from different sources under identical conditions. *Appl. Environ. Microbiol.* 72, 212–220. doi: 10.1128/aem.72.1.212-220.2006

Le Bissonnais, Y., Singer, M., and Bradford, J. (1993). "Assessment of soil erodibility: the relationship between soil properties, erosion processes and susceptibility to erosion," in *Farm Land Erosion in Temperate Plains Environment and Hills*, ed. S. Wicherek (Amsterdam: Elsevier), 87–96. doi: 10.1016/b978-0-444-81466-1.50012-5

Lehmann, A., Zheng, W., and Rillig, M. C. (2017). Soil biota contributions to soil aggregation. *Nat. Ecol. Evol.* 1, 1828–1835. doi: 10.1038/s41559-017-0344-y

Lindhout, T., Lau, P. C. Y., Brewer, D., and Lam, J. S. (2009). Truncation in the core oligosaccharide of lipopolysaccharide affects flagella-mediated motility in *Pseudomonas aeruginosa* PAO1 via modulation of cell surface attachment. *Microbiology* 155(Pt 10), 3449–3460. doi: 10.1099/mic.0.030510-0

Logue, J. B., Stedmon, C. A., Kellerman, A. M., Nielsen, N. J., Andersson, A. F., Laudon, H., et al. (2016). Experimental insights into the importance of aquatic bacterial community composition to the degradation of dissolved organic matter. *ISME J.* 10, 533–545. doi: 10.1038/ismej.2015.131

Louca, S., Polz, M. F., Mazel, F., Albright, M. B., Huber, J. A., O'connor, M. I., et al. (2018). Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.* 2, 936–943. doi: 10.1038/s41559-018-0519-1

Lu, X., Lu, X., and Liao, Y. (2018). Effect of tillage treatment on the diversity of soil arbuscular mycorrhizal fungal and soil aggregate-associated carbon content. *Front. Microbiol.* 9:2986. doi: 10.3389/fmicb.2018.02986

Lueders, T., Manefield, M., and Friedrich, M. W. (2004). Enhanced sensitivity of DNA- and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. *Environ. Microbiol.* 6, 73–78. doi: 10.1046/j.1462-2920.2003.00536.x

Menzel, P., Ng, K. L., and Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7:11257. doi: 10.1038/ncomms11257

Mikha, M. M., Vigil, M. F., and Benjamin, J. G. (2013). Long-term tillage impacts on soil aggregation and carbon dynamics under wheat-fallow in the central Great Plains. *Soil Sci. Soc. Am. J.* 77, 594–605. doi: 10.2136/sssaj2012.0125

Mistry, J., Finn, R. D., Eddy, S. R., Bateman, A., and Punta, M. (2013). Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* 41:e121. doi: 10.1093/nar/gkt263

Murdoch, D., and Chow, E. D. (2018). *ellipse: Functions for Drawing Ellipses and Ellipse-Like Confidence Regions. R Package Version 0.4.1*. Available online at: https://CRAN.R-project.org/package=ellipse (accessed January 05, 2018).

Murer, E. J., Baumgarten, A., Eder, G., Gerzabek, M. H., Kandeler, E., and Rampazzo, N. (1993). "An improved sieving machine for estimation of soil aggregate stability (SAS)," in *Soil Structure/Soil Biota Interrelationships*, eds L. Brussaard, and M. J. Kooistra, (Amsterdam: Elsevier), 539–547. doi: 10.1016/b978-0-444-81490-6.50043-1

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., et al. (2018). *vegan: Community Ecology Package. R Package Version 2.5-1*. Available online at: https://CRAN.R-project.org/package=vegan (accessed September 1, 2019).

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290. doi: 10.1093/bioinformatics/btg412

Pereira, S. B., Mota, R., Santos, C. L., De Philippis, R., and Tamagnini, P. (2013). Assembly and export of extracellular polymeric substances (EPS) in cyanobacteria: a phylogenomic approach. *Adv. Bot. Res.* 65, 235–279. doi: 10.1016/b978-0-12-394313-2.00007-x

R Core Team, (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Redmile-Gordon, M., Brookes, P., Evershed, R., Goulding, K., and Hirsch, P. (2014). Measuring the soil-microbial interface: extraction of extracellular polymeric substances (EPS) from soil biofilms. *Soil Biol. Biochem.* 72, 163–171. doi: 10.1016/j.soilbio.2014.01.025

Redmile-Gordon, M., Gregory, A., White, R., and Watts, C. (2020). Soil organic carbon, extracellular polymeric substances (EPS), and soil structural stability as affected by previous and current land-use. *Geoderma* 363:114143. doi: 10.1016/j.geoderma.2019.114143

Reed, H. E., and Martiny, J. B. (2013). Microbial composition affects the functioning of estuarine sediments. *ISME J.* 7, 868–879. doi: 10.1038/ismej.2012.154

Reeve, W., Terpolilli, J., Melino, V., Ardley, J., Tian, R., De Meyer, S., et al. (2013). Genome sequence of the lupin-nodulating *Bradyrhizobium* sp. strain WSM1417. *Stand. Genomic Sci.* 9, 273–282. doi: 10.4056/sigs.4518260

Rho, M., Tang, H., and Ye, Y. (2010). FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res.* 38:e191. doi: 10.1093/nar/gkq747

Rosenberg, E., Delong, E. F., Lory, S., Stackebrandt, E., and Thompson, F. (2014). *The Prokaryotes: Actinobacteria*. Heidelberg: Springer.

Rutherford, P. M., and Juma, N. G. (1992). Influence of texture on habitable pore space and bacterial-protozoan populations in soil. *Biol. Fertil. Soils* 12, 221–227. doi: 10.1007/bf00336036

Schmieder, R., and Edwards, R. (2011). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PLoS One* 6:e17288. doi: 10.1371/journal.pone.0017288

Schubert, M., Lindgreen, S., and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* 9:88. doi: 10.1186/s13104-016-1900-2

Sheehy, J., Regina, K., Alakukku, L., and Six, J. (2015). Impact of no-till and reduced tillage on aggregation and aggregate-associated carbon in Northern European agroecosystems. *Soil Tillage Res.* 150, 107–113. doi: 10.1016/j.still.2015.01.015

Singh, P., Heikkinen, J., Mitra, S., and Regina, K. (2016). "Impact of tillage and residue management practices on soil aggregates and soil organic carbon," in *Geostatistical and Geospatial Approaches for the Characterization of Natural Resources in the Environment*, ed. N. J. Raju (Cham: Springer), 381–384. doi: 10.1007/978-3-319-18663-4_58

Six, J., Bossuyt, H., Degryze, S., and Denef, K. (2004). A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. *Soil Tillage Res.* 79, 7–31. doi: 10.1016/j.still.2004.03.008

Stockfisch, N., Forstreuter, T., and Ehlers, W. (1999). Ploughing effects on soil organic matter after twenty years of conservation tillage in Lower Saxony, Germany. *Soil Tillage Res.* 52, 91–101. doi: 10.1016/s0167-1987(99)00063-x

Strickland, M. S., Lauber, C., Fierer, N., and Bradford, M. A. (2009). Testing the functional significance of microbial community composition. *Ecology* 90, 441–451. doi: 10.1890/08-0296.1

Suresh Kumar, A., Mody, K., and Jha, B. (2007). Bacterial exopolysaccharides–a perception. *J. Basic Microbiol.* 47, 103–117. doi: 10.1002/jobm.200610203

Sutherland, I., and Thomson, S. (1975). Comparison of polysaccharides produced by *Myxococcus* strains. *Microbiology* 89, 124–132. doi: 10.1099/00221287-89-1-124

Sutherland, I. W. (2001). Biofilm exopolysaccharides: a strong and sticky framework. *Microbiology* 147, 3–9. doi: 10.1099/00221287-147-1-3

Totsche, K. U., Amelung, W., Gerzabek, M. H., Guggenberger, G., Klumpp, E., Knief, C., et al. (2018). Microaggregates in soils. *J. Plant Nutr. Soil Sci.* 181, 104–136.

Vance, E. D., Brookes, P. C., and Jenkinson, D. S. (1987). An extraction method for measuring soil microbial biomass C. *Soil Biol. Biochem.* 19, 703–707. doi: 10.1016/0038-0717(87)90052-6

Vestergaard, G., Schulz, S., Schöler, A., and Schloter, M. (2017). Making big data smart–how to use metagenomics to understand soil quality. *Biol. Fertil. Soils* 53, 479–484. doi: 10.1007/s00374-017-1191-3

Walkley, C., and Black, I. (1934). An examination of Degtjareff method of determining soil organic matter and proposed modification of the method of the chromic acid titration method. *Soil Sci.* 37, 29–39.

Wang, Y., Xu, J., Shen, J., Luo, Y., Scheu, S., and Ke, X. (2010). Tillage, residue burning and crop rotation alter soil fungal community and water-stable aggregation in arable fields. *Soil Tillage Res.* 107, 71–79. doi: 10.1016/j.still.2010.02.008

Weil, R. R., and Brady, N. C. (2017). *The Nature and Properties of Soils*.Harlow: Pearson Education.

Whitfield, C., and Trent, M. S. (2014). Biosynthesis and export of bacterial lipopolysaccharides. *Annu. Rev. Biochem.* 83, 99–128. doi: 10.1146/annurev-biochem-060713-035600

Wilcox, R. R., and Schönbrodt, F. D. (2014). *The WRS Package for Robust Statistics in R*. R Package Version 0.24.

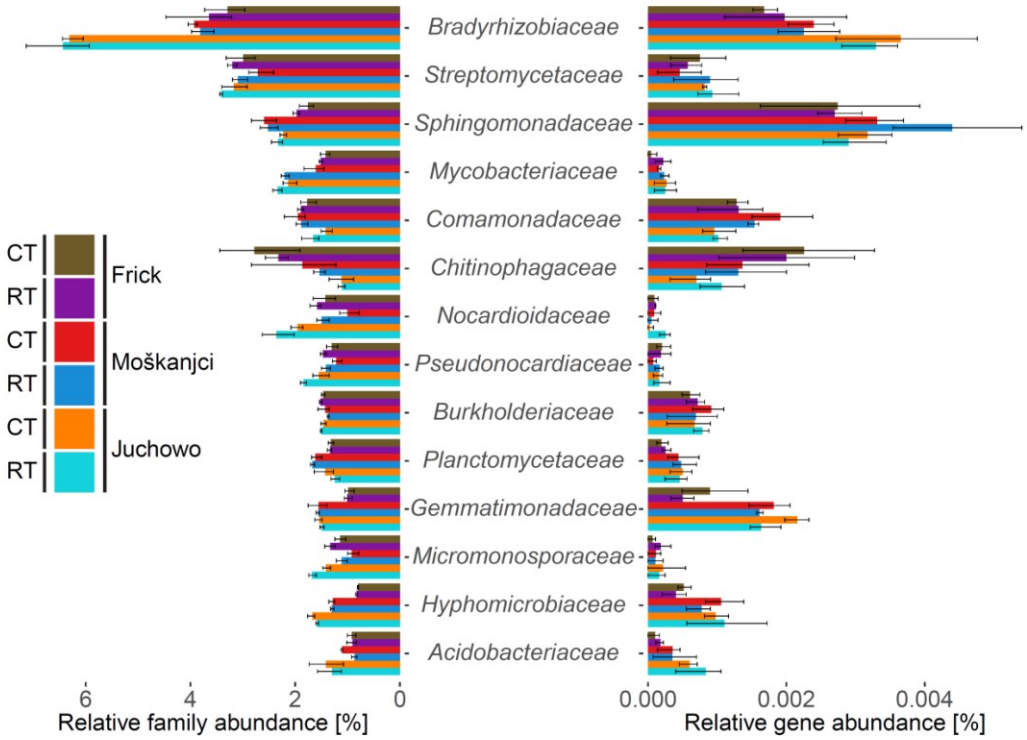
Wright, S., Starr, J., and Paltineanu, I. (1999). Changes in aggregate stability and concentration of glomalin during tillage management transition. *Soil Sci. Soc. Am. J.* 63, 1825–1829. doi: 10.2136/sssaj1999.6361825x

Zhang, Z., Qu, Y., Li, S., Feng, K., Wang, S., Cai, W., et al. (2017). Soil bacterial quantification approaches coupling with relative abundances reflecting the changes of taxa. *Sci. Rep.* 7:4837. doi: 10.1038/s41598-017-05260-w

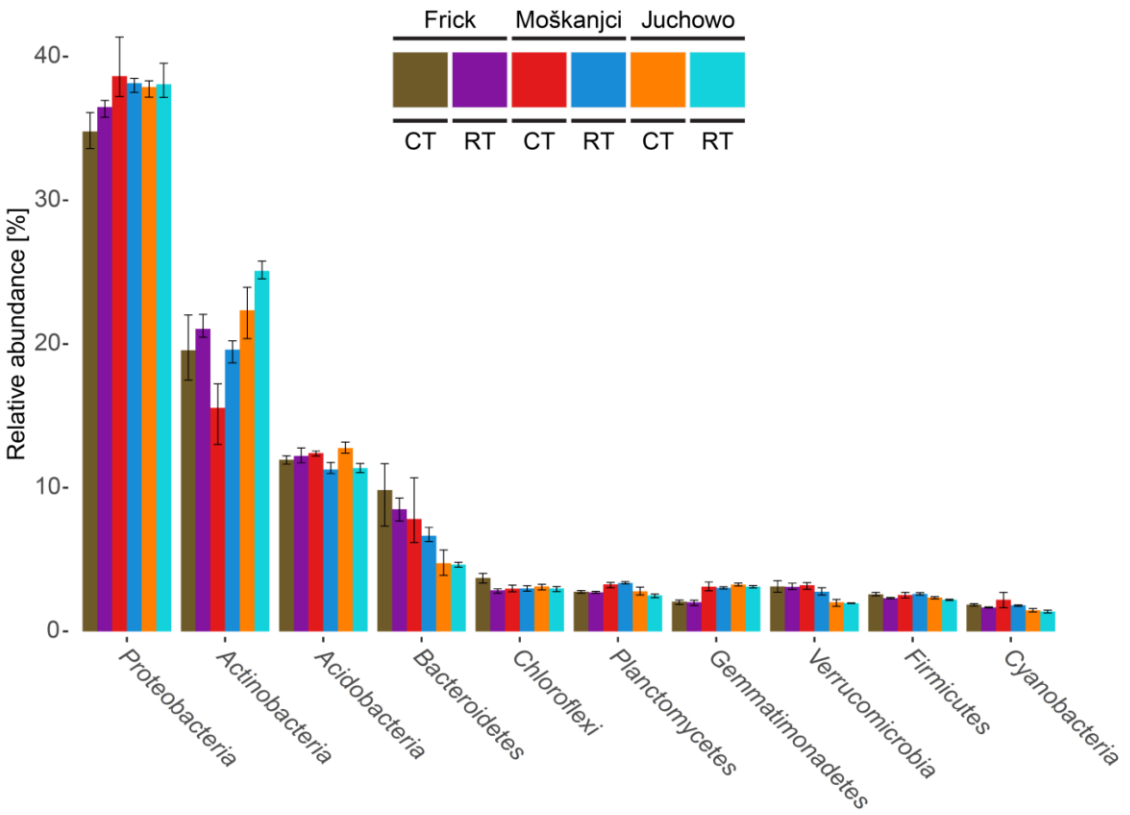
Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers

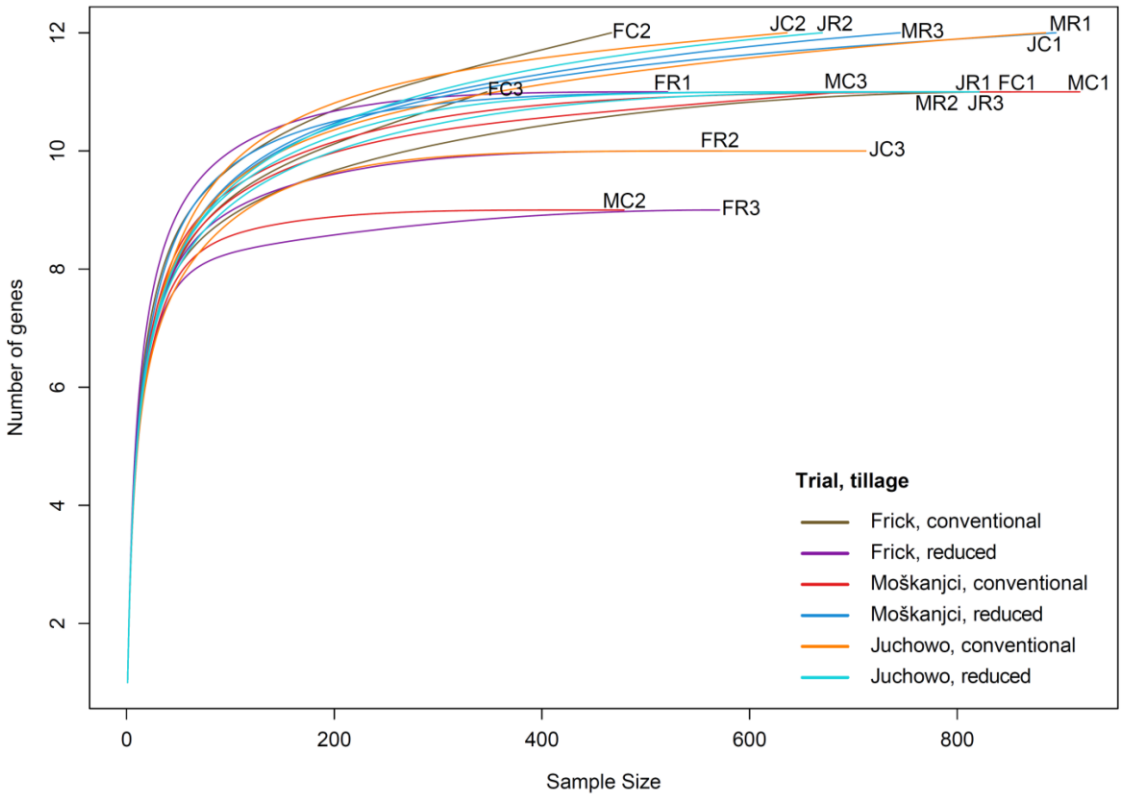
Supplementary Material



Supplementary Figure 1. Comparison of the relative abundances of dominant bacterial families and the relative abundances of their sequences related to EPS and LPS formation. Error bars show standard deviations.



Supplementary Figure 2. Relative abundances of the ten most dominant phyla. Error bars show standard deviations.



Supplementary Figure 3. Number of assigned genes related to the formation of EPSs and LPSs depicted as a function of sequencing depth. "F", "M" and "J" at the beginning of the sample names stand for "Frick", "Moškanjci" and "Juchowo", respectively. The following letter "C" or "R" stands for conventional or reduced tillage. The number distinguishes the replicates.

Supplementary Table 1. Details of the raw and filtered sequencing data. "F", "M" and "J" at the beginning of the sample names stand for "Frick", "Moškanjci" and "Juchowo", respectively. The following letter "C" or "R" stands for conventional or reduced tillage. The number distinguishes the replicates.

Frick						
Raw data	FC1	FC2	FC3	FR1	FR2	FR3
Number of reads	3 666 752	2 103 714	1 648 448	2 516 708	2 335 156	2 311 802
Average length of reads [bp]	301.00	301.00	301.00	301.00	301.00	301.00
Filtered data						
Number of reads	3 666 146	2 102 832	1 647 552	2 516 294	2 334 850	2 311 356
Average length of reads [bp]	297.71	293.97	291.52	297.59	297.91	297.87

Moškanjci						
Raw data	MC1	MC2	MC3	MR1	MR2	MR3
Number of reads	3 381 440	2 009 488	2 860 256	3 701 668	2 644 092	2 913 230
Average length of reads [bp]	301.00	301.00	301.00	301.00	301.00	301.00
Filtered data						
Number of reads	3 380 688	2 009 194	2 859 498	3 701 194	2 643 614	2 912 892
Average length of reads [bp]	297.72	297.17	298.04	297.98	297.94	298.12

Juchowo						
Raw data	JC1	JC2	JC3	JR1	JR2	JR3
Number of reads	3 270 470	2 402 428	2 856 936	3 058 810	2 678 352	3 178 412
Average length of reads [bp]	301.00	301.00	301.00	301.00	301.00	301.00
Filtered data						
Number of reads	3 269 872	2 401 852	2 856 422	3 058 184	2 677 828	3 177 714
Average length of reads [bp]	297.81	297.76	297.91	297.51	297.64	297.55

Supplementary Table 2. Bacterial families whose relative abundances and potential to produce EPSs or LPSs were significantly affected by soil trial, tillage or their interaction. Significant differences were determined by a robust 2-way ANOVA (n = 3, $p < 0.05$).

Factor	Families whose abundance was affected	Families whose potential to produce EPS or LPS was affected
Trial	*Acaryochloridaceae, Acetobacteraceae, Acidiferrobacteraceae, Acidimicrobiaceae, Acidobacteriaceae, Acidothermaceae, Actinopolysporaceae, Actinospicaceae, Aerococcaceae, Aeromonadaceae, Akkermansiaceae, Alcaligenaceae, Alteromonadaceae, Amoebophilaceae, Anaerolineaceae, Anaeromyxobacteraceae, Aphanothecaceae, Aquificaceae, Archangiaceae, Ardenticatenaceae, Aurantimonadaceae, Bacillaceae, Bacteriovoracaceae, Bacteroidaceae, Balneolaceae, Bdellovibrionaceae, Beijerinckiaceae, Beutenbergiaceae, Bogoriellaceae, Bradyrhizobiaceae, Brevibacteriaceae, Brevinemataceae, Brucellaceae, Caldilineaceae, Campylobacteraceae, Candidatus Actinomarinaceae, Candidatus Brocadiaceae, Candidatus Desulfofervidaceae, Candidatus Midichloriaceae, Cardiobacteriaceae, Carnobacteriaceae, Catalimonadaceae, Catenulisporaceae, Cellulomonadaceae, Cellvibrionaceae, Chamaesiphonaceae, Chitinophagaceae, Chlorobiaceae, Chlorogloeopsidaceae, Chromobacteriaceae, Chroococcaceae, Chroococcidiopsidaceae, Chrysiogenaceae, Chthoniobacteraceae, Chthonomonadaceae, Clostridiaceae, Coleofasciculaceae, Colwelliaceae, Comamonadaceae, Competibacteraceae, Conexibacteraceae, Corynebacteriaceae, Coxiellaceae, Crocinitomicaceae, Cryomorphaceae, Cryptosporangiaceae, Cyanothecaceae, Cyclobacteriaceae, Cytophagaceae, Dermacoccaceae, Desulfobacteraceae, Desulfobulbaceae, Desulfohalobiaceae, Desulfomicrobiaceae, Desulfovibrionaceae, Desulfuromonadaceae, Dietziaceae, Ectothiorhodospiraceae, Enterobacteriaceae, Enterococcaceae, Erwiniaceae, Erysipelotrichaceae, Erythrobacteraceae, Eubacteriaceae, Ferrimonadaceae, Fibrobacteraceae, Fimbriimonadaceae, Flammeovirgaceae, Flavobacteriaceae, Frankiaceae, Fusobacteriaceae, Gaiellaceae, Gallionellaceae, Gemmataceae, Gemmatimonadaceae, Geobacteraceae, Geodermatophilaceae, Gomontiellaceae, Gordoniaceae, Hahellaceae, Halobacteriovoraceae, Hapalosiphonaceae, Helicobacteraceae, Herpetosiphonaceae, Holophagaceae, Hydrogenophilaceae, Hyellaceae, Hymenobacteraceae, Hyphomicrobiaceae, Idiomarinaceae, Ignavibacteriaceae, Immundisolibacteraceae, Intrasporangiaceae, Isosphaeraceae, Jiangellaceae, Jonesiaceae, Kangiellaceae, Kineosporiaceae, Kiritimatiellaceae, Kofleriaceae, Ktedonobacteraceae, Lachnospiraceae, Legionellaceae, Lentimicrobiaceae, Lentisphaeraceae, Leptolyngbyaceae, Leptospiraceae, Magnetococcaceae, Marinifilaceae, Marinilabiliaceae, Melioribacteraceae, Methylococcaceae, Methylocystaceae, Methylophilaceae, Methylothermaceae, Microbulbiferaceae, Micrococcaceae, Microcoleaceae, Micropepsaceae, Microthrixaceae, Moraxellaceae, Moritellaceae, Mycoplasmataceae, Myxococcaceae, Nakamurellaceae, Nannocystaceae, Nitriliruptoraceae, Nitrosomonadaceae, Nitrospinaceae, Nitrospiraceae, Nocardiaceae, Nocardiopsaceae, Nostocaceae, Oceanospirillaceae, Odoribacteraceae, Oleiphilaceae, Opitutaceae, Orbaceae, Oscillatoriaceae, Oscillospiraceae, Oxalobacteraceae, Pasteurellaceae, Patulibacteraceae, Pectobacteriaceae, Peptococcaceae, Peptoniphilaceae, Persicobacteraceae, Phycisphaeraceae, Phyllobacteriaceae, Piscirickettsiaceae, Planctomycetaceae, Polyangiaceae, Porphyromonadaceae, Porticoccaceae, Prevotellaceae, Prolixibacteraceae, Propionibacteriaceae, Pseudoalteromonadaceae, Pseudomonadaceae, Pseudonocardiaceae, Psychromonadaceae, Puniceicoccaceae, Rhizobiaceae, Rhodanobacteraceae, Rhodobiaceae, Rhodocyclaceae, Rhodospirillaceae, Rhodothermaceae, Rickettsiaceae, Rikenellaceae, Ruaniaceae, Rubritaleaceae, Rubrobacteraceae, Ruminococcaceae, Saccharospirillaceae, Sandaracinaceae, Sanguibacteraceae, Saprospiraceae, Schleiferiaceae, Scytonemataceae, Shewanellaceae, Solibacteraceae, Solirubrobacteraceae, Sphingobacteriaceae, Sphingomonadaceae, Spirochaetaceae, Spongiibacteraceae, Sporichthyaceae, Streptococcaceae, Streptosporangiaceae, Succinivibrionaceae, Sutterellaceae, Symphyonemataceae, Synechococcaceae, Synergistaceae, Syntrophaceae, Syntrophobacteraceae, Syntrophorhabdaceae, Thermaceae, Thermoanaerobacteraceae, Thermoanaerobacterales Family IV. Incertae Sedis, Thermomonosporaceae, Thermonemataceae, Thiotrichaceae, Tolypothrichaceae, Trueperaceae, Tsukamurellaceae, Verrucomicrobia subdivision 3, Verrucomicrobia subdivision 6, Verrucomicrobiaceae, Vibrionaceae, Vulgatibacteraceae, Wenzhouxiangellaceae, Williamsiaceae, Woeseiaceae, Xanthobacteraceae, Xenococcaceae*	*Acidobacteriaceae, Bdellovibrionaceae, Bradyrhizobiaceae, Comamonadaceae, Ectothiorhodospiraceae, Gemmatimonadaceae, Hyphomicrobiaceae, Micrococcaceae, Nocardiaceae, Opitutaceae, Sinobacteraceae*
Tillage	*Glycomycetaceae*	-
Trial + tillage	*Dermatophilaceae, Micromonosporaceae, Mycobacteriaceae, Nocardioidaceae*	-
Trial x tillage	*Acidaminococcaceae, Chromatiaceae, Halieaceae, Halobacteroidaceae, Halomonadaceae, Methylobacteriaceae, Parvularculaceae, Sphaerobacteraceae*	*Balneolaceae, Halieaceae, Mariprofundaceae, Oxalobacteraceae, Prolixibacteraceae, Rhodospirillaceae*

Supplementary Table 3. Influence of soil trial, tillage and their interaction on the relative abundances of the dominant bacterial families, as determined by a robust 2-way ANOVA. Effect sizes (ω^2) and significance levels were calculated based on triplicates (n = 3). Significance levels are represented by the amount of stars: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$.

Robust ANOVA (effect sizes and significance levels)						
Family	Trial		Tillage		Trial x Tillage	
Bradyrhizobiaceae	0.90	***	0.00	ns	0.00	ns
Streptomycetaceae	0.27	ns	0.22	ns	0.00	ns
Sphingomonadaceae	0.77	***	0.00	ns	0.01	ns
Mycobacteriaceae	0.67	***	0.15	***	0.07	ns
Comamonadaceae	0.48	***	0.02	ns	0.02	ns
Chitinophagaceae	0.58	***	0.01	ns	0.00	ns
Nocardioidaceae	0.68	***	0.14	***	0.00	ns
Pseudonocardiaceae	0.55	***	0.24	ns	0.00	ns
Burkholderiaceae	0.30	ns	0.00	ns	0.06	ns
Planctomycetaceae	0.64	***	0.00	ns	0.05	ns
Gemmatimonadaceae	0.88	***	0.00	ns	0.00	ns
Micromonosporaceae	0.70	***	0.17	*	0.00	ns
Hyphomicrobiaceae	0.98	***	0.00	ns	0.00	ns
Acidobacteriaceae	0.52	**	0.03	ns	0.00	ns

Supplementary Table 4. Influence of trial, tillage and their interaction on the relative abundances of bacterial reads assigned to genes encoding for proteins involved in EPS and LPS formation, as determined by a robust 2-way ANOVA. Effect sizes (ω^2) and significance levels were calculated based on triplicates (n = 3). Significance levels are represented by the amount of stars: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$.

Robust ANOVA (effect sizes and significance levels)						
Gene	Trial		Tillage		Trial x Tillage	
wza	0.40	***	0.00	ns	0.07	ns
wcaB	0.13	ns	0.00	ns	0.02	ns
wcaF	0.00	ns	0.06	ns	0.01	ns
wcaK/amsJ	0.03	ns	0.00	ns	0.02	ns
kpsE	0.07	ns	0.00	ns	0.03	ns
algE	0.00	ns	0.00	ns	0.00	ns
algJ	0.00	ns	0.00	ns	0.08	ns
sacB	0.03	ns	0.00	ns	0.15	ns
wzt	0.00	ns	0.00	ns	0.11	ns
lptC	0.20	ns	0.00	ns	0.01	ns
lptF	0.17	ns	0.00	ns	0.00	ns
lptG	0.27	ns	0.00	ns	0.00	ns

Supplementary Table 5. General influence of trial, tillage and their interaction, as well as site-specific influence of tillage in Frick, Moškanjci and Juchowo on the estimated absolute abundances of the EPS and LPS genes harbored by the dominant bacterial families, as determined by a robust 2-way ANOVA and a robust t-test, respectively. Corresponding effect sizes (ω^2 and r) and significance levels were calculated based on triplicates (n = 3). Significance levels are represented by the amount of stars: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$.

Robust ANOVA (effect sizes and significance levels)

Gene	Trial		Tillage		Trial x Tillage	
wza	0.84	***	0.04	ns	0.03	ns
wcaB	0.90	***	0.02	ns	0.00	ns
wcaF	0.82	***	0.03	ns	0.00	ns
wcaK/amsJ	0.00	ns	0.00	ns	0.27	ns
kpsE	0.90	***	0.03	ns	0.02	ns
algE	0.10	ns	0.00	ns	0.00	ns
algJ	0.52	***	0.01	ns	0.00	ns
sacB	0.28	***	0.01	ns	0.00	ns
wzt	0.85	***	0.04	*	0.04	ns
lptC	0.75	***	0.03	ns	0.02	ns
lptF	0.85	***	0.03	ns	0.01	ns
lptG	0.84	***	0.03	ns	0.01	ns

Robust t-test (effect sizes and significance levels)

Gene	Frick		Moškanjci		Juchowo	
wza	0.83	***	0.91	***	0.14	ns
wcaB	0.69	ns	0.82	***	0.25	ns
wcaF	0.53	ns	0.89	***	0.35	ns
wcaK/amsJ	0.81	ns	0.66	ns	0.38	ns
kpsE	0.88	***	0.53	ns	0.37	ns
algE	0.18	ns	0.64	ns	0.10	ns
algJ	0.46	ns	0.77	***	0.41	ns
sacB	0.28	ns	0.80	***	0.94	***
wzt	0.89	***	0.81	***	0.31	ns
lptC	0.65	ns	0.58	ns	0.23	ns
lptF	0.71	ns	0.85	***	0.01	ns
lptG	0.69	ns	0.89	***	0.00	ns

Supplementary Table 6. Influence of trial, tillage and their interaction on the relative abundances of the EPS and LPS genes harbored by the dominant bacterial families, as determined by a robust 2-way ANOVA. Effect sizes (ω^2) and significance levels were calculated based on triplicates (n = 3). Significance levels are represented by the amount of stars: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$.

Robust ANOVA (effect sizes and significance levels)						
Family	Trial		Tillage		Trial x Tillage	
Bradyrhizobiaceae	0.55	***	0.00	ns	0.00	ns
Streptomycetaceae	0.00	ns	0.00	ns	0.03	ns
Sphingomonadaceae	0.26	ns	0.00	ns	0.05	ns
Mycobacteriaceae	0.09	ns	0.06	ns	0.02	ns
Comamonadaceae	0.47	***	0.00	ns	0.00	ns
Chitinophagaceae	0.31	ns	0.00	ns	0.00	ns
Nocardioidaceae	0.04	ns	0.10	ns	0.30	ns
Pseudonocardiaceae	0.00	ns	0.00	ns	0.00	ns
Burkholderiaceae	0.00	ns	0.00	ns	0.00	ns
Planctomycetaceae	0.29	ns	0.00	ns	0.00	ns
Gemmatimonadaceae	0.72	***	0.08	ns	0.00	ns
Micromonosporaceae	0.00	ns	0.00	ns	0.00	ns
Hyphomicrobiaceae	0.39	***	0.00	ns	0.00	ns
Acidobacteriaceae	0.52	***	0.00	ns	0.00	ns

Supplementary Table 7. Influence of tillage on the overall relative abundance and the abundance of EPS and LPS genes of families whose potential for the formation of EPSs and LPSs was affected by tillage in each trial. Effect sizes (ω^2) and significance levels were calculated based on triplicates (n = 3). Significance levels are represented by the amount of stars: 1 – $p < 0.05$, 2 – $p < 0.01$, 3 – $p < 0.001$.

Robust t-test (effect sizes and significance levels)					
Trial	Family	Relative abundance of families		Relative abundance of EPS/LPS genes	
Frick	*Cytophagaceae*	0.50	ns	0.84	***
	Alteromonadaceae	0.12	ns	0.81	***
	Porphyromonadaceae	0.52	ns	0.75	***
	Flavobacteriaceae	0.59	ns	0.71	***
	Geodermatophilaceae	0.28	ns	0.92	***
	Acetobacteraceae	0.64	ns	1.00	***
	Aurantimonadaceae	0.26	ns	0.95	***
	Enterobacteriaceae	0.47	ns	0.89	***
	Myxococcaceae	0.64	ns	0.96	***
	Cellulomonadaceae	0.42	ns	0.94	***
	Rhodospirillaceae	0.48	ns	0.97	***
Moškanjci	*Oxalobacteraceae*	0.80	***	0.91	***
	Cellvibrionaceae	0.59	ns	0.89	***
	Opitutaceae	0.49	ns	0.84	***
	Cytophagaceae	0.31	ns	0.75	***
	Mycobacteriaceae	0.94	***	0.78	***
	Rhodobiaceae	0.31	ns	0.73	***
	Alcaligenaceae	0.20	ns	0.71	***
	Frankiaceae	0.75	***	0.99	***
	Isosphaeraceae	0.18	ns	0.86	***
	Balneolaceae	0.63	***	0.97	***
	Gemmataceae	0.68	ns	0.83	***
	Hapalosiphonaceae	0.43	ns	0.87	***
Juchowo	*Desulfobulbaceae*	0.31	ns	0.93	***
	Mariprofundaceae	0.31	ns	0.99	***
	Gemmatimonadaceae	0.43	ns	0.84	***
	Flavobacteriaceae	0.00	ns	0.88	***
	Halieaceae	0.89	***	1.00	***
	Oxalobacteraceae	0.29	ns	0.91	***
	Acetobacteraceae	0.25	ns	0.69	***
	Nocardioidaceae	0.80	ns	0.92	***
	Polyangiaceae	0.82	***	0.76	***