Effect of tomato genotype on microbiome composition

Simão Pedro Martins Correia

Master's in Agricultural Engineering
Faculty of Sciences of the University of Porto
Department of Geosciences, Environment and Spatial Plannings
2017

Supervisor
Susana Maria Pinto de Carvalho, Professor, Faculty of Sciences of the University of Porto

Co-Supervisors
Birgit Mitter, PhD, Austrian Institute of Technology
Paula Maria Lima Castro, Professor, Faculty of Biotechnology of the Portuguese Catholic University
Todas as correções determinadas pelo júri, e só essas, foram efetuadas. O Presidente do Júri,
Porto, ______/______/_________
“You live in intimate association with bacteria, and you couldn't survive without them”

Bonnie Bassler
Acknowledgements

I want to thank my family for the huge unconditional support they gave me. I want to thank Terezinha Dias, Sara Amorim, Ana Meireles, David Russo and Tiago Gundar for giving me the strength and advice for going forward with the decision to do an Erasmus Internship. I want to thank my friends for giving me what I needed not what I wanted. I want to thank Sebastian for all the hours of suffering, boredom, laughter and learning while working at the AIT. I want to thank AIT for giving me the opportunity of work. I want to thank Nikolic Branislav for helping me in times of questioning. I want to thank Birgit Mitter for giving me the push and the tools to go forward in my thesis. I want to thank Professor Susana Carvalho and Professor Paula Castro for giving me the opportunity, advice and knowledge to know more. I want to thank everybody who helped me reach this far in my evolution as a constant student of life.
Abstract

Plants live in close association with complex microbial assemblages that play a central role in plant nutrient uptake, pathogen protection, and stress tolerance. There is an increasing interest by the agricultural industry to integrate beneficial plant-microbe interactions in crop production. However, despite high expectations, there are still only few microbial products implemented in agricultural practice. One reason is that implemented inoculants show plant genotype dependent differences in the effect on plant vigor. Current microbial screening programs aim at identifying strains showing the desired plant-beneficial traits together with low plant genotype specificity, which involves enormous screening efforts. Alternatively, a better microbiome understanding could help to develop strategies for the successful application of microbial inoculants. One hypothesis in this context is that different plant genotypes host different microbial assemblages, which differently affects the establishment and activity of microbial inoculants. Thus, the objective of this study was to understand the effect of 18 tomato (Solanum lycopersicum) plant genotypes on the plant microbiome composition, and in second place, to understand the differences in microbiome composition of different fruit types (salad and cocktail tomatoes) and different plant compartments (root and rhizosphere). 16S rRNA gene fragment amplicon sequencing using Illumina MiSeq technology was employed to analyse the bacterial microbiome in rhizosphere and roots of field grown tomato plants. Data processing and biostatistics were performed to elucidate alpha- and beta-diversity as well as to phylogenetically describe and compare the bacterial assemblages. Additionally, current knowledge on found bacterial groups was briefly summarized. This study concluded, that the tomato genotype has no effect on the microbiome composition, showing no significant differences between genotypes, with an operational taxonomic units richness mean of 1700, a Simpson index mean of 0.93, and a Bray-Curtis Dissimilarly index mean between 0.34 and 0.44. However, significant differences were found between rhizosphere and roots microbiomes, with an operational taxonomic units richness mean of 210 on roots and 4300 on rhizosphere, a Simpson index mean of 0.74 on roots and 0.99 on rhizosphere, and a Bray-Curtis Dissimilarly index mean of 0.35 on roots and 0.18 on rhizosphere. More studies are needed, for the analysis of microbiota on species or even strain level to fully understand the complex interaction between soil, plant and microbial inoculum.

Keywords: Cocktail, endophyte, epiphyte, exudate, microbiome, microorganism, PGPB, rhizosphere, root, salad, Solanum lycopersicum.
Resumo

As plantas vivem em associações complexas com microrganismos desempenhando um papel central na absorção de nutrientes, proteção contra patogénicos e tolerância aos stresses. A indústria agrícola demonstra um interesse crescente em integrar interações benéficas entre plantas e microrganismos nas culturas hortícolas. No entanto, apesar de altas expectativas, ainda existem poucos produtos microbianos implementados na prática agrícola. Uma razão para tal é que inóculos demonstram diferenças dependentes do genótipo das plantas sobre o efeito no vigor da planta. Os programas atuais de triagem microbiana visam a identificação de estirpes com efeitos benéficos para a planta, juntamente com baixa especificidade genotípica da planta, envolvendo elevados custos económicos. Alternativamente, uma melhor compreensão do microbioma poderia ajudar a desenvolver estratégias para a aplicação bem-sucedida de inóculos microbianos. Uma hipótese neste contexto é que diferentes genótipos de plantas hospedam diferentes associações microbianas, que afetam de forma diferente o estabelecimento e atividade de inóculos microbianos. Assim, os objetivos deste estudo foram compreender o efeito de 18 genótipos de plantas de tomate (Solanum lycopersicum) na composição do microbioma da planta e, em segundo lugar, compreender as diferenças na composição do microbioma associado a diferentes tipos de frutas (tomates de salada e cocktail) e diferentes compartimentos da planta (raiz e rizosfera). Para analisar o microbioma bacteriano em raízes e rizosfera de plantas de tomate cultivadas em campo foi utilizada a sequenciação de fragmentos de gene 16S rRNA utilizando a tecnologia Illumina MiSeq. O processamento de dados e a bioestatística foram realizados para elucidar a diversidade alfa e beta, bem como para descrever e comparar as assemblagens bacterianas filogeneticamente. Adicionalmente, foi feito um breve resumo sobre os grupos bacterianos encontrados. Este estudo concluiu que os genótipos do tomate não têm efeito na composição do microbioma, demonstrando diferenças não significativas entre genótipos, com uma média de riqueza de unidades taxonómicas operacionais de 1700, uma média de índice de Simpson de 0.93. e uma média de índice de Bray-Curtis Dissimilarity entre 0.34 e 0.44. No entanto, foram encontradas diferenças significativas entre microbiomas das raízes e rizosfera, com uma média de riqueza de unidades taxonómicas operacionais de 210 em raízes e 4300 em rizosfera, uma média de índice de Simpson de 0.74 em raízes e 0.99 em rizosfera, e uma média de índice de Bray-Curtis Dissimilarity de 0.35 em raízes e 0.18 em rizosfera. Mais estudos ao nível de espécie ou ao nível de estirpe microbiana são necessários para se entender melhor a interação complexa entre solo, planta e inóculo microbiano.

Palavras-chave: Cocktail, endófita, epífita, exsudato, microbioma, microrganismo, PGPB, raiz, rizosfera, salada, Solanum lycopersicum.
Table of Contents

Acknowledgements ........................................................................................................ iv
Abstract ......................................................................................................................... v
Resumo ........................................................................................................................... vi
List of abbreviations ...................................................................................................... ix
List of figures .................................................................................................................. xi
List of tables ................................................................................................................... xiii
1 Introduction ................................................................................................................ 1
  1.1 Plant microbiome ................................................................................................... 1
  1.2 A triad interaction – rhizosphere-root-microbe .................................................... 3
    1.2.1 Rhizosphere .................................................................................................. 3
    1.2.2 Root ............................................................................................................. 4
    1.2.3 Microbe ...................................................................................................... 6
  1.3 Tomato and tomato breeding industry ................................................................. 8
  1.4 Objective .............................................................................................................. 11
2 Material and methods ............................................................................................... 12
  2.1 Plant material ...................................................................................................... 12
  2.2 Location and edaphoclimatic conditions .............................................................. 14
  2.3 Experimental layout, growth conditions and crop treatment ............................. 14
  2.4 Sampling ............................................................................................................. 15
  2.5 Laboratory procedure ........................................................................................ 15
    2.5.1 DNA isolation ............................................................................................ 15
    2.5.2 16S rRNA gene fragment amplification for sequencing .............................. 16
    2.5.3 Electrophoresis .......................................................................................... 17
    2.5.4 Dilutions .................................................................................................... 18
  2.6 Bioinformatics and data processing .................................................................... 18
  2.7 Statistical analysis ............................................................................................... 19
  2.8 Sequence analysis ............................................................................................... 20
3 Results ....................................................................................................................... 23
  3.1 OTU count overview .......................................................................................... 23
  3.2 Alpha-diversity analysis ..................................................................................... 24
  3.3 Beta-diversity analysis ....................................................................................... 26
  3.4 Taxonomy ........................................................................................................... 29
    3.4.1 By plant compartment .............................................................................. 29
    3.4.2 By genotype ............................................................................................... 31
  3.5 Core and unique OTU’s ...................................................................................... 33
3.5.1 Core OTU’s .................................................................33
3.5.2 Unique OTU’s ...............................................................34
4 Discussion ................................................................................37
  4.1 OTU’s count overview ..........................................................37
  4.2 Alpha-diversity analysis .......................................................37
  4.3 Beta-diversity analysis .........................................................38
  4.4 Taxonomy ...........................................................................39
  4.5 Core and unique OTU’s .......................................................40
  4.6 Sequencing data quality ......................................................40
5 Conclusions ............................................................................42
6 References ...............................................................................43
Attachment A | Modified DNA Isolation manual from Van der Beek .................56
Attachment B | Taxonomy summary of core and unique OTU’s .........................58
Attachment C | Species Description ..................................................59
List of abbreviations

2,4-DAPG – 2,4-diacylphloroglucinol
ACC – 1-aminocyclopropane-1-carboxylate
AHLs – N-acyl-L-homoserine lactones
ATP – Adenosine triphosphate
bp – Base pair
Ca – Calcium
CaCl₂ – Calcium chloride
CAL – Calcium-acetate-lactate
CAP – Constrained analysis of principle coordinates
CO₂ – Carbon dioxide
CTAB – Cetrimonium bromide
Cu – Copper
ddH₂O – Double distilled water
DMSO – Dimethyl sulfoxide
DNA – Deoxyribonucleic acid
dNTP’s – Deoxynucleotide triphosphates
e.g. – Exempli gratia (for example)
EDTA – Ethylenediaminetetraacetic acid
et al – And others
F – Forward
FDR – False discovery rate
Fe – Iron
fw2.2 – Loci fruit weight 2.2
ha – Hectare
i.e. – id est (that is)
IAA – Indole-3-acetic acid
K – Potassium
Mg – Magnesium
Mn – Manganese
N – Nitrogen
Na – Sodium
NaCl – Sodium chloride
NH₃ – Ammonia
OTU – Operational taxonomic unit
P – Phosphorus
P (p-value) – Probability value
PCR – Polymerase chain reaction
PE – Paired-end
PERMANOVA – Permutational multivariate analysis of variance
PGP – Plant-growth-promoting
PGPB – Plant growth promoting bacteria
pH – Potential of hydrogen
PVP – Polyvinylpyrrolidone
QTLs – Quantitative trait loci
R – Reverse
rcf – Relative centrifugal force
rDNA – Ribosomal deoxyribonucleic acid
rpm – Rotations per minute
rRNA – Ribosomal ribonucleic acid
S – Sedimentation rate
STE – Sodium Tris-EDTA
TBE – Tris borate ethylenediaminetetraacetic acid
TrisHCl – Tris hydrochloride
Zn – Zinc
List of figures

Figure 1 – Adapted from the plant microbiome model from Marnie, 2014. Belowground habitats include microbes living in areas surrounding the plant roots at the rhizosphere and rhizoplane. Microbes that live adherent to the root surface, epiphyte, and microbes living inside the plant tissue, endophytes. Bulk soil microbes are under influence of environmental factors without influence of plant exudates. The microbe action can be beneficial, prejudicial or neutral to the plant................................................................. 2

Figure 2 – Root anatomy with aspects from belowground habitats like the rhizosphere, rhizoplane and bacterial mucilage taken from Maier, Pepper, and Gerba, 2017. .................. 5

Figure 3 – Tomato genome map showing several QTL’s at the 12 chromosomes of the genome (Barrantes et al., 2016). .................................................................................. 10

Figure 4 – Comparison between the wild type Solanum pimpinellifolium L. and the cultivated tomato Solanum lycopersicum L (Pratta, Rodríguez, Zorzoli, Picardi, & Valle, 2011). ..........11

Figure 5 – Tomato leaf mould caused by Cladosporium fulvum (MoyaTeplica.ru, 2016). ......11

Figure 6 – Cocktail tomato, in this case Black cherry tomato (W. Atlee Burpee & Co., 2017). ........................................................................................................ 12

Figure 7 – Salad tomato, in this case Auriga tomato (Culinaris - Saatgut für Lebensmittel, 2017). .......................................................................................................... 12

Figure 8 – Daily temperature and relative humidity during the experiment. The data was recorded by using An EBI 20-TH Data Logger (ebro Electronic GmbH & Co. KG). ..........14

Figure 9 – Amplification scheme done by PCR. On denaturation (1), high temperature is applied to separate the strands from each other. On annealing (2) and elongation (3), primers complement to the target sequence and DNA polymerase synthesizes, from the end of the primer, new strands of DNA complementary to the target sequence. The process gets repeated exponentially (Microbe Online, 2017)................................................................. 17

Figure 10 – Index PCR adds a unique synthetic oligonucleotide index sequence to the end of DNA strands in each sample, by which samples can be distinguished after sequencing (GenDx, 2017)...................................................................................................... 17

Figure 11 – NCBI page showing the procedure and options to choose..................................21

Figure 12 – Resulting graphic with blast hits producing sequence significant alignments. ....21

Figure 13 – Sequences producing significant alignments show empty parts with query cover below 95%, without all the base pairs......................................................... 22

Figure 14 – Venn diagram showing the number of OTU’s present only in roots (green), in both regions (dark blue) and only in rhizosphere (blue). ..................................................24

Figure 15 – Venn diagram showing the number of OTU’s present only in Cocktail tomatoes (green), in both regions (dark blue) and only in Salad tomatoes (blue).........................24
Figure 16 – Alpha diversity within roots and rhizospheres, from the 18 tomato cultivars tested, measured by observed OTU richness and Simpson index of diversity. A permutation pairwise comparison (RVAideMemoire R package) showed that the richness and the diversity values were significantly different ($p = 0.0002$, perm = 9999).

Figure 17 – Alpha diversity of the 18 cultivars tested in terms of observed OTU richness and Simpson index of diversity. A permutation pairwise comparison showed ($p = 0.97$, perm = 9999) that the richness and the diversity values were not significant.

Figure 18 – Bray-curtis beta diversity among genotype samples. A permutation test assessed the significance of the treatment on a Constrained Analysis of Principal Coordinates (CAP) ($p > 0.415$, perm = 999) (vegan R package).

Figure 19 – Average Bray-curtis distance to centroid as an indication of beta diversity in the roots from the 18 genotypes tested. A permutation test assessed the significance of the treatment on a Constrained Analysis of Principal Coordinates (CAP) ($p > 0.389$, perm = 999).

Figure 20 – Beta diversity, as indicated by the distance to centroid (Bray Curtis index), in the rhizosphere from the 18 tomato cultivars tested. A permutation test assessed the significance of the treatment on a Constrained Analysis of Principal Coordinates (CAP) ($p > 0.339$, perm = 999).

Figure 21 – Beta diversity, as indicated by the distance to centroid (Bray Curtis index), among plant compartment. A permutation test assessed the significance of the treatment on a Constrained Analysis of Principal Coordinates (CAP) ($p < 0.001$, perm = 999).

Figure 22 – Beta diversity, as indicated by the distance to centroid (Bray Curtis index), among fruit type samples. A permutation test assessed the significance of the treatment on a Constrained Analysis of Principal Coordinates (CAP) ($p = 1$, perm = 999).

Figure 23 – Relative OTU abundance at (A) Phylum, (B) Class and (C) Order level in roots (left) and rhizosphere (right).

Figure 24 – Relative OTU abundance at (A) Phylum, (B) Class and (C) Order level in all eighteen cultivars.
List of tables

Table 1 – Characterization of the 18 tomato cultivars tested in terms of variety name, year of release, breeding background, fruit weight, fruit type and suitable production system. ........13
Table 2 – Overview of the number of Operational Taxonomic Units (OTU’s) in total, only in root, only in rhizosphere, in both regions, unique to the specific cultivar and common to all the 18 cultivars tested.................................................................23
Table 3 – Number of core and unique OTU’s with chimeras and without chimeras............33
Table 4 – Taxonomic classification and number of OTU’s included in each classification.....34
1 Introduction

The plant-associated microbes, the plant microbiota, can help plants to fend off diseases, stimulate growth, occupy space that would otherwise be taken up by pathogens, promote stress resistance and additionally influence crop yield and quality (Berg, Zachow, Müller, Philipps, & Tilcher, 2013). There is increasing interest to integrate beneficial plant-microbe interactions in current crops. However, as plant-microbe interactions are highly specific and strongly driven by the host plant genotype and physiology, effects of a given microbial strain on different varieties of a crop species are often highly variable and unpredictable (Mitter, Pfaffenbichler, & Sessitsch, 2016).

A better understanding of plant-microbe interactions is needed, especially on the relationships between plant genotype and the composition of its naturally microbiota community, for the creation of best plant-microbe combinations for optimal plant production in face of current challenges, such as degradation of soil (Montgomery, 2007), depletion of water supplies (Rosegrant, Cai, & Cline, 2002), and losses of biodiversity (Pimentel, Lach, Zuniga, & Morrison, 2000). Human demands are beyond food production. Agriculture and human urbanization compete for land usage. Balancing the two will inevitably require harnessing resources from fragile ecosystems or marginal landscapes (Tilman et al., 2009) while simultaneously maximizing plant productivity. The plant microbiome can play a prominent role towards achieving this goal (Marnie, 2014).

1.1 Plant microbiome

Microorganisms are essential to life, macroorganisms like plants and animals cannot live without them. In 1994, Richard Jefferson suggested that the evolutionary selection unit is not a single organism but a macroorganism (e.g. a plant or animal) and all its associated microorganisms. According to this theory, the entire suite of genes, the genome of the host and the genomes of all its microbes (i.e. the microbiome), determine the vigor of an organism.

The importance of the plant microbiome for plant health and development is well acknowledged (Bulgarelli, Schlaeppi, Spaepen, van Themaat, & Schulze-Lefert, 2013; Hardoim et al., 2015). It comprises the set of genes present in a given habitat associated with a given plant, comprising genomes vastly more complex than that of the plant alone, and by the nature of microbial interactions, these genomes serve as an extension of the plant genetics (Berendsen, Pieterse, & Bakker, 2012). In this case, the habitat refers to the whole plant, regions of the organism (roots, leaves, seeds, etc.) and zones of interaction between the organism and the surrounding environment. One of those habitats is the rhizosphere, a
zone of high microbial activity created by the roots through their exudates and by the exudates of the microorganisms within the soil matrix, being not merely the soil in contact with the plant (Marnie, 2014).

Microbes interact in an adherent (epiphytic) or internal (endophytic) manner in the belowground habitats within the plant microbiome, bulk soil, rhizosphere and rhizoplane (Figure 1). Bulk soil is an area where the roots do not penetrate. It’s a region out of reach for plant influence by root exudation, i.e. higher concentrations of organic compounds in the rhizospheres compared with bulk soils are consistently reported (Jones, 1998). The rhizosphere is a key habitat, containing a vast microbial diversity, (Egamberdieva et al., 2007; Mendes et al., 2011) and the soil in the rhizosphere acts as a medium for complex signalling among microbes and plants done by exudates, creating in this way a zone of interaction between the roots and the near soil.

The rhizosphere is influenced by climatic factors, which in turn impact the plants and the microbiota that utilize this habitat (Bais, Park, Weir, Callaway, & Vivanco, 2004). Closer to the plant is the rhizoplane habitat, which is the surface of the plant tissues directly in contact with the soil (roots and rhizomes). Microorganisms have different lifestyles, not restricted to a specific habitat, existing as facultative epiphytes and endophytes influenced by biotic and abiotic factors (Marnie, 2014).

![Image](image_url)

Figure 1 – Adapted from the plant microbiome model from Marnie, 2014. Belowground habitats include microbes living in areas surrounding the plant roots at the rhizosphere and rhizoplane. Microbes that live adherent to the root surface, epiphyte, and microbes living inside the plant tissue, endophytes. Bulk soil microbes are under influence of environmental factors without influence of plant exudates. The microbe action can be beneficial, prejudicial or neutral to the plant.

Because of this interplay between epiphytic-endophytic lifestyles and the plant microbiome, the rhizosphere is a dynamic environment. Microbiome structure is both influenced by and has an influence on the rhizosphere, contributing to some of the major differences between
rhizosphere and bulk soil (Neumann & Römheld, 2002). The pH of the rhizosphere can differ up to 2–3 units from bulk soil as a direct result of biological activity, which can likewise impact the relative solubility of essential nutrients (Marnie, 2014) such as phosphorus, which occurs in soils abundant in insoluble inorganic forms that can be solubilized through the actions of plants and microbes (Neumann & Römheld, 2002).

Rhizosphere biota varies and adapts (Rout & Callaway, 2009, 2012). The plant microbiome influences ecosystem functions, increasing the available forms of carbon and nitrogen to the plants. Microbes mediate the ecosystem functioning by soil formation, decomposition of organic matter, nutrient mineralization, and primary productivity (de Bello et al., 2010). The microbiome of the rhizosphere has an impact on plant productivity, e.g. nodulating soybean (*Glycine max* L.) cultivars manipulated to enhance yield through alterations to their interactions with *Rhizobium* microbial partners (Harris, Pacovsky, & Paul, 1985; Kiers & Denison, 2008; Heath & Tiffin, 2009). Additionally, PGP (plant growth promoting) activities can support plant nutrition and growth by nitrogen (N$_2$) fixation (James, 2000; Martinez-Romero, 2006), phosphate solubilisation, and production of plant growth hormones (Hardoim, van Overbeek, & Elsas, 2008). Plants are no longer considered standalone entities, it is well established that they live in a symbiotic way with microbes in order to survive.

### 1.2 A triad interaction – rhizosphere-root-microbe

In order to address rhizosphere-root-microbe interactions, the relationships between belowground and aboveground events were ignored. This includes protection from herbivory, pollination, seed predation and pathogen attack from aboveground structures as recommend in literature (Friesen et al., 2011).

#### 1.2.1 Rhizosphere

As previously stated, the rhizosphere refers to the zone of influence created by the roots through their exudates and by the exudates of the microorganisms within the soil matrix, and as such it is not merely the soil in contact with the roots (Marnie, 2014).

The rhizosphere microbial community is in constant change and represents the main source of microorganisms, which migrate into plants via the root and establish sub-communities inside plants (Hardoim et al., 2015). Differences between rhizosphere and bulk soils are mainly due to abiotic and biotic stresses. With drought, pathogenic attack, and metal toxicities the differences between zones are apparent and important to understand. In arid environments, for example, low microbial abundance and diversity are expected, but the contrary happens more often (Ben-David, Zaady, Sher, & Nejidat, 2011; Aguirre-Garrido et al., 2012). Rhizosphere soils have higher water-holding capacity, increased nutrient availability, and greater microbial biomass compared to bulk soils (Schade & Hobbie, 2005).
The rhizosphere region of the plant microbiome is informative about plant health (Marnie, 2014). A healthy status can be reflected in the composition of the rhizosphere microbiome (Burdon & Thrall, 2009).

Exudates in the rhizosphere can be defined as chemicals released by plants or microorganisms. They allow communication between them, and assist in environmental adaptation to alleviate stresses, such as pathogen attack, drought-limiting nutrient acquisition and metal toxicity. Through exudates, the plant microbiome serves as an extension of the plant genome (Marnie, 2014).

Exudates can vary strongly and are often species-specific to the plant and/or the microorganism in action. They include high-molecular-weight molecules (sugars, lipids, and proteins) and low-molecular-weight signalling molecules (organic molecules and amino acids) (Badri et al., 2009). Many exudates serve as an energy source for the microbiota; prokaryotes can utilize many different sources of carbon, including many of the plant exudates (Marnie, 2014). Sugars and amino acids are in the majority of plant root exudates (Jaeger, Lindow, Miller, Clark, & Firestone, 1999), serving a variety of functions as antimicrobials, allelopathic molecules (i.e. that influence another organism) and pathogen/herbivore defence agents (Marnie, 2014). *Sorghum halepense*, an invasive grass, has a dominant chemical exudate (sorgoleone) from root hairs (Czarnota, Paul, Dayan, Nimbal, & Weston, 2001; Dayan, Kagan, & Rimando, 2003) with allelopathic properties (Czarnota, Rimando, & Weston, 2003). Studies of the feedbacks between the influence of the microbiome and plant production of exudates are important, yet very rare (Marnie, 2014).

### 1.2.2 Root

The root is the “gut” of a plant. It provides absorption of water and inorganic nutrients, support and anchoring of plant structure to the ground, a region for storage of food and nutrients, vegetative reproduction and competition with other plants. Roots function to uptake nutrients and chemical signalling molecules from the rhizosphere while simultaneously depositing nutrients and chemical signalling molecules into the rhizosphere (Marnie, 2014).

These exudates serve as a chemical currency to the plant, providing services to the plant microbiome in form of nutrient substrate, protection and much more. Released chemical components are usually unique to a genus or family (Lesuffleur, Paynel, Bataillé, Le Deunff, & Cliquet, 2007) ranging in concentration and composition due to several factors (Carvalhais et al., 2011). A schematic overview of the root and rhizosphere is shown in Figure 2.
Terpenoids, flavonoids, and isoflavonoids comprise many of the plant’s antimicrobial defence molecules (Hardoim et al., 2008). Terpenoids, as primary metabolites, regulate cellular processes and their derivatives are essential for photosynthesis (photosynthetic pigments), and seed emergence (gibberellic and abscisic acids). As secondary metabolites, they provide a range of benefits to plants in the form of defences against pathogens and niche expansion through allelopathic molecules (Marnie, 2014).

Roots are continuously attacked by pathogens with the potential for rapid and selective responses to plant defence mechanisms, therefore roots must continuously produce and secrete defence compounds into the rhizosphere (Bais, Loyola-Vargas, Flores, & Vivanco, 2001; Flores, Vivanco, & Loyola-Vargas, 1999). One mechanism to cope with pathogens is the production of border cells, which act as a buffer for pathogen attack (Hawes, Gunawardena, Miyasakab, & Zhaoa, 2000).

Elevated levels of CO$_2$, drought, and nutrient deprivation (mainly nitrogen and phosphorus) influence root exudate composition and quantity. For maize plants grown under either nitrogen (N) or phosphorus (P) limitations, decreased quantities of exuded amino and organic acids from roots were found (Carvalhais et al., 2011). These factors can have profound impacts on the phytochemistry released as exudates (Marnie, 2014). Experiments with CO$_2$-fertilization showed that increased carbon allocation to roots resulted in shifts in exudate composition and concentration that varied according to plant species (Cheng & Gershenson, 2007). Not all environmental stressors trigger the same responses from plant roots and nutrient deprivation have an inhibitory effect on root exudation (Marnie, 2014). The microbiome can ensure plant protection through exudates in addition to plant-growth promotion (Marnie, 2014).
High-molecular-weight compounds require active transport mechanisms to cross the plasma membrane and gain release into the rhizosphere (Badri et al., 2009), as is the case of the ATP-binding cassette transporter, which is important in root exudation composition and concentration (Badri et al., 2009). In contrast, low-molecular-weight molecules can be released through membrane diffusion or through protein channels (Badri et al., 2009). Because plant exudation is influenced by many factors, estimating the costs to the plant becomes impossible and irrelevant outside the context of a given system.

Plants are often the beneficiaries of antimicrobial exudates provided by microbial assemblages from the rhizosphere. Research indicates that plants can actively construct the rhizosphere microbiome and that this community is, to some degree, regulated or recruited by the plant to serve protection against pathogens (Friesen et al., 2011). One example is the production of enzymes that degrade N-acyl-L-homoserine lactones (AHLs). Such compounds are commonly found among several prokaryotic genera, including plant pathogens and their role is the regulation of gene expression in response to fluctuations in cell-population density (quorum sensing). Research showed plants capability of recruiting beneficial bacteria expressing high levels of AHL-degrading enzymes when exposed to a pathogen, suppressing virulence gene expression in this way (Teplitski, Robinson, & Bauer, 2000; Reading & Sperandio, 2006).

Some fluorescent pseudomonads are capable of producing antimicrobials 2,4-diacetylphloroglucinol (2,4-DAPG) (Marnie, 2014). These bacteria are common to the rhizosphere of many plant species (Mavrodi, Mavrodi, Parejko, Weller, & Thomashow, 2011) and their antimicrobial metabolites provide a huge spectrum of protection against a wide range of plant pathogens (Raaijmakers, Paulitz, Steinberg, Alabouvette, & Moënne-Loccoz, 2009). The 2,4-DAPG build-up in the rhizosphere showed correlation with the suppression of the disease known as "take-all" in wheat (Mendes et al., 2011). Research also suggests that 2,4-DAPG reduces mineral content in the rhizosphere, perform plant regulatory and signalling functions (altering exudate profiles), and play a role in the induction of plant systemic resistance (Doornbos et al., 2012).

Plant phenotypic expression can also be the result of the interplay with the microbiome. Plant functional processes can be mediated by microorganisms by providing biochemical capabilities and through alteration of plant metabolic pathways (Marnie, 2014). Microorganisms are able to synthesize biologically active chemicals that can mimic those produced by plants (e.g. hormones), or are totally novel to plants (e.g. specific antimicrobials).

1.2.3 Microbe

For microbes, root exudates are the primary means of survival within the soil environment. Preference for organic and amino acids released by plant roots is reflected in chemotaxis
ability (movement of an organism in response to a chemical stimulus) in many rhizosphere microbes (Nelson, 2004).

Root exudates induce gene expression in rhizosphere bacteria, by means of, chemotaxis and stress tolerance (Amador, Canosa, Govantes, & Santero, 2010), polychlorinated biphenyl degradation (Toussaint, Pham, Barriault, & Sylvestre, 2012), modulation of genes involved in competence and sporulation (Mäder et al., 2002), and biofilm formation on plant roots (Rudrappa, Czymmek, Parè, & Bais, 2008). Amino acids released by roots are involved in biofilm formation and disassembly of rhizosphere bacteria (Kolodkin-Gal et al., 2010). Adherence of microbes into the surface and internal structures of plants are prevented by exudates like terpenoids, flavonoids, and isoflavonoids (Hardoim et al., 2008). Sugars, organic acids and amino acids are abundant in the rhizosphere (Jones, 1998), making this part of the plant microbiome a hot-spot in microbial diversity (Berendsen et al., 2012).

Microbial respiration increases with availability of small molecules, especially organic acids (Hees et al., 2005). This suggests that the microbial community changes when varying concentration levels of readily available nutrients that require minimal energy to assimilate. Amino acids are obtained more easily by the rhizosphere microorganisms than by the plants (Kielland, 1994; Owen & Jones, 2001). It was shown for Pseudomonas putida that in order to maximize the utilization of amino acids, sugar metabolism must be deviated (Moreno, Martínez-Gomariz, Yuste, Gil, & Rojo, 2009). This PGPB utilizes the same protein for amino acid uptake as it does for downregulating gene transcription necessary for sugar uptake (Moreno et al., 2009).

Biogeochemical cycling of soil nutrients is the outcome of microbial uptake and exudation in the rhizosphere (Marnie, 2014). Plant exudates can be catalysts for the chemical transformations in soils, related to plant nutrients (N, P, Zn, Fe, Cu and Mn) and alkaline metals (Ca, Mg, K, and Na) (Stevenson & Cole, 1999). Nitrogen transformation into oxidation states is essential for higher organisms to survive. Rhizosphere microbiota is responsible for this transformation. Conversion of nitrogen gas (N\textsubscript{2}) into ammonia (NH\textsubscript{3}) is made possible by N\textsubscript{2}-fixing prokaryotes, thus making nitrogen bioavailable to primary producers. All plant hormones currently known can be produced by microorganisms (Friesen et al., 2011). For example, research on IAA hormone (auxin) production showed that 80% of the IAA was from bacterial taxa in plant rhizospheres (Loper, 1986). Low concentrations of IAA promote root growth in many plants (Glick, 1995), whereas high concentrations of IAA inhibit plant growth and cause developmental perturbations typical of the effect of plant pathogenic bacteria (Sarwar & Kremer, 1995). PGPB that influence plant traits by increasing tolerance to abiotic stress and plant performance are favourably selected by the plant (Friesen, 2013). PGPB can also have a direct impact on plant traits by modulating gene expression (Alfano et al., 2007; Da, Nowak, & Flinn, 2012). IAA hormone and 1-aminocyclopropane-1-carboxylate
(ACC) deaminase, an enzyme involved in regulating plant hormone signalling, can be imitated by microbes. Research showed that inoculation of wheat with rhizosphere bacteria expressing ACC deaminase, increased root development and, consequently, nutrient uptake (Shaharoona, Naveed, Arshad, & Zahir, 2008).

Microbe exudation of low-molecular-weight effector molecules triggers the immune response in plants through pathogen-associated molecular patterns (Boller & He, 2009). Therefore microbes in the rhizosphere can act on other microbes by suppressive mechanisms or act directly on the plant to stimulate immune responses or to trigger plant exudation (Marnie, 2014). With further understanding of microbe interactions with the surrounding soil and plant microbiota, improvements in quality, productivity and crop sustainability will be achieved.

### 1.3 Tomato and tomato breeding industry

Tomato (*Solanum lycopersicum*) originates from the Andes region (Chile, Bolivia, Ecuador, Colombia and Peru) and its domestication is estimated to have begun in 200 B.C, in the region where nowadays Mexico is. Further domestication at an intense level was done after the discovery in the 15th century by the Europeans, in the 18th and 19th centuries (Sims, 1980). Since the 20th century, the single species *S. lycopersicum* has been extensively modified by plant breeding, resulting in numerous cultivars that present considerable morphological differences.

*S. lycopersicum* is one of more than 3000 species of the Solanaceae family, which was widespread in the “Old World” (eggplant in China and India) and the “New World” (pepper, potato and tomato in Central and South America) (Knapp, 2002). While *S. lycopersicum* is the only domesticated species, it has 12 wild relatives (I.E. Peralta, Knapp, & Spooner, 2006) which have a large genetic diversity, especially within the self-incompatible species like *S. chilense* and *S. peruvianum* (Rick, 1988). The genetic variability of the wild species has been investigated for specific traits and is being exploited in tomato breeding programs (Rick & Chetelat, 1995; Walter, 1967). In comparison with the wild species, the cultivated tomato is poor in genetic variability, with an estimated variation of less than 5% within the genomes (Miller & Tanksley, 1990).

In addition, most wild tomatoes are endemic in narrow geographical regions in the Andes, with very small populations, being vulnerable to extinction. Thus, the discovery and maintenance of new species in diverse regions is extremely important (Peralta, Knapp, & Spooner, 2005).

With the introduction of tomato in Europe, coming from America, tomato domestication experienced severe genetic bottlenecks, since selection was done on a single plant basis and with small numbers of selected plants. In a predominantly inbreeding species, genetic
variation tends to decrease, even without selection. Consequently, genetic drift occurs reducing genetic variation (Bai & Lindhout, 2007).

With domestication a wide range of morphological and physiological traits can be distinguish from those of their wild ancestors. These characteristics are collectively referred to as the domestication syndrome (Frary & Doganlar, 2003). Studies on the domestication syndrome and domestication process revealed that numerous traits that distinguish cultivars from their wild relatives are often controlled genetically by a relatively small number of loci with effects of unequal magnitude (Frary & Doganlar, 2003). Genes and quantitative trait loci (QTLs) found to underlie domestication syndrome are traits for growth (self-pruning, plant height and precocity) and fruit development (set, size, shape, colour and morphology) (Grandillo & Tanksley, 1996; Doganlar, Frary, & Tanksley, 2000; Frary & Doganlar, 2003; Tanksley, 2004).

Fruit size, for example, was heavily affected, given the massive increase that occurred. While wild tomatoes have small fruits with the purpose of propagation of the species, the cultivars were selected to feed human beings.

The crossing of a wild and a cultivated tomato, showed that mutations in about six QTLs might be responsible for transforming the small berries of wild tomatoes to the large fruit from the modern cultivars (Tanksley, 2004). For example, the locus “fruit weight 2.2” (fw2.2) is responsible for changes in the fruit weight by up to 30% (Bai & Lindhout, 2007) (Figure 3). It was suggested that mutation(s) in the fw2.2 locus was the first step to domestication and responsible for a key transition during that process (Alpert, Grandillo, & Tanksley, 1995; Frary et al., 2000).

With commercial development, breeding shifted from open pollinated cultivars to hybrids combining good traits from both parents that segregate in the progeny (Bai & Lindhout, 2007). The art of tomato breeding identifies and combines the specific traits for each market. In the past few decades, breeding goals have gone through four phases: breeding for yield in the 1970s, for shelf-life in the 1980s, for taste in the 1990s and for nutritional quality currently (Bai & Lindhout, 2007). For the tomato breeding program to be successful, certain principles must be taken into account. Growers must produce a high yield of high-quality fruit, while holding production costs as low as possible. Therefore, many of the breeding goals are based on characteristics that reduce production costs or ensure reliable production of high yields with high-quality fruits (Bai & Lindhout, 2007).

Tomato production is sub-optimal over large parts of the tomato crop-growing areas, due to unfavourable environmental conditions caused by abiotic factors including high or low temperatures, excessive water or drought, soil salinity and alkalinity (Bai & Lindhout, 2007). In tomato breeding, germplasm offers high genetic variation for stress tolerance (C.M. Rick &

Figure 3 – Tomato genome map showing several QTL’s at the 12 chromosomes of the genome (Barrantes et al., 2016).

One of the most prominent issues in tomato breeding is improving resistance to the most destructive pests and pathogens. Tomato hosts about 200 species of a wide variety of pests and pathogens that can cause significant economic losses (Bai & Lindhout, 2007). Usually, the use of synthetic agrochemicals for the control of pests and diseases are effective, but this causes several problems such as raising of production costs and potential risk to growers, consumers and the environment. Nature has provided a great wealth of mechanisms of resistances that are available in the wild species. Many of those mechanisms are simply inherited, but also remarkable successes have been made in transferring disease-resistance genes into cultivated tomato (Bai & Lindhout, 2007), as for example, genes from *S. pimpinellifolium* (Figure 4) which demonstrated to increase resistance to fungal pathogens such as *Cladosporium fulvum* (Figure 5) (Walter, 1967).
Faculty of Sciences of the University of Porto
Effect of tomato genotype on microbiome composition

Figure 4 – Comparison between the wild type *Solanum pimpinellifolium* L. and the cultivated tomato *Solanum lycopersicum* L. (Pratta, Rodríguez, Zorzoli, Picardi, & Valle, 2011).

Fruit quality is dependent of human perception, a combination of visual stimuli like size, shape and colour, and sensory factors like sugar content, acidity and aroma. This perception is also heavily influenced by the origin of the product (e.g. biological product vs transgenic product) (Bai & Lindhout, 2007). Research on the genetic control of fruit quality traits has been dominated by studies of the ripening process and determination of soluble solid content (C.M. Rick & Chetelat, 1995).

Figure 5 – Tomato leaf mould caused by *Cladosporium fulvum* (MoyaTeplica.ru, 2016).

Much of the genetic variation was due to spontaneously occurring mutations that were rapidly introduced into new cultivars if these provided benefit. Sooner or later, a limit of the potential of tomato breeding by only using the cultivated germplasm will be reached (Bai & Lindhout, 2007). Therefore, the integration of alternative resources may have to be used in future, such as microorganisms with desired traits to improve tomato production.

**1.4 Objective**

Because of the genetic characteristics of the tomato plant, one hypothesis for this study is that different plant genotypes host different microbial assemblages, which differently affects the establishment and activity of microorganisms. As such, as a first step, this thesis was set up to investigate, the impact of the plant genotype on the microbiome in rhizosphere and roots of tomato plants, and secondarily, differences in the microbiome composition in different fruit types (salad and cocktail tomatoes) and different plant compartments (root and rhizosphere) were analysed.
2 Material and methods

2.1 Plant material

Eighteen different tomato cultivars (Table 1) were tested in this study, including cocktail and salad tomatoes. Cocktail tomatoes (Figure 6) are round, small (e.g. 2.5 cm in diameter) and they grow in large clusters. In turn, salad tomatoes (Figure 7) are larger (5 to 7.5 cm in diameter) and more succulent, being used in sandwiches, salads and tomato sauce (Masley, 2017).

Figure 6 – Cocktail tomato, in this case Black cherry tomato (W. Atlee Burpee & Co., 2017).

Figure 7 – Salad tomato, in this case Auriga tomato (Culinaris - Saatgut für Lebensmittel, 2017).
The cultivars tested have either conventional or organic breeding background. The fruit types analysed were derived from separate breeding programs because consumer expectations on fruit quality differ with fruit type.

Table 1 – Characterization of the 18 tomato cultivars tested in terms of variety name, year of release, breeding background, fruit weight, fruit type and suitable production system.

<table>
<thead>
<tr>
<th>Cultivar name</th>
<th>Year of release</th>
<th>Breeding background</th>
<th>Average fruit weight (g)</th>
<th>Fruit type</th>
<th>Production systems</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annamay F1</td>
<td>2010</td>
<td>Conventional</td>
<td>46.0</td>
<td>Cocktail</td>
<td>I</td>
<td>European Commission (2015)</td>
</tr>
<tr>
<td>Black Cherry</td>
<td>2009</td>
<td>Uncertain</td>
<td>23.0</td>
<td>Cocktail</td>
<td>I</td>
<td>European Commission (2015)</td>
</tr>
<tr>
<td>Goldita</td>
<td>1997</td>
<td>Conventional</td>
<td>15.6</td>
<td>Cocktail</td>
<td>I</td>
<td>Arche Noah (2016)</td>
</tr>
<tr>
<td>Harzfeuer F1</td>
<td>1959</td>
<td>Conventional</td>
<td>76.4</td>
<td>Salad</td>
<td>O, I</td>
<td>Bundessortenan (2015)</td>
</tr>
<tr>
<td>Resi</td>
<td>2010</td>
<td>Uncertain</td>
<td>17.3</td>
<td>Cocktail</td>
<td>O</td>
<td>European Commission (2015)</td>
</tr>
<tr>
<td>Sakura F1</td>
<td>1999</td>
<td>Conventional</td>
<td>23.7</td>
<td>Cocktail</td>
<td>I</td>
<td>European Commission (2015)</td>
</tr>
</tbody>
</table>

1 Uncertain: traditional cultivars, genetic resources, and seed-savers exchange
2 Calculated from the experiment data
3 Suitable production system of the cultivars:
   - O = outdoor, I = indoor, H = hydroponic, * = hardly grown anymore (old cultivars)
2.2 Location and edaphoclimatic conditions

The experiment was carried out at Reinshof Experimental Station of the University of Göttingen, Germany, (51°30'17.6"N, 9°55'16.2"E) at 152 m elevation on a natural organic field of alluvial soil, from May to October 2016. Faba bean (Vicia faba L.) was cultivated as preceding crop. The soil pH was 7, humus content 1.89%, magnesium 40 mg g\(^{-1}\) (CaCl\(_2\) extraction), phosphorus 70 mg g\(^{-1}\) (calcium-acetate-lactate (CAL) extraction), and potassium 90 mg g\(^{-1}\) (CAL extraction) (Janssen, 2016). At the beginning of the field season (June 2\(^{nd}\)), at the mid-season (September 1\(^{st}\)), and at the end of the field season (October 12\(^{th}\)), the plant available nitrogen in the top soil (0-30 cm) was 111.5 kg ha\(^{-1}\), 23.5 kg ha\(^{-1}\), and 7.0 kg ha\(^{-1}\) and in the sub-soil (30-60 cm) 54.3 kg ha\(^{-1}\), 88.6 kg ha\(^{-1}\), and 5.2 kg ha\(^{-1}\) respectively. Average temperature and relative humidity during the experiment were 19.00±4.60 °C and 75.00±10.70 % with average diurnal deviation of 5.43 °C and 14.16 % correspondingly (Figure 8).

![Figure 8](image.png)

Figure 8 – Daily temperature and relative humidity during the experiment. The data was recorded by using An EBI 20-TH Data Logger (ebro Electronic GmbH & Co. KG).

2.3 Experimental layout, growth conditions and crop treatment

The tomato plants were grown according to organic standards. The experiment was designed as randomized complete block with 20 indeterminate tomato cultivars (S. lycopersicum) with eight replications and one plant per plot. From the whole 20 cultivars the eighteen tested in this study were chosen later. The two cultivars where discarded because of bias results. On March 2016, tomato seeds were germinated in polyethylene plastic nursery multi-pot trays QP 96 (Hermann Meyer KG, Germany) filled with organic substrate (topora® Bio, coconut fibre, pine and green compost) (Kleeschulte GmbH & Co. KG,
Twenty days later, each seedling was placed in a pot (11x11x11 cm) filled with organic substrate with peat-perlite mixture (HAWITA-Gruppe GmbH, Germany) and maintained in the greenhouse at a temperature of 20 °C and 18 °C in day and night time, respectively, until the first inflorescence emerged. On May 2016, all seedlings were transplanted to the field with a space of 50 cm within row and 100 cm between rows. The growing system was designed to avoid major pathogens, such as Phytophthora infestans and Cladosporium fulvum to allow the assessment of sensitive cultivars. In this sense, a well-ventilated rainout shelter with open sides was used. Border effects were minimized by excluding from the test three plants at the edges of each row. Between-row spaces were covered with dark plastic foil for weed control. Two weeks after transplanting, the tomato plants were trellised to overhead wires to vertically support the vine and continued fixing routinely. Pruning was performed weekly by pinching out side-shoots emerging from leaves, inflorescences, and internodes, so that only single leader stem was remained. Neither fertilizer nor pesticide were applied in this experiment. Drip irrigation was supplied weekly (two times per week during hot periods) through lateral polytube drip lines (John Deere GmbH & Co.KG, Germany) with 50 cm emitter spacing in moderate level based on assessment of temperature.

2.4 Sampling

Sampling was performed by removing the entire plant, after which they were gently shaken and soil adhering to the roots (rhizosphere) was recovered. The entire roots were gently washed with sterile water and aseptically cut in small pieces for proper conditioning and study. Rhizosphere and root samples were stored at -80 °C in cryogenic vials.

2.5 Laboratory procedure

2.5.1 DNA isolation

The Erasmus internship for the master thesis in discussion began with the laboratory procedure. Genomic DNA isolation was done from 18 different tomato cultivars; 3 replicates for rhizosphere DNA (54 samples) and 3 replicates for root DNA (54 samples), totaling 108 samples. Rhizosphere was isolated using the FastDNA™ SPIN Kit for soil with bead-beating and the FastPrep® Instrument (MP Biomedicals, Santa Ana, CA) according to the manufacturer’s protocol. bead-beating is used to avoid the preferential extraction of Gram-negative bacteria or bacteria with thinner cell walls (Feinstein, Sul and Blackwood 2009). For DNA isolation from roots, roots were washed, as described above, and crushed using a ball mill until a fine powder was obtained. DNA was extracted following an adapted protocol from Van der Beek (Van der Beek et al, 1992) (Attachment A). DNA (5 ml) was separated and
visualy tested for quality by electrophoresis. DNA concentration was measured using a Nanodrop spectrophotometer (Thermo Scientific).

2.5.2 16S rRNA gene fragment amplification for sequencing

2.5.2.1 First amplification
Amplicon preparation for 16S rRNA gene sequencing was done in a two-step process involving PCR amplifications (Figure 9). Initially, the primers 799-forward (5'-AACMGGATTAGATACCCKG-3') and 1392-reverse (5'-ACGGGCGGTGTGTRC-3') (Chelius & Triplett, 2001) were used. Amplification with the primer pair 799F and 1392R allows exclusion of the chloroplast 16S rDNA and, exclusion in co-amplification of bacterial and mitochondrial ribosomal gene fragments of about 600 bp and 1,000 bp amplicon size, respectively (Mitter et al., 2017). PCR amplification was performed in a peqSTAR thermocycler (peQlab, Erangen, Germany). Twenty five μl PCR reaction mixture contained 1 μl of template DNA, 0.25 μl of KAPA HiFi DNA Polymerase (Kapa Biosystems, Wilmington, MA, USA), 5 μl of each primer with a concentration of 1.5 μM, 075 μl of dNTP’s with a concentration of 10 mM, 5 μl of 5x KAPA HiFi fidelity buffer with Mg, 1.25 μl of DMSO (5%) and 6.75 μl of PCR grade water. The cycling parameters for PCR amplification were as follows: denaturation at 95 °C for 3 min (5 min for root samples); 30 cycles of denaturation at 98 °C for 20 s, annealing at 51 °C for 20 s (55 °C for 30 s for root samples) and extension at 72 °C for 30 s; final extension was at 72 °C for 1 min (3 min for root samples). PCR amplicons were separated by gel electrophoresis and the bands deriving from bacterial 16S rRNA genes (about 600 bp) were excised from the agarose gels. The gel pieces were put in a filter tip that was placed in a fresh tube and DNA was collected by centrifugation for 8 min at 6,000 rpm (Grey & Brendel, 1992).

2.5.2.2 Second amplification
The second amplification, intended to reduce non-specific binding due to the amplification of unexpected primer binding sites (nested PCR), was executed with the primers 799-forward illumina (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAACMGGATTAGATACCCKG-3') and 1175-reverse illumina (5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACACGCTCRCTCCDCCTTCTC-3') (Illumina Inc., San Diego, USA). The PCR reaction parameters were as described above, without the use of DMSO (5%). The cycling parameters for PCR amplification were as follows: denaturation at 95 °C for 5 min.; 30 cycles of denaturation at 98 °C for 20 s, annealing at 55 °C for 20 s and extension at 72 °C for 30 s; final extension was at 72 °C for 3 min. The resulting amplicons showed bands of about 500 bp in electrophoresis.

2.5.2.3 Index PCR, purification and DNA pulling
Index PCR (Figure 10), which adds a unique synthetic oligonucleotide index sequences to the end of DNA strands in each sample, by which samples can be distinguished after
sequencing (Son & Taylor, 2011), was performed with Nextera XT Index Kit (Illumina Inc., San Diego, CA, USA) according to the manufacturer’s protocol. Resulting amplicons were purified using AMPure XP beads (Beckman-Coulter, Brea, CA, USA) following the manufacturer’s protocol. Subsequently, amplicon concentration was assessed by measuring the intensity of the bands in the gel picture by the software Image Lab™ Software Version 4.0.1 (Bio-Rad Laboratories Lda, Hercules, California, USA) and equal quantities of each sample were pooled to generate the amplicon library for sequencing. This step was necessary to ensure equal amount of DNA from each sample to prevent uneven sequencing output of the samples. DNA quality and quantity of the pooled library were tested with an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Library denaturing, addition of internal control DNA (PhiX, Illumina) and sample loading were done according to the Illumina protocol. Sequencing was performed on a MiSeq desktop sequencer (Illumina Inc., San Diego, USA).

Figure 9 – Amplification scheme done by PCR. On denaturation (1), high temperature is applied to separate the strands from each other. On annealing (2) and elongation (3), primers complement to the target sequence and DNA polymerase synthesizes, from the end of the primer, new strands of DNA complementary to the target sequence. The process gets repeated exponentially (Microbe Online, 2017).

Figure 10 – Index PCR adds a unique synthetic oligonucleotide index sequence to the end of DNA strands in each sample, by which samples can be distinguished after sequencing (GenDx, 2017).

2.5.3 Electrophoresis

All genomic DNA preparations and all PCR amplicons were separated by agarose gel electrophoresis (100 V for 1h) in 1% (w/v) TBE agarose gels (Biozym Biotech Trading, Vienna, Austria) and visualized afterwards with a BioSpectrum imaging system (UVP, Upland, CA, USA). After weighing the required amount of agarose powder, it was dissolved
(1:100) in TBE buffer to achieve 1% (w/v) and heated up in the microwave until it was completely dissolved. After cooling down for a short time, 5 μl of Midori Green solution was added. The obtained fluid mixture was transferred to the electrophoresis chamber (Bio-Rad Laboratories Lda, Hercules, California, USA). After gel got solid the chamber was filled with TBE Buffer. DNA samples (5 μl) were mixed with 6x loading buffer (1% (w/v) of bromophenol blue, 30% (v/v) of glycerol in ddH$_2$O, 4% of 20mM EDTA pH 8) and loaded into the slots. The first lane was loaded with a DNA ladder (5 μl) with bands from 250 bp to 10,000 bp (Thermo Fisher Scientific, Waltham, MA, USA).

### 2.5.4 Dilutions

Dilutions of 1:10 of rhizosphere samples were carried out because the quality test by electrophoresis didn’t show bands of pure DNA isolate on the gel. Dilution of the PCR reagents was performed to decrease the amount of PCR-inhibitions, low molecular contaminants such as humic acids by agarose plug diffusion (Moreira 1998).

### 2.6 Bioinformatics and data processing

MiSeq raw data quality was checked in FastQC (Andrews, 2010) and reads were screened for PhiX contamination using Bowtie 2.2.6 (Langmead & Salzberg, 2012). Forward and reverse primers were stripped employing Cutadapt 1.8.3 (Martin, 2011). A Bayesian clustering for error correction was applied (Nikolenko, Korobeynikov, & Alekseyev, 2013; Schirmer et al., 2015) before merging the PE (paired-end) reads using PEAR 0.9.6 (p<0.001) (Zhang, Kobert, Flouri, & Stamatakis, 2014) and quality filtering performed in USEARCH v8.0.1517 (maximum expected error=0.5) (Edgar, 2013; Edgar & Flyvbjerg, 2015). METAXA2 was then used to target the extraction and to verify that the filtered sequences represent 16S rRNA V7-V9 region amplicons (Bengtsson-Palme et al., 2015). Targeted reads were labelled according to the genotype sample name and combined in QIIME (Caporaso et al., 2010). Sequences were dereplicated, sorted and clustered at 97 % of similarity using VSEARCH 1.1.1 (Rognes, 2017). Chimeras, sequences with mixed amplicons, *ergo*, amplicons of another origin, were checked adopting both a de-novo and a reference based approach, as routine of the above mentioned tool and filtered out. The RDP classifier training set v15 (09/2015) was used as a database for the reference based chimera detection. An optimal global alignment was applied afterwards in VSEARCH and a BIOM table generated. Taxonomy assignment was performed employing the naïve Bayesian RDP classifier v2.10 (Q. Wang, Garrity, Tiedje, & Cole, 2007) in QIIME with a minimum confidence of 0.8 against the SILVA database release 128 (Quast et al., 2013). After filtering out genomic DNA and non-amplicon sequences, a total of 120,607 sequences with 360 bp were obtained. R studio software (Rstudio, Boston, MA, USA) was used for filtering the Operational Taxonomic Unit (OTU) sequences, which each sequence is an OTU because it’s
not possible to claim for sure that any sequence belongs to a certain microorganism. To
determine if OTU’s were consistently found among the three replicates of a sampling
collection the term “reproducibly occurring OTU’s” was defined - OTU’s detected in at least
66% (2 of 3) of each genotype sample. To avoid biases originated from uneven number of
OTU’s per sample, only OTU’s that had a relative abundance of minimum 0.1% of the OTU’s
per sample were included. The resulting OTU table was processed by 1) splitting the OTU’s
according to the genotype; 2) collecting OTU’s that were present in at least two of three
replicates command; and 3) merging the resulting OTU’s in one table.

2.7 Statistical analysis
Statistical analysis was done using R studio software (Rstudio, Boston, MA, USA). The final
OTU count (OTU’s more abundant) was 1,842 in total.
An α-diversity analysis, which describes a community at one time and place composed by
the number of species (richness) and the shape of the distribution (evenness) of an individual
sample, was performed by calculating total OTU richness and calculating the Simpson index
of diversity - values between 0 (high diversity) and 1 (low diversity) by giving the probability
that two entities taken at random represent the same entity (\( \sum (n/N)^2 \); n = the total number of
organisms of a particular species; N = the total number of organisms of all species),
measuring a relative abundance of different species that enrich an area (Simpson, 1949).
Values were compared between samples by means of permutational pairwise comparisons
in the RVAideMemoire R package (Hervé, 2015) and PERMANOVA using the vegan
package (Hartmann et al., 2012; Oksanen et al., 2015). The resulting \( \rho \) values were adjusted
by False Discovery Rate (FDR). Richness and evenness values were then plotted into
boxplots via ggplot2 (Wickham & Chang, 2015) package in R.
A β-diversity analysis, which describes the diversity extent between different samples
composed by the variation in species composition and community variation, was performed
by a constrained analysis of principle coordinates (CAP), distance to centroid or distance-
based redundancy analysis by applying the capscale function; PERMANOVA using the vegan
package; and by ANOVA using Tukey post-hoc analysis to estimate the effects of the
genotype on the microbiome composition. Individuation and correlation of OTU’s responsible
for shaping the diversity structure were called core OTU’s (most prominent OTU’s from
reproducibly occurring OTU’s) and unique OTU’s, being predicted using Bray Curtis
dissimilarity (\( \frac{\sum \text{abs}(x_{ij} - x_{ji})}{\sum(x_{ii}+x_{jj})} \); x = variable; i, j = objects) (Bray & Curtis, 1957) for inference
measurement of OTU’s and abundance analysis.
Diagrams, illustrating the overlap in OTU occurrence between selected sample types, were
created using JVenn (Bardou, Mariette, Escudié, Djemiel, & Klopp, 2014).
2.8 Sequence analysis

Core OTU's and OTU's unique for each genotype, were analysed manually with the RDP classifier for taxonomy, and the NCBI database for highest similarities. In RDP classifier, each analysed sequence had a percentage in every taxonomic level. Below 95% the specific taxonomic level is not considered. Taxonomy analysis was done for relative abundant OTU's for all scientific classifications except for genus and species because of high ambiguity. Chimeras were checked with the use of BLAST at the NCBI database on the core and unique sequences. While working with plant material, sequences originating from chimeric amplification of plant chloroplast 16S rDNA or mitochondria 18S rDNA with bacteria 16S rDNA can be found. Chimera check was done as explained: 1) a sequence is checked by performing blast in NCBI site, on default settings, with the exclusion of “Uncultured/environmental sample sequences” on “Choose Search Set” (Figure 11), because only cultured bacteria was considered; 2) positive for chimera if the resulting graphic (Figure 12) with blast hits producing sequence significant alignments (Figure 13) shows empty parts with query cover below 95%, without all the base pairs. For example, in the Figures 11, 12 and 13, the sequence (GTAGTCCACGGCTAAACGATGTGACCTTTGGAGGTGTGGCTTCTCGGAGCTAAGCGGTAAAGTCGACCGCCTGGGGAGTACGGCGCCGAAGGGTAAACTCAATGAAATGGACGGGGGCGAACAACGCGAGCGAGCATGCGGATAATTGATGACAGCGAAGAACCTCATGAGCTTGACATGAACTGGAAACACCTAGAAATAGGTGCCCCGCTGGTCGGTTTACAGGTGGTGACATGGGCAAAGCTTACATGGCGTAGATGTTTTCGTAGTCCTGTAAAGCTAGGTGTTAAGTCCCGTAACGAGCGCAACCCTTGTCTAGTACCAGACATGTGAGGGCATTAGAGAGACT) from the OTU 16145, a unique sequence from the genotype primavera, is a chimera.
Figure 11 – NCBI page showing the procedure and options to choose.

Figure 12 – Resulting graphic with blast hits producing sequence significant alignments.
Faculty of Sciences of the University of Porto

Effect of tomato genotype on microbiome composition

Figure 13 – Sequences producing significant alignments show empty parts with query cover below 95%, without all the base pairs.
3 Results

3.1 OTU count overview

Community sequencing of rhizosphere and root samples of 18 tomato cultivars resulted in a total number of 1,842 OTU’s. The number of OTU’s per cultivar ranged from 785 in “Lyterno F1” to 1076 in “Roterno F1” (Table 2). In general, sequencing of rhizosphere resulted in a higher number of OTU’s than sequencing of root samples (Table 2). The number of unique OTU’s in the rhizosphere ranged from 453 in “Lyterno F1” to 796 in “Roterno F1” (Table 2), while the number of OTU’s in root samples ranged from 148 in “Roterno F1” to 303 in “Goldita”.

Table 2 – Overview of the number of Operational Taxonomic Units (OTU’s) in total, only in root, only in rhizosphere, in both regions, unique to the specific cultivar and common to all the 18 cultivars tested.

<table>
<thead>
<tr>
<th>Cultivar name</th>
<th>Number of OTU’s</th>
<th>Total</th>
<th>Root</th>
<th>Rhizosphere</th>
<th>Both regions</th>
<th>Unique</th>
<th>Common</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amoroso F1</td>
<td>999</td>
<td>232</td>
<td>23,2 %</td>
<td>642</td>
<td>64,3 %</td>
<td>125</td>
<td>12,5 %</td>
</tr>
<tr>
<td>Annamay F1</td>
<td>896</td>
<td>169</td>
<td>18,9 %</td>
<td>628</td>
<td>70,1 %</td>
<td>99</td>
<td>11,0 %</td>
</tr>
<tr>
<td>Auriga</td>
<td>793</td>
<td>207</td>
<td>26,1 %</td>
<td>510</td>
<td>64,3 %</td>
<td>76</td>
<td>9,6 %</td>
</tr>
<tr>
<td>Bartelt F1</td>
<td>1004</td>
<td>225</td>
<td>22,4 %</td>
<td>615</td>
<td>61,3 %</td>
<td>164</td>
<td>16,3 %</td>
</tr>
<tr>
<td>Benarys Gartenfreude</td>
<td>947</td>
<td>272</td>
<td>28,7 %</td>
<td>540</td>
<td>57,0 %</td>
<td>135</td>
<td>14,3 %</td>
</tr>
<tr>
<td>Black Cherry</td>
<td>960</td>
<td>286</td>
<td>29,8 %</td>
<td>562</td>
<td>58,5 %</td>
<td>112</td>
<td>11,7 %</td>
</tr>
<tr>
<td>Bocati F1</td>
<td>887</td>
<td>284</td>
<td>32,0 %</td>
<td>492</td>
<td>55,5 %</td>
<td>111</td>
<td>12,5 %</td>
</tr>
<tr>
<td>Cappricia F1</td>
<td>992</td>
<td>288</td>
<td>29,0 %</td>
<td>572</td>
<td>57,7 %</td>
<td>132</td>
<td>13,3 %</td>
</tr>
<tr>
<td>Goldita</td>
<td>896</td>
<td>303</td>
<td>33,8 %</td>
<td>473</td>
<td>52,8 %</td>
<td>120</td>
<td>13,4 %</td>
</tr>
<tr>
<td>Green Zebra</td>
<td>913</td>
<td>211</td>
<td>23,1 %</td>
<td>592</td>
<td>64,8 %</td>
<td>110</td>
<td>12,0 %</td>
</tr>
<tr>
<td>Harzfeuer F1</td>
<td>1031</td>
<td>274</td>
<td>26,6 %</td>
<td>639</td>
<td>62,0 %</td>
<td>118</td>
<td>11,4 %</td>
</tr>
<tr>
<td>Lyterno F1</td>
<td>785</td>
<td>232</td>
<td>29,6 %</td>
<td>453</td>
<td>57,7 %</td>
<td>100</td>
<td>12,7 %</td>
</tr>
<tr>
<td>Primabella</td>
<td>953</td>
<td>263</td>
<td>27,6 %</td>
<td>548</td>
<td>57,5 %</td>
<td>142</td>
<td>14,9 %</td>
</tr>
<tr>
<td>Primavera</td>
<td>1010</td>
<td>252</td>
<td>25,0 %</td>
<td>633</td>
<td>62,7 %</td>
<td>125</td>
<td>12,4 %</td>
</tr>
<tr>
<td>Resi</td>
<td>1035</td>
<td>203</td>
<td>19,6 %</td>
<td>713</td>
<td>68,9 %</td>
<td>119</td>
<td>11,5 %</td>
</tr>
<tr>
<td>Roterno F1</td>
<td>1076</td>
<td>148</td>
<td>13,8 %</td>
<td>796</td>
<td>74,0 %</td>
<td>132</td>
<td>12,3 %</td>
</tr>
<tr>
<td>Sakura F1</td>
<td>983</td>
<td>168</td>
<td>17,1 %</td>
<td>653</td>
<td>66,4 %</td>
<td>162</td>
<td>16,5 %</td>
</tr>
<tr>
<td>Supersweet 100 F1</td>
<td>1059</td>
<td>207</td>
<td>19,5 %</td>
<td>705</td>
<td>66,6 %</td>
<td>147</td>
<td>13,9 %</td>
</tr>
</tbody>
</table>

Rhizosphere and roots share 522 OTU’s, while 570 OTU’s are specific to roots and 750 OTU’s are unique in the rhizosphere (Figure 14).
Comparison of fruit types revealed that 247 OTU’s are specific to cocktail tomatoes and 130 OTU’s are specific to salad tomatoes, while 1,465 OTU's were present in both fruit types (Figure 15).

3.2 Alpha-diversity analysis

Alpha-diversity is defined by calculating total OTU richness and Simpson index of diversity. Total OTU richness and Simpson index of diversity showed a clear and significant difference between roots and rhizosphere (Figure 16), confirmed by PERMANOVA (P<0.001).

The average of roots OTU richness was 210, while rhizosphere OTU richness was 4300. Roots show less microbial richness compared with rhizosphere. In contrast to observed OTU richness, Simpson index of diversity takes also the abundances of OTU’s into account to calculate diversity in a sample. The average of roots Simpson index of diversity was 0.74 while rhizosphere Simpson index of diversity was 0.99. Lower values indicate more diversity while higher values indicate less diversity.
Effect of tomato genotype on microbiome composition

Figure 16 – Alpha diversity within roots and rhizospheres, from the 18 tomato cultivars tested, measured by observed OTU richness and Simpson index of diversity. A permutation pairwise comparison (RVAdiEMeroire R package) showed that the richness and the diversity values were significantly different (p = 0.0002, perm = 9999).

Between different genotypes, total OTU richness and Simpson index of diversity showed no significance, which was confirmed by PERMANOVA (P=1). As such, R studio was only able to show the mean of all genotypes together (Figure 17) for the observed OTU richness and Simpson index of diversity. The average of the OTU richness was 1,700 and 50 % of the values ranged from 100 to 4,100. In terms of Simpson index, the average was 0.93 and 50% of the samples ranged from 0.70 to 1.

Figure 17 – Alpha diversity of the 18 cultivars tested in terms of observed OTU richness and Simpson index of diversity. A permutation pairwise comparison showed (p = 0.97, perm = 9999) that the richness and the diversity values were not significant.

In the comparison of the two fruit types, total OTU richness and Simpson index of diversity showed no significance, which was confirmed by PERMANOVA (P=0.9787). Because there was no significance, R studio was only able to show the mean of all fruit type samples together for the observed OTU richness and Simpson index of diversity. When considering fruit type samples, the result was the same as the genotype analysis (Figure 17) because of the lack of significance, being all samples considered instead.
3.3 Beta-diversity analysis

Beta-diversity is a measure of the differences in diversity between the communities in different samples. It takes into account the variation in species composition and abundance. Variation between genotype samples are illustrated in Figure 18. The distance to centroid (i.e. the distance of the group centroids to the principal coordinates) were not significantly different between genotypes, which was confirmed by PERMANOVA (P=1), which showed an average distance to centroid between 0.40 and 0.43. This means that in terms of community composition the microbiome in different tomato genotypes was similar.

Tukey post-hoc analysis, which gives a set of confidence intervals on the differences between the means of all genotypes with a specified family-wise probability of coverage, showed no significant difference between genotypes with a family-wise probability of 95 % confidence level. This means that the genotypic factor had essentially no effect on the bacterial community.

![Bray-curtis beta diversity among genotype samples. A permutation test assessed the significance of the treatment on a Constrained Analysis of Principal Coordinates (CAP) (p > 0.415, perm = 999) (vegan R package).](image)

All genotypes grown in the same field and with the same soil characteristics can be expected to have been similar in microbiome composition across the field. Thus, similar microbiome
compositions in the rhizosphere of different genotypes are not surprising. However, one would expect a higher influence of the plant genotype. In order to confirm whether or not there was an effect of the genotype on the microbiome from root samples (Figure 19) and rhizosphere samples (Figure 20), individual analyses were made. As a result, no significant effect of the plant genotype on the beta-diversity on both root and rhizosphere was found. PERMANOVA (P=1) confirmed that there was no significant difference between root and rhizosphere samples of different tomato genotypes, which showed an average distance to centroid between 0.34-0.44 and 0.36-0.41, respectively.

Tukey post-hoc analysis with a family-wise probability of 95% confidence level was not significantly relevant in both cases. Variance in roots and rhizosphere was not affected by the plant genotype.

Figure 19 – Average Bray-curtis distance to centroid as an indication of beta diversity in the roots from the 18 genotypes tested. A permutation test assessed the significance of the treatment on a Constrained Analysis of Principal Coordinates (CAP) (p > 0.389, perm = 999).
Comparison of the roots and rhizosphere microbiomes independently from plant genotype revealed that species composition and community variation (Figure 21) were significantly different, confirmed by PERMANOVA (P=0.001), which showed an average distance to centroid of 0.35 and 0.18, respectively.

Tukey post-hoc analysis with a family-wise probability of 95 % confidence level, giving a set of confidence intervals on the differences between the means of roots and rhizosphere, showed significance with a difference of 18 %.

In terms of beta-diversity in cocktail and salad tomato plants, no significant differences were found between those groups, confirmed by PERMANOVA (P=1), which showed an average
distance to centroid of 0.43 and 0.46, respectively. Tukey post-hoc analysis with a family-wise probability of 95 % confidence level showed also no significant difference.

3.4 Taxonomy

For phylogenetic description of the bacterial assemblages in roots and rhizosphere of tomato plants OTU’s were classified according to their matches with available databases. Since the majority of OTU’s classification at the Family or Genus level was not possible (too low confidence at these taxonomic levels), the phylogenetic description of the communities was done at higher levels, namely Kingdom, Phylum, Class and Order (Figures 23 and 24).

3.4.1 By plant compartment

Rhizosphere samples contained both Bacteria and Archaea (13 % of total community), while only Bacteria could be detected in roots.

Phylogenetic profiling revealed that roots and rhizosphere had distinct bacterial communities (Figure 23). In roots, the predominant Phyla were Proteobacteria (50 %), Bacteroidetes (24 %) and Actinobacteria (20 %), while in the rhizosphere the groups that prevailed were Actinobacteria (40 %), Proteobacteria (35 %), and Chloroflexi (10 %) (Figure 23A).

At Class level, roots were dominated by Betaproteobacteria (25 %), Actinobacteria (23 %) and Flavobacteria (18 %) while the rhizosphere was dominated by Actinobacteria (25 %), Betaproteobacteria (15 %) and Alphaproteobacteria (13 %) (Figure 23B).

Analysis on Order level revealed that the bacterial assemblage in roots was dominated by Burkholderiales (20 %), Micrococcales (15 %) and Flavobacteriales (17 %) while the rhizosphere was dominated by Burkholderiales (13 %), Micrococcales (10 %) and Rhizobiales (10 %) (Figure 23C).
Figure 23 – Relative OTU abundance at (A) Phylum, (B) Class and (C) Order level in roots (left) and rhizosphere (right).
3.4.2 By genotype

The bacterial communities colonizing rhizosphere and roots of different tomato genotypes were very similar. Genotypes did not show distinct communities (Figure 24). In general, *Proteobacteria, Actinobacteria* and *Bacteroidetes* (Figure 24A) prevailed among the bacterial communities. The predominant Classes were *Actinobacteria, Betaproteobacteria* and *Flavobacteria* (Figure 24B) and at Order level the bacterial communities were predominantly composed by *Burkholderiales, Micrococcales* and *Flavobacteriales* (Figure 24C).
Figura 24 – Relative OTU abundance at (A) Phylum, (B) Class and (C) Order level in all eighteen cultivars.
3.5 Core and unique OTU’s

After chimera check, the number of core and unique OTU’s decreased remarkably (Table 5). Chimera check on the core and unique sequences showed that about 37 % of sequences were considered chimeras. From the 276 core and unique OTU’s, 174 were chimera free. One can only assume that the other 1,566 not checked OTU’s, had the same problem. The core OTU’s are per definition the proportion of the microbiome, which is common to all samples of a given sample set. In the present study, the core means the most prominent OTU’s present in all 18 different plant genotypes. Unique OTU’s are such OTU’s that are exclusive to individual genotypes. The OTU’s were analysed on RDP classifier. The total 174 OTU’s were summed up in 41 Genus, 16 Families, 5 Orders, 1 Sub-order, 4 Classes, 3 Phyla and 1 Kingdom classified using the RDP classifier (Attachment B).

Table 3 – Number of core and unique OTU’s with chimeras and without chimeras.

<table>
<thead>
<tr>
<th>Cultivar name</th>
<th>Number of OTU’s</th>
<th>With chimeras (Total: 276)</th>
<th>Without chimeras (Total:174)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Core Unique</td>
<td>Core Unique</td>
<td></td>
</tr>
<tr>
<td>Amoroso F1</td>
<td>9</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Annamay F1</td>
<td>9</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Auriga</td>
<td>11</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Bartelly F1</td>
<td>13</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Benarys Gartenfreude</td>
<td>10</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Black Cherry</td>
<td>14</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Bocati F1</td>
<td>10</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Capriccia F1</td>
<td>18</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Goldita</td>
<td>17</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Green Zebra</td>
<td>15</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Harzfeuer F1</td>
<td>15</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Lyterno F1</td>
<td>7</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Primabella</td>
<td>16</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Primavera</td>
<td>22</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Resi</td>
<td>16</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Roterno F1</td>
<td>31</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Sakura F1</td>
<td>5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Supersweet 100 F1</td>
<td>15</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

3.5.1 Core OTU’s

Out of 22 sequences comprising the core OTU’s (Table 4), 15 could be assigned to a Genus. Eight of them were classified as Microbacterium, Pseudoxanthomonas, Rhizobacter, Pseudomonas, Variovorax, Polaromonas, Sphingobium and Ensifer. Two OTU’s each were classified as Arthrobacter and Streptomyces, respectively. Three OTU’s were identified as Flavobacterium.
Three OTU’s could be assigned to Families, one from Oxalobacteraceae Family and two from Comamonadaceae Family. The remaining OTU’s could be identified as members of Rhizobiales and Actinomycetales (Order), or Micrococcineae (Sub-order) and Gammaproteobacteria (Class).

Table 4 – Taxonomic classification and number of core OTU’s included in each classification.

<table>
<thead>
<tr>
<th>Genus</th>
<th>Family</th>
<th>Suborder</th>
<th>Order</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Microbacterium</td>
<td>1 Oxalobacteraceae</td>
<td>1 Micrococcineae</td>
<td>1 Rhizobiales</td>
<td>1 Gammaproteobacteria</td>
</tr>
<tr>
<td>1 Pseudoxanthomonas</td>
<td>2 Comamonadaceae</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Rhizobacter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Pseudomonas</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Variovorax</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Polaromonas</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Sphingobium</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Ensifer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Arthrobacter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Streptomyces</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Flavobacterium</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.5.2 Unique OTU’s

The two unique OTU’s present in “Amorosi F1” belonged to the Family of Chitinophagaceae, with one of the sequences being classified at Genus level as Taibaiella.

Out of three OTU’s present only in “Annamay F1”, two were classified at Family level as Flavobacteriaceae whereas the other was classified at Order level as Burkholderiales.

From the six OTU’s unique to “Auriga”, one was classified at Genus level as Rothia, two at Family level as Paenibacillaceae and Acetobacteraceae, two at Family level as Flavobacteriaceae, and one at Phylum level as Firmicutes.

Out of eight OTU’s unique to “Bartelly F1”, one was classified at Genus level as Alicyclobacillus, one at Family level as Moraxellaceae, five at Order level as Actinomycetales, and one was classified at Sub-order level as Micrococcineae.

From the eight OTU’s present only in “Benarys Gartenfreude”, two were classified at Genus level as Shinella and Alicyclobacillus, two at Family level as Flavobacteriaceae and Sphingomonadaceae, one at Order level as Flavobacterium, one at Sub-order level as Micrococcineae, and two were classified at Class level as Gammaproteobacteria.

The microbiome of “Black Cherry” harboured eleven unique OTU’s. Six of them were classified at Genus level as Marinobacter, Alicyclobacillus, Chitinophaga, Agrococcus, Paenibacillus and Rheinheimera, one at Family level as Flavobacteriaceae, three at Class level as Gammaproteobacteria, and one remained unidentified with assignment to the Kingdom Bacteria only.

From the five OTU’s unique to “Bocati F1”, one was classified at Genus level as Brachybacterium, one at Family level as Micrococcaceae, one at Order level as Cytophagales, and two were classified at Order level as Actinomycetales.
Eight OTU’s were unique to “Cappricia F1”. Five of them were classified at Genus level as *Sphingobacterium*, *Kyrolidia*, *Sphingomonas*, *Romboutsia* and *Dyadobacter*, one at Family level as Flavobacteriaceae, and two were classified at Class level as *Alphaproteobacteria* and *Betaproteobacteria*. 

Out of eight OTU’s present only in “Goldita”, one was classified at Genus level as *Microbispora*, one at Family level as *Micrococcaceae*, one at Order level as *Actinomycetales*, two at Class level as *Alphaproteobacteria* and *Betaproteobacteria*, and three at Phylum level as *Proteobacteria*.

From the nine OTU’s unique to “Green Zebra”, four OTU’s were classified at Genus level as *Taibaiella*, *Methylocaldum*, *Pseudomonas* and *Paenibacillus*, three at Sub-order level as *Micrococcineae*, one at Phylum level as *Proteobacteria*, and one remained unidentified with assignment to the Kingdom *Bacteria* only.

Ten OTU’s were found only in “Harzfeuer F1”. One was classified at Genus level as *Ohtaekwangia*, one at Family level as *Flavobacteriaceae*, two at Family level as *Paenibacillaceae*, one at Order level as *Actinomycetales*, two at Order level as *Cytophagales*, one at Sub-order level as *Micrococcineae*, one at Phylum level as *Proteobacteria*, and one remained unidentified with assignment to the Kingdom *Bacteria* only.

From the five OTU’s present only in “Lyterno F1”, one was classified at Sub-order level as *Micrococcineae*, one at Phylum level as *Bacteroidetes*, and three at Kingdom level as *Bacteria*.

From the fourteen OTU’s unique to “Primabella”, five were classified at Genus level as *Asticcacaulis*, *Dyadobacter*, *Chryseobacterium*, *Kaistia* and *Flavobacterium*, one at Family level as *Comamonadaceae*, two at Family level as *Flavobacteriaceae*, two at Family level as *Sphingobacteriaceae*, two at Class level as *Alphaproteobacteria* and *Betaproteobacteria*, and two at Phylum level as *Proteobacteria* and *Bacteroidetes*.

Ten OTU’s were unique to “Primavera”. Two of them were classified at Genus level as *Bacillus* and *Desulfospirorosinus*, one at Family level as *Geodermatophilaceae*, three at Order level as *Burkholderiales*, *Rhizobiales* and *Actinomycetales*, one at Class level as *Actinobacteria*, two at Class level as *Betaproteobacteria*, and one remained unidentified with assignment to the Kingdom *Bacteria* only.

The microbiome in “Resi” harboured eleven unique OTU’s. Three of them were classified at Genus level as *Arthrobacter*, *Pedobacter* and *Paenibacillus*, two at Family level as *Streptosporangiacaeae* and *Flavobacteriaceae*, one at Order level as *Rhizobiales*, two at Order level as *Cytophagales*, one at Class level as *Betaproteobacteria*, one at Phylum level as *Proteobacteria*, and one at Kingdom level as *Bacteria*. 
With 22, “Roterno F1” harboured the highest number of unique OTU’s among all genotypes analysed. Seven of them were classified at Genus level as *Exiguobacterium, Geobacillus, Delftia, Corynebacterium, Schlegelella, Roseomonas* and *Pandoraea*. Five were identified at Family level as *Neisseriaceae, Moraxellaceae, Pseudonocardiaceae, Microbacteriaceae* and *Burkholderiaceae*, two at Family level as *Microbacteriaceae*, one at Order level as *Rhizobiales*, two at Order level as *Burkholderiales*, two at Class level as *Betaproteobacteria* and *Gammaproteobacteria*, one at Phylum level as *Bacteroidetes*, and two remained unidentified with assignment to the Kingdom *Bacteria* only.

Out of three OTU’s present only in “Sakura F1”, one was classified at Genus as *Dyadobacter*, and two at Family level as *Comamonadaceae* and *Micrococcaceae*. From the nine OTU’s unique to “Supersweet 100 F1”, one was classified at Genus as *Paenibacillus*, one at Family level as *Sneathiellaceae*, two at Family level as *Flavobacteriaceae*, one at Order level as *Burkholderiales*, two at Class level as *Gammaproteobacteria*, one at Phylum level as *Proteobacteria*, and one remained unidentified with assignment to the Kingdom *Bacteria* only.
4 Discussion

4.1 OTU’s count overview

16S rRNA gene fragment amplicon sequencing and subsequent data quality filtering resulted in a total number of 1,842 OTU’s. The number of OTU’s in the different tomato varieties were not different, indicating that the plant genotype may have no effect on the number of OTU’s in the microbiome of tomato roots and rhizosphere.

The number of observed OTU’s in the rhizosphere was in general higher than in the corresponding root. This is not surprising, as it is common knowledge that the rhizosphere is a hot spot for microbial activity and the region from where the microorganisms migrate into plants via the root and establish sub-communities inside plants (Hardoim et al., 2015). The root exudates and mucilage attract and select a high number of different soil microorganisms (Mitter, Pfaffenbichler, & Sessitsch, 2016) making some of those microbes permanent endophytes of the root. It has been suggested that the root microbial assemblage is a subset of the rhizosphere microbiota.

When comparing different fruit types (cocktail or salad tomatoes) no difference in OTU counts were observed, indicating that breeding for certain tomato fruit types, may have no effect on the number of different bacteria colonizing roots and rhizosphere of tomato plants. However, for example, a study with three different potato cultivars showed that a portion of the detected OTU’s was cultivar-specific (Weinert et al., 2011). Similar cultivar-dependent effects were observed for the rhizobacterial communities in the rhizosphere of young potato plants (İnceoğlu, Al-Soud, Salles, Semenov, & van Elsas, 2011). In another study, OTU richness was affected by maize genotypes (Peiffer et al., 2013).

4.2 Alpha-diversity analysis

Biodiversity indexes were different in rhizosphere and root, indicating a richer microbial assemblage in the rhizosphere when compared with roots. This was expected and often shown before (Hardoim et al., 2015).

The lower microbial richness of root microbiota can be explained by the specific recruitment of microbes made by the plant, for protection and growth, increasing plant available forms and soil stocks of carbon and nitrogen by mediating the ecosystem by ways of plant functionality traits (de Bello et al., 2010).

Against the expectations, microbial diversity was lower in the rhizosphere than in the roots. With such high richness it should be expected that the evenness would be low. The rhizosphere microbiome is responsible for protection against pathogen attack and growth, contributing to overall fitness of the plant. And since the rhizosphere is the zone of influence
created by the roots through their exudates and by the exudates of the microorganisms within the soil matrix, only the microbes "agreed" in the exchange are present in the rhizosphere. Plants can modulate their rhizosphere microbiome in a host-dependent way (Pérez-Jaramillo, Mendes, & Raaijmakers, 2016). The higher microbial diversity in roots could maybe be explained by the characteristics of the root habitat, which could favour spatial distribution and restricts population size, leading to a higher biodiversity.

A deeper insight on the microbiota in rhizosphere and root could give a broader knowledge on how rhizosphere-root relationship functions. Which microbes are present, how and why they contribute to the relationship is largely unknown and needs further analysis.

Furthermore, richness and diversity of the microbial community in roots and rhizosphere of tomato plants was not affected by the genotype. With constant agronomic advances, cultivars are developed to be grown with low inputs. This study suggests that the development of different tomato cultivars do not significantly modify the plants microbiome. Plant breeding acts on the genetic of the plant, not on the plant microbiome.

Different genotypes of the same species do not necessarily have different microbiomes. This was the case for the 18 tomato cultivars tested in this study. In the same direction, as genetic variation is in general low within cultivated tomato also plant physiology might not be that different when comparing different genotypes, as it was the case. Differences in the plant physiology might be important in diversification of plant microbiomes. Thus, plant species with higher genotypic variations could eventually show correlation between genotype and microbiome. The hypothesis that genetic modification and variation of plants effects microbiome richness and diversity is raised.

### 4.3 Beta-diversity analysis

Bacterial community structure was not significantly different between the tested genotypes, also not when considering only root or rhizosphere samples between genotypes. But when comparing the bacterial communities in the different plant compartments, rhizosphere and roots, without considering the genotype, significant differences in structure were found. Genotypes do not affect microbiome composition in tomato plants and thus tomato breeding has seemingly little effect on the tomato microbiome. However, for example, one study showed that plant genotype strongly modifies the structure and growth of maize rhizosphere microbial communities (Aira, Gómez-Brandón, Lazcano, Bååth, & Domínguez, 2010). Another study concluded with their findings that a small change in plant genotype can have complex and unforeseen effects on the plant microbiome (Turner, James, & Poole, 2013). More research has demonstrated that host genotype also influences the overall composition of both rhizosphere and root communities (Dayakar Badri, Chaparro, Zhang, Shen, & Vivanco, 2013; Bulgarelli et al., 2013, 2015).
As expected, species composition and community variation were significantly different between rhizosphere and roots.

4.4 Taxonomy

The root microbiota was dominated by the Phylum *Proteobacteria*, with *Burkholeriales* and *Rhizobiales* being the most prominent Orders. *Proteobacteria* are gram-negative typical endophytes playing key roles in the carbon, sulfur and nitrogen cycles (Kersters, De Vos, et al., 2006), being the largest and most diverse Phylum in the *Bacteria* domain (Marin, 2011). *Rhizobiales* live symbiotically in the roots of leguminous plants and play a key role in nitrogen fixation in the biosphere (Kersters, Lisdiyanti, Komagata, & Swings, 2006; Gyaneshwar et al., 2011; Pini, Galardini, Bazzicalupo, & Mengoni, 2011), which is a reason for being present in the rhizosphere. *Burkholderiales* are known for their biodegradative capacities and antagonistic properties towards multiple soil-borne fungal pathogens (Benitez & Gardener, 2009; Chebotar et al., 2015; Kumar et al., 2017).

*Bacteroidetes* are typical endophytes, were the second most predominant group of bacteria in the root microbiota, with *Flavobacteria* being the most prominent Class among them. *Flavobacteria* are gram-negative microbes specialized in the degradation of complex organic matter in the biosphere, especially in the form of polysaccharides and proteins (Church, 2008).

It was speculated that root plays an important role in the selection of the microbiota which is more useful for the plant health and development (Bulgarelli et al., 2013; Hardoim et al., 2015), while the composition in the rhizosphere is mainly driven by the exchange of exudates (Hardoim et al., 2015).

In the rhizosphere microbiota, the *Actinobacteria* prevailed, with the order *Micrococcales* being predominant. *Actinobacteria* are ubiquitously distributed in both aquatic and terrestrial ecosystems, playing an important role as symbionts and as pathogens in plant-associated microbial communities (Barka et al., 2016). They are gram-positive microorganisms capable of colonizing the rhizosphere through antagonistic and competitive characteristics concerning other soil microorganisms (Bulgarelli et al., 2013; Barka et al., 2016) preventing the harmful effects of one or more deleterious microorganisms through biocontrol or antagonism towards soil plant pathogens. Contributing also with nitrogen fixation, siderophore synthesis, phytohormone synthesis, and solubilization of minerals to make them available for plant uptake and use (Glick, 1995).

The Phylum *Chloroflexi* was specific to the rhizosphere and not present in roots. This type of bacteria are mostly gram-negative, containing aerobic thermophiles, which use oxygen and grow well in high temperatures and anoxygenic phototrophs, which use light for photosynthesis but do not produce oxygen as a byproduct (Hanada, 2014). The presence of
these bacteria in the rhizosphere might be due to the low oxygen content and they might produce byproducts that can be used by the plant.

Apart from the predominance of bacteria in both rhizosphere and root microbiota, the rhizosphere contained about 12.5 % of Archaea, which are mostly anaerobic (Valentine, 2007) and play an important role in chemical cycling producing, for example, nitrite, which other microbes then oxidize to nitrate that plants and other organisms consume (Leininger et al., 2006).

It is difficult, when not impossible, to take conclusions from phylogeny of microbiota in roots and rhizosphere relatively to their function in plant-microbe interaction. However, given the importance of rhizosphere and root microbial community on plant health and development, one can speculate, that plant health or illness could be predicted from prevalent taxa in the plant microbiome.

### 4.5 Core and unique OTU’s

The core microbiome is the part of the microbiome which is consistently found dominant in a given habitat, independently from changing environmental conditions or, in case of plant microbiomes, the plant genotype (Pfeiffer et al., 2016). The unique OTU’s might represent those members of the microbiome that are specific to a given plant genotype, implying a close relationship relying on genotype-genotype combinations. In view of the hypothesis underlying this study, the core microbiome might not have a strong influence on the efficacy of microbial inoculants on different varieties of the same plant species in plant production and protection. The unique microbiota, which are plant genotype specific, in contrast, could eventually affect establishment and activity of microbial inoculants. Possible mechanisms include direct antagonism, out-competition in space and use of nutrients or indirect via modulating plant genetic and metabolism counteracting establishment and activity of the inoculum. However, more comprehensive studies including inoculation experiments and functional analysis of individual members of the microbiota in different tomato varieties are needed to verify this hypothesis. A description of the found species can be seen on Attachment C.

### 4.6 Sequencing data quality

The sequencing data set contained a high number of chimeras between bacterial 16S rRNA genes and chloroplast and mitochondria small subunit RNA genes, respectively. This results from incorrect binding of short amplicons during PCR. Bacterial 16S rRNA gene and chloroplast 16S rRNA gene as well as mitochondria 18S rRNA genes are in many sequence areas highly similar and they are co-amplified in PCR using conserved primers. In this study, the chimeras almost invariably consisted of recombinations between short stretches of about 20-50 bp deriving from plant organelles and large stretches of bacteria 16S rRNA fragments.
This could be one reason why these chimera were not discovered and filtered out in the automated chimera check during sequence quality filtering. Raw sequence reads were initially subjected to standard quality filtering as described in Material and Methods and this includes an automated chimera check, which is based on aligning the reads to 16S rRNA gene databases. The databases used are restricted to bacterial 16S rRNA genes and thus chloroplast and mitochondria sequences might remain unidentified in this analysis. New and advanced chimera check tools are needed. Those that would take into account also plant organelle genes to ensure chimera free data sets from sequencing plant colonizing bacterial communities. Another approach would be to reduce the proportion of plant DNA during sample preparation. In early microbiome studies, based on 16S rRNA gene PCR cloning and Sanger sequencing, bacterial cells were mechanically dislodged from plant tissue (Garbeva, van Overbeek, van Vuurde, & van Elsas, 2001). In a more recent study, this resulted in a reduction of plant organelle derived amplicons to less than 4% (Drage et al., 2014).
5 Conclusions

In this study, the plant genotype had no significant effect on the microbiome composition. There was also no significant difference in the microbial assemblage between cocktail and salad tomato. As expected, the rhizosphere and root harboured significantly different bacterial communities. In sum, the present study concluded, that the tomato genotype has no effect on the microbiome composition, showing no significant differences between genotypes, but significant differences between rhizosphere and roots microbiomes. The taxonomic analysis in this study suggests that plant-bacteria associations, although quite variable, are composed by a conservative set of taxa. Furthermore, the high number of chimera in the sequencing showed, that the sample preparation for 16S rRNA gene amplicon sequencing of plant material needs further improvement. The present study represents an initial step to test the hypothesis that plant genotype might have an effect on the microbiome, and that this could be due to differences in the microbiota of different genotypes which, in turn, condition in different ways the establishment and activity of the microorganisms. This hypothesis is based on the assumption that the microbiome in rhizosphere and roots is significantly influenced by the plant genotype. This was not the case in this study. However, the results presented here derived from sequencing only a short fragment of the 16S rRNA gene, which does not allow for identification of microbiota on species level. It cannot be excluded that the microbiome of different genotypes is similar on family level but differs in species level only. More studies are needed, also such allowing for the analysis of microbiota on species or even strain level to fully understand the complex interaction between soil, plant and microbial inoculum.
6 References


Effect of tomato genotype on microbiome composition


Effect of tomato genotype on microbiome composition


Attachment A | Modified DNA Isolation manual from Van der Beek

**Chemicals:**

<table>
<thead>
<tr>
<th>STE Buffer</th>
<th>500 mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sorbitol</td>
<td>0,35M</td>
</tr>
<tr>
<td>TrisHCl</td>
<td>0,1M</td>
</tr>
<tr>
<td>EDTA</td>
<td>5mM</td>
</tr>
<tr>
<td>Na Bisulfite</td>
<td>20mM</td>
</tr>
<tr>
<td>PVP-40</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lysis buffer</th>
<th>500 mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>TrisHCl</td>
<td>0,2M</td>
</tr>
<tr>
<td>EDTA</td>
<td>0,05M</td>
</tr>
<tr>
<td>NaCl</td>
<td>2M</td>
</tr>
<tr>
<td>CTAB</td>
<td>2%</td>
</tr>
</tbody>
</table>

| Sarkosyl     | 5%     | 5 g for 100 mL |

<table>
<thead>
<tr>
<th>Chloroform/Isoamylalcohol (24:1)</th>
<th>250 mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chloroform</td>
<td>240 mL</td>
</tr>
<tr>
<td>Isoamylalcohol</td>
<td>10 mL</td>
</tr>
</tbody>
</table>

**Protocol:**

- Grind 100 mg (= 0.1g) of plant material in 2 ml lysing matrix tube E (under liquid nitrogen) - must be frozen until buffer is added;
- Add 310 µL STE buffer (shortly before STE, Na bisulfite and PVP-40 mix);
- Add 440 µL Lysis buffer;
- Add 150 µL Sarkosyl;
- Vortex 10-15 seconds;
- Use the Fast Prep Instrument for 40 seconds with speed setting of 6.0;
- 1 h at 65 ° C. in a water bath;
- Centrifugation 14,000xg 15min rcf (relative centrifugal force);
- Transfer liquid into 2ml eppendorf tubes;
- 5min at room temperature;
- Add 900 µL Chlorofom/Isoamylalcohol (24:1);
- Centrifuge at 10 000 rpm for 10 min;
- Add 900 µL of isopropanol to the mixture, mix slowly (DNA filaments form);
- At RT for 1h (or overnight at -20 degrees freeze and thaw for 1h);
- Centrifuging max. Speed for 10 min;
- Remove supernatant;
- Add 300 µL 70% ethanol to pellet and centrifuge 5 min at max;
- Remove supernatant and allow pellet to dry (approx. 1h, less if possible);
- Dissolve the pellet in 100 µL TE (tris-EDTA);
- 1 μL RNase (20 μg / mL) (gel electrophoresis already possible);
- 1 h at 37 °C. in a warming cabinet, followed by gel electrophoresis;
- Store DNA samples at 4 °C.
Attachment B | Taxonomy summary of core and unique OTU’s

- At Genus level: Agrococcus, Alicyclobacillus, Arthrobacter, Asticcacaulis, Bacillus, Brachybacterium, Chitinophaga, Chryseobacterium, Corynebacterium, Delftia, Desulfosporosinus, Dyadobacter, Ensifer, Exiguobacterium, Flavobacterium, Geobacillus, Kaistia, Kyrpidia, Marinobacter, Methylocaldum, Microbacterium, Microbispore, Ohtaekwangia, Paenibacillus, Pandoraea, Pedobacter, Polaromonas, Pseudomonas, Pseudoxanthomonas, Rheinheimera, Rhizobacter, Romboutsia, Roseomonas, Schlegelella, Shinella, Sphingobacterium, Sphingobium, Sphingomonas, Streptomyces, Taibaiella and Variovorax;
- At Family level: Acetobacteraceae, Burkholderiaceae, Chitinophagaceae, Comamonadaceae, Flavobacteriaceae, Geodermatophilaceae, Microbacteriaceae, Micrococcaceae, Moraxellaceae, Neisseriaceae, Oxalobacteraceae, Paenibacillaceae, Pseudonocardiaeae, Sneathiellaceae, Sphingomonadaceae and Streptosporangiaceae;
- At Order level: Actinomycetales, Burkholderiales, Cytophagales, Flavobacterium and Rhizobiales;
- At Sub-order level: Micrococcineae;
- At Class level: Actinobacteria, Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria;
- At Phylum level: Bacteroidetes, Firmicutes and Proteobacteria;
- At Kingdom level: Bacteria.
Attachment C | Species Description

*Acetobacteraceae* Family harbours many species of biotechnological importance, being described as nitrogen fixing bacteria able to promote plant growth by a variety of mechanisms (Reis & Teixeira, 2015). They are Gram-negative and aerobic that conduct an incomplete oxidation of sugars and alcohols to produce organic acids as final product of their metabolism (Reis & Teixeira, 2015). They can grow in very acidic environments with pH close to 3.0–3.5, being the optimum range between 5.0–6.5 (Kersters, Lisdiyanti, et al., 2006).

*Burkholderiaceae* Family are endophytic microbes with the capacity to inhibit fungal growth (Acton, 2012)

*Chitinophagaceae* Family is known for chitin degradation having antifungal activities (Hargreaves, Williams, & Hofmockel, 2015)

*Comamonadaceae* Family are key players for mineralization of carbon bound sulfur, allowing cycling of soil sulfate between organic and inorganic forms (Schmalenberger et al., 2008)

*Flavobacteriaceae* Family promotes plant growth with auxin production, the ability of using 1-Amino Cyclopropan-1-Carboxylate (ACC) as nitrogen source and phosphate-solubilization (Soltani et al., 2010)

*Geodermatophilaceae* Family contributes with nitrogen fixation, siderophore synthesis, phytohormone synthesis, and solubilization of minerals (Glick, 1995).

*Microbacteriaceae* Family contains plant pathogens but can also contribute to the control of bacterial and fungal pathogens (Evtushenko & Takeuchi, 2006; Hallmann, Quadt-Hallmann, Mahaffee, & Kloepper, 1997; Reiter, Pfeifer, Schwab, & Sessitsch, 2002)

*Micrococcaceae* Family are capable of colonizing the rhizosphere through antagonistic and competitive characteristics concerning other soil microorganisms (Barka et al., 2016; Bulgarelli et al., 2013).

*Moraxellaceae* Family is known for its pathogenicity in animals, belonging to the Phylum *Proteobacteria* the largest and most diverse in the *Bacteria* Kingdom (Marin, 2011). There isn’t relevant literature about the activity of this Family in the soil or in the plant.

*Neisseriaceae* Family is known for its pathogenicity in humans. Displays good bioremediation capacity against crude oil, naphthalene and xylene (Xu, Deng, Huang, & Song, 2014). However, there isn’t relevant literature about the activity of this Family in the soil-plant relationship.

*Oxalobacteraceae* Family belongs to the Order *Burkholderiales* known for biodegradative capacities and antagonistic properties towards multiple soil-borne fungal pathogens (Benítez & Gardener, 2009; Chebotar et al., 2015; Kumar et al., 2017). *Oxalobacteraceae* in general
were seldom specifically examined in rhizosphere studies (Dohrmann & Tebbe, 2005; Green, Michel, Hadar, & Minz, 2007; Ofek, Hadar, & Minz, 2012).

*Paenibacillaceae* Family promotes plant growth by nitrogen fixation, phosphate-solubilization, iron acquisition, phytohormone production and biocontrol (Grady, MacDonald, Liu, Richman, & Yuan, 2016).

*Pseudonocardiaeae* Family plays an important role in the degradation of hemicellulose, cellulose, and chitin substances,(Yeager et al., 2017) being mostly endophytes (Qin et al., 2015).

About the *Sneathiellacea* family, there isn't relevant literature related to the activity of this Family in the soil-plant relationship. Because this Family belongs to the Class *Alphaproteobacteria*, possibly it lives symbiotically in the roots of leguminous plants. *Sphingomonadaceae* Family is a legume symbiont (Beattle, 2007), the establishment of this Family in the rhizosphere has been shown to be promoted by plant root exudates (Haichar et al., 2008).

*Streptosporangiaceae* Family was found on soil (Kudo, Itoh, Miyadoh, Shomura, & Seino, 1993). The role of member of this bacterial Family in the soil-plant relationship remains elusive.