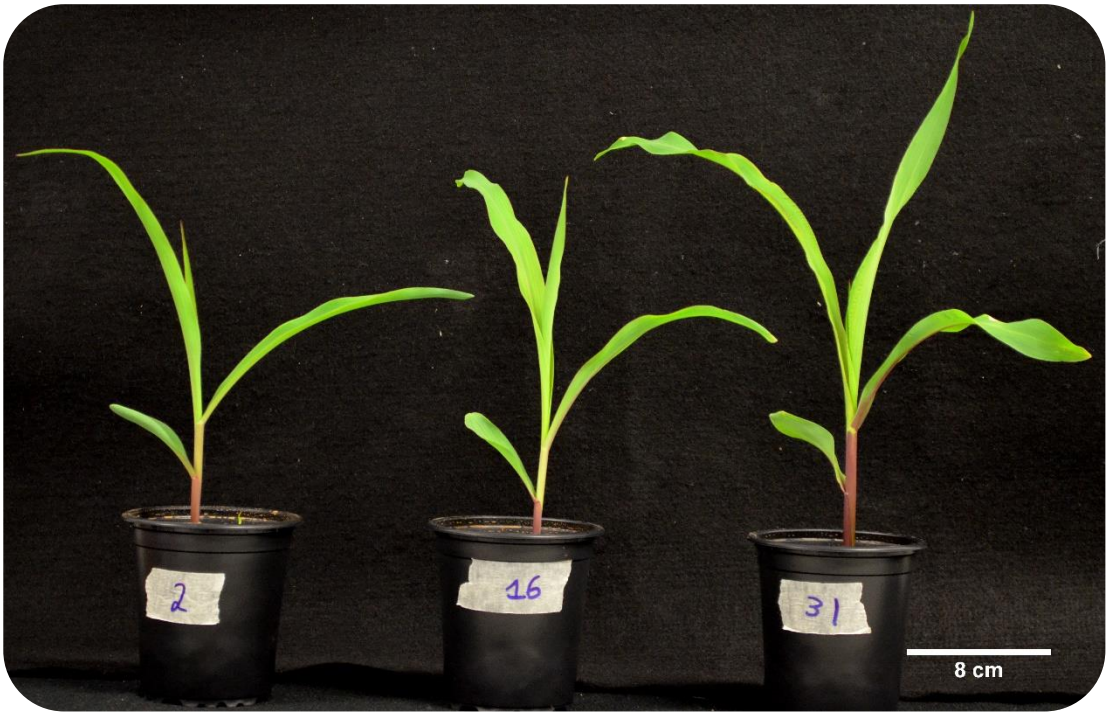
2023

The role of soil microbiota and environmental cues on the regulation of maize development

Valéria dos Santos Custódio



Dissertation presented to obtain the **Ph.D degree in Molecular Bioscience**

PhD programme PGCD and MolBioS

Oeiras, July, 2023

The role of soil microbiota and environmental cues on the regulation of maize development

Valéria dos Santos Custódio

Thesis submitted as partial fulfilment of the requirements for the degree of Doctor (Ph.D) in Molecular Biosciences (by the Universidade NOVA de Lisboa)

Oeiras, July 2023

PhD Program molecular biosciences

PGCD
PROGRAMA DE PÓS-GRADUAÇÃO
CIÊNCIA PARA O DESENVOLVIMENTO

ITQB nova

This work was performed at:

Plant Microbiome and Nutrition Laboratory
University of Nottingham, Plant Science, Sutton Bonington Campus, Leicestershire, LE12 5RD, UK

Plant Functional Genomics Laboratory
Genomic of Plant Stress Unit
Instituto de Tecnologia Química e Biológica António Xavier – Universidade Nova de Lisboa
Av. da República, Quinta do Marquês, 2780-157 Oeiras
Portugal

Plant Nutrition Laboratory
University of Nottingham, Plant Science, Sutton Bonington Campus, Leicestershire, LE12 5RD, UK

Instituto Nacional de Investigação e Desenvolvimento Agrário (INIDA)
CP 84, 7600 Praia, Cabo Verde

Supervisors:

Dr. Gabriel Castrillo – Plant microbiome and Nutrition Laboratory

Prof. Dr. M. Margarida Oliveira - Plant Functional Genomics Laboratory

Prof. Dr. Martin R. Broadley – Plant Nutrition Laboratory

Thesis committee:

Dr. Isabel Abreu – ITQB NOVA

Prof. Dr. Corina Carranca - INIAV

This work has been financially supported by “Fundação para a Ciência e a Tecnologia” through PhD fellowship SFRH/BD/128384/2017 awarded to Valéria dos Santos Custódio and the R&D Unit Bioresources for Sustainability GREEN-IT (UID/Multi/04551/2013, UIDB/04551/2020) and by the Future Food Beacon – University of Nottingham, BBSRC-NSF (no. BB/V011294/1), and by Fundação Calouste Gulbenkian (n.: 146400)

À minha família,

Acknowledgments

This thesis would not have been possible without the precious contribution and support of my supervisors, collaborators, colleagues and the care of my friends and family. Therefore, I want to express my gratitude to those who have contributed to my research and personal development during these years.

To my father, Osvaldo, and Mama “prossora”, a deep thank you for always being present despite the distance. All I’ve achieved so far is by wanting to be like you. To my uncle, Valério, that wherever he is, his support and affection were always unconditional.

To Gabriel Castrillo, a supervisor, a mentor, and more importantly, a very good friend. My special thank you and gratitude for the opportunity to be at your lab, where your vision and approach to science have shown me that dedication, discipline, transparency, audacity and creativity are important values to achieve excellence in science. I would like to thank you for all the support and for always being present in the good and, most important, in the bad moments to help me thrive throughout the challenges. I hope that one day, I will be a brilliant scientist as you are.

To Margarida Oliveira for accepting me in her lab and for all the scientific support. For always having time to discuss results with me and for your tremendous attention to detail. Your enthusiasm, work capacity and care have shaped me into a better professional and personal version. Thank you so much for everything.

To Martin Broadley for not only accepting me as a PhD student but also for receiving me at your lab. For the good feedback and for always understanding what I want to achieve, and give me space and good advice to follow my goals.

To Isai Salas-González, for not only teaching me how to analyse all this complex data but also for getting my back when data analysis was too much for me to handle. A special thank you also for always having the appropriate advice to help me find my way.

To Paulina Flis, for being an excellent chemical analyst that I always trust with my ionome analysis. Notably, a special thank you for your support and friendship. I learned with you every day how to be a good person and a good friend. Thank you for taking care of my mental health

during the pandemic with the amazing virtual Thursdays tea and the delight conversations.

To my best friends Carmen Garcia, Claudete Neves and Claudinete Neves for all your support and, more important, for embracing and loving my good, bad and villain beings. To my friend Joana 'Companheira' Furtado for always being present since we start our adventure in Cantanhede and for the beautiful walks in London and for always receiving me at your and Quentin's place with warmth and love. To my friends, Alexandre 'Petxas' Fernandes, Luis 'Suru' Cabral, Amaro 'Djueka' Rocha, Raquel Nadine, Suzanna Hopffer Andrade and Yamili Hernandez for always being present despite the distance and for always having time for me when I'm at home. To Cristihan González, for all the amazing dinners and *tarta de Santiago* that made my stay at Kegworth sweeter and for using your skills to help us have an amazing figure arrangement.

To my PGCD friends, Joel Amaque and Irina Fonseca with whom I have started this journey. Your support and friendship helped me during the long hours of work. Even when sometimes I needed to do some coding at 10 p.m. (Sorry! Joel, hehehhe). To Elves Heleno Duarte and his beloved wife, Eneida Mascarenhas, for all the conversations, discussions and for bringing Cabo Verde to me in UK. Finally, to my PGCD Silva-Melo family, Fredilson Melo and Eunice Silva, for all the support with the figures design and thesis writing, for your friendship, and for taking care of my social life in Oeiras.

To Ângela Moreno and Régla Amórós, for all the support during the work at INIDA. To Dona Ivete, Emanuel 'Lapa' Vaz, Wando, Nelito, Nelson and all the INIDA staff that helped me during all the field work.

To my friends, Yiqun Gao and Peiying Xue, and my colleagues, Guillem Reyt, Riccardo Fusi, Niokhor Bakhoum, and Elina Chrysantheou, for their friendship, support and dedication. To Lolita Wilson and Abdul Mossa for all your help and friendship.

To Elton Semedo, Louisa-Beth Pearce-Mooney, João Sousa, and Christalla Mannoli, for being the best undergraduate that I could wish for. You have been so important to this work by helping me and pushing me to improve every day. I could not be more proud of all of you.

To my *madrinha* Sónia Lima and my cousins Jéssica Alfama and Elber Alfama for always caring for me when I'm home. To my cousins, Inês Direito and Miguel Direito for all the support and friendship. To my friend Felipe Direito for the friendship and for taking care of my cultural life in the UK. To all my brothers and sisters who have been so patient during these years and for always pampering and taking care of me.

And, of course, a big thank you to everyone in Plant Sciences and Glasshouse team, especially to Hayley Newman, Nick Leftley, Zoe Philips, Sue Flint and Mark Merchant for all their support that makes my work go as smoothly as possible.

To my colleagues at Kegworth, Carolina Maia and Matthew Kent, for all the good moments and for taking care of my social life.

To all my lab colleagues in Portugal, especially Helena Sapeta, for all the advice and caring about me, and always trying to help.

To everyone here, I am truly thankful for the kindness and generosity I have experienced.

Valéria Custódio

The author of this thesis, Valéria dos Santos Custódio, hereby declares to have actively participated in the following publications:

Custódio, Valéria, Mathieu Gonin, George Stabl, Niokhor Bakhoum, M. Margarida Oliveira, Caroline Gutjahr & Gabriel Castrillo. 2022. Sculpting the soil microbiota. *The Plant Journal*, *109*(3), 508-522.

Valéria Custódio participated in the design of the original outline and wrote the Review. This manuscript includes the work described in Chapter I.

Dindas, Julian, Thomas A. DeFalco, Gang Yu, Lu Zhang, Pascale David, Marta Bjornson, Marie Christine Thibaud, **Valéria Custódio**, Gabriel Castrillo, Laurent Nussaume, Alberto P. Macho & Cyril Zipfel. 2022. "Direct Inhibition of Phosphate Transport by Immune Signaling in Arabidopsis." *Current Biology* 32(2):488-495.

Valéria Custódio analyzed the microbiome composition of phosphate mutants (plant growth, DNA extraction, microbiome sequencing).

Salas-González, Isai, Guilhem Reyt, Paulina Flis, **Valéria Custódio**, David Gopaulchan, Niokhor Bakhoum, Tristan P. Dew, Kiran Suresh, Rochus Benni Franke, Jeffery L. Dangl, David E. Salt & Gabriel Castrillo. 2021. "Coordination between Microbiota and Root Endodermis Supports Plant Mineral Nutrient Homeostasis." *Science* 371(6525).

Valéria Custódio conducted the soil collection, SynCom preparation, harvesting, DNA extraction, and library preparation. This manuscript describes all the methods used in this thesis.

List of Abbreviations

ABA	abscisic acid
ACC	1-aminocyclopropane-1-carboxylate
ANOVA	Analysis of Variance
ASV	Amplicon sequence variants
CGIAR-CSI	Consultative Group for International Agricultural Research Consortium for Spatial Information
CAP	Canonical analysis of Principal coordinates
cDNA	complementary DNA
CDS	coding sequence
CZ	Central zone
DADA2	Divisive Amplicon Denoising Algorithm
DAG	days after germination
DAS	days after sowing
DEG	differentially expressed genes
DNA	deoxyribonucleic acid
DTPA	Diethylenetriaminepentaacetic acid
EDTA	Ethylenediamine tetraacetic acid
ET_c	Water requirement
ET_0	Reference evapotranspiration
EtOH	Ethanol
FC	fold change
g	gram
GLM	Generalized linear model
GO	gene ontology
ha	Hectare
ICP-MS	Inductively Coupled Plasma Mass Spectrometry
INIDA	Instituto Nacional de Investigação e Desenvolvimento Agrário
JA	jasmonic acid
K_c	Crop coefficient
KEGG	Kyoto Encyclopedia of genes and genomes
L	Litre
LB	Luria Broth
mRNA	messenger Ribonucleic Acid

NaOAc	Sodium acetate
padj	adjusted p-value
PCA	principal component analysis
PCo	Principal coordinates
PCR	polymerase chain reaction
PEG	polyethylene glycol
PERMANOVA	Permutational analysis of variance
PZ	Peripheral zone
R2A	Reasoner's 2A agar
RF	Random Forest
RNA	ribonucleic acid
rRNA	Ribosomal RNA
RZ	Rib zone
SA	salicylic acid
SAM	Shoot apical meristem
SWC	Soil water content
SynCom	Synthetic community
TAW	Total available water
TRIS-HCl	Tris (hydroxymethyl) aminomethane hydrochloride
v	volume
w	weight
WCA	West and Central Africa
WT	wild type

Summary

As a result of climate change, the amount of greenhouse gases in the atmosphere is changing, temperatures are rising, precipitation patterns are changing, and extreme weather events are occurring more frequently. These changes accounted for 25% of average yield losses for low-latitude maize. Several strategies have been developed to tackle the reduction of productivity, such as altering planting and harvesting time, sowing crops with a short life cycle, applying crop rotation and irrigation techniques, and introducing variation in cropping systems. Besides these strategies, a knowledge-driven approach to manage the soil microbiota rather than merely applying microorganisms has been described as promising to sustainably increase crop yield.

Taking this into consideration, we have designed a holistic approach that combines the characterization of a collection of agricultural soils, soil bacterial composition, plant microbiota, and nutrition availability to understand how plant mechanisms and microbiota coordinate to ensure the adaptative growth of leaves in the model plant *Zea mays* (maize).

For two years, we monitored free-living soil bacteria communities, soil mineral content, and climatic conditions in 24 agricultural soils in the Sahel region (Cabo Verde). We observed strong structuring of the free-living bacteria according to the soil mineral content but not with the climatic conditions. Since plants recruit their symbionts from the surrounding free-living microbes, we have further evaluated the factors driving microbiota assembly in root and maize single leaves in three different soils. We found that individual maize leaves harbour distinct ionome and microbiota compositions. Then we designed a synthetic community (SynCom) to evaluate the effects of bacteria on single-leaf growth. We found that leaves respond differentially to the SynCom by suppressing defence genes.

Collectively, our work will help to understand the biotic and abiotic cues regulating maize development and will help to design microbial synthetic communities with a predictable effect on the plant host. This will support the development of more efficient and environmentally-friendly agricultural practices.

Resumo

Como resultado das alterações climáticas, a quantidade de gases com efeito estufa na atmosfera tem aumentado, a temperatura tem subido, os padrões de precipitação têm-se alterado e os eventos climáticos extremos têm ocorrido com maior frequência. Estas mudanças justificam, em média, 25% das perdas de rendimento das culturas de milho de baixa latitude. Para combater a redução da produtividade, foram desenvolvidas várias estratégias, como a alteração da altura da sementeira e da colheita, o uso de variedades de ciclo de vida curto, o recurso a técnicas de rotação e irrigação e outros ajustes nos sistemas de cultivo. Além destas estratégias, a modelação do microbiota do solo, em abordagens baseadas no conhecimento em vez da aplicação cega de microrganismos, tem sido referida como uma estratégia promissora para aumentar, de modo sustentável, a produtividade das culturas.

Posto isto, neste projecto adoptámos uma abordagem holística que combina a caracterização de uma coleção de solos agrícolas, da composição bacteriana do solo, do microbiota das plantas e da nutrição do solo, de modo a entender como os mecanismos das plantas e o microbiota se coordenam para garantir o crescimento adaptativo das folhas em *Zea mays* (milho).

Durante dois anos, monitorizámos as comunidades bacterianas do solo, a composição mineral do solo e as condições climáticas em 24 solos agrícolas na região do Sahel (Cabo Verde). Observámos uma forte estruturação das bactérias do solo de acordo com o conteúdo mineral do solo, mas não em função das condições climáticas. Dado que as plantas recrutam a comunidade bacteriana a partir do solo circundante, também avaliámos quais os factores que promovem o recrutamento microbiano em raizes e folhas de milho, cultivado em três solos diferentes. Descobrimos que folhas individuais de milho abrigam composições distintas de ionoma e de microbiota. Em seguida, construímos comunidades sintéticas (SynCom) para avaliar os efeitos das bactérias no crescimento das folhas. Os dados obtidos revelaram que a diferente resposta das folhas ao SynCom a que a planta é exposta se deve à supressão de genes de defesa da folha.

Globalmente, este trabalho é um importante contributo para a compreensão dos sinais bióticos e abióticos que regulam o desenvolvimento do milho, ajudando assim a desenhar comunidades microbianas com efeito previsível na planta, em suporte de práticas agrícolas mais eficientes e ecológicas.

Table of Contents

Chapter I

Introduction, hypothesis and objectives

This chapter was adapted from:

Custódio, V., Gonin, M., Stabl, G., Bakhoum, N., Oliveira, M.M., Gutjahr, C. and Castrillo, G. (2022), Sculpting the soil microbiota. *Plant Journal*, 109: 508-522. https://doi.org/10.1111/tpj.15568

Author's contribution

Custódio, V., Gonin, M., Stabl, G., Bakhoum, N., Oliveira, M.M., Gutjahr, C. and Castrillo, G. designed the original outline for the review. Stabl, G made the figures and legends, with input from the rest of the authors. All authors contributed to the text and corrected the different versions of the article.

1 Introduction

Demands on agriculture and natural resources are increasing unprecedentedly due to the exponential population growth (Godfray *et al.*, 2010). Additionally, around one billion people are chronically malnourished, suffering from acute hunger (insufficient caloric supply) or hidden hunger (insufficient minerals and vitamins), with severe impacts on health and life quality. Furthermore, agricultural practices have increasingly degraded the ecosystem, especially soil health, which is perceived as directly related to human nutrition (Blum *et al.*, 2019). Fundamentally, the soil health concept is the idea that soil is a living ecosystem and the well-being of soil is essential for achieving ecosystem services, including high-quality air and water, promoting diverse biotic and microbial community structure, supporting a high level of crop productivity and promoting human health (Blum *et al.*, 2019). The degradation of soil health leads to a dramatic loss of biodiversity and organic matter and to greenhouse gas emissions, with concomitant negative consequences for food production and climate change (Borrelli *et al.*, 2020). As a natural medium supporting plant growth and development, the soil is essential to understanding the abiotic and biotic factors that modulate plant performance and ecological balance, such as its mineral composition, the organisms living within it, and how they cross-interact.

1.1 Soil as a living ecosystem

It has been estimated that 1 g of soil harbours approximately 10^8–10^{10} microbial cells (Zhang et al., 2017), among other more complex organisms. These complex communities consist of hundreds to thousands of species of bacteria, fungi, oomycetes, archaea, viruses and nematodes and vary in diversity and composition across different soil types depending on soil properties. In turn, microbial activity can influence the soil's chemical, physical, and biological characteristics. Any loss of balance among factors determining soil health through agricultural intensification reduces microbial diversity compared with native soils (Roesch *et al.*, 2007). Although we still lack a complete understanding of the ecological mechanisms in soil, it appears that 80–90% of soil processes are mediated by microorganisms (Benbi and Nieder, 2003). They are involved in

ecological interactions that impact the soil structure, the recycling of nutrients, the decomposition of organic matter, the release of mineral nutrients, or the complementation of plant functions (Figure 1).

Soil structure results from a structural cluster of soil particles into aggregates that form a vast range of physicochemical niches, which provide distinct microhabitats for diverse microbial communities (Fox *et al.*, 2018). During soil aggregate formation, microbes present in the surrounding bulk soil are enclosed within the structures of the aggregates. Once formed, the soil aggregates constitute an isolated environment that significantly differs from the surrounding soil (Ebrahimi and Or, 2016; Bocking and Blyth, 2018) (Figure 1D). In response to changing environmental conditions, soil aggregates can be restructured, releasing the enclosed microbial communities and enabling the colonization of new soil niches (Cania *et al.*, 2019). Prominent examples of environmental perturbations that modulate the spatial isolation of microorganisms in soil aggregates in nature are floodings or intense rainfalls. With the resulting soil water saturation, microorganism-containing in aggregates can be dispersed by water movement reaching new soil areas and resources (Dechesne *et al.*, 2010; Ebrahimi and Or, 2015).

During drought events, the water content in soil pores is dramatically reduced, affecting the connectivity among soil areas. This lack of connectivity results in functionally disconnected resource islands across the soil structure, with low organic carbon decomposition and respiration rates (Canarini *et al.*, 2017; Schimel, 2018). These changes are associated with compositional adjustments in microbial populations. As the soil dries, the bacterial abundance changes (Barnard *et al.*, 2015). However, many bacteria can survive these changes by, for example, entering a dormancy state until more favourable conditions arise and growth is resumed (Zhu *et al.*, 2019).

Favourable conditions arise during the rainfall, where resource islands are reconnected as soil pores are filled with water. Furthermore, rewetting promotes a microbial activity pulse, leading to a burst of respiration and nitrification, known as the Birch effect (Birch, 1958) (Figure 1b).

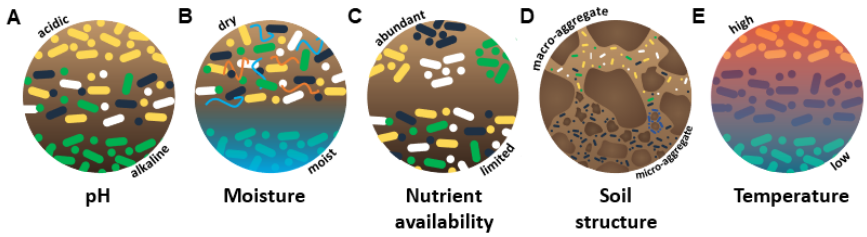


Figure 1: Abiotic factors shape soil microbial communities. (a) Soil microbial communities change according to pH. Commonly, microbial diversity is restricted at acidic and alkaline soil pH levels and is increased at moderate soil pH levels (Cao *et al*., 2016; Thompson *et al*., 2017; Wu *et al*., 2017). **(b)** Soil moisture affects the microbial community composition. The brown colour indicates dry soil and the cyan indicates moist soil. Under dry conditions, the soil harbours a broad variety of bacteria. In moist conditions, the composition of bacteria changes to species favoured by high humidity. **(c)** Nutrient availability influences interactions among microbes and abundance. **(d)** Soil structure is determined by clusters of mineral particles (macro- and micro-aggregates). Once formed, the soil aggregates provide distinct microhabitats for diverse microbial communities. **(e)** Soil temperature influences bacterial community composition, as different bacteria grow optimally at different soil temperatures. Red indicates high and blue indicates low temperatures, respectively. For all panels, the differently coloured bacteria indicate different bacterial genotypes (taxa, species, strains).

The composition of the soil microbiota is often modulated by the availability of nutrients in the soil (Fierer *et al.*, 2012) (Figure 1C). Fluctuations in nutrient composition and concentrations lead to complex metabolic interactions within microbial communities, resulting in a change in soil microbial composition (Vetsigian *et al.*, 2011; Leff *et al.*, 2015; Velez *et al.*, 2018). Under *in vitro* conditions, nutrient availability changes the spatial structure and genetic diversity of microbial colonies (Mitri *et al.*, 2016). In experimental mixed-genotype colonies of *Pseudomonas aeruginosa* grown with abundant nutrient resources, different bacterial lineages remained well mixed but lost diversity as the colony expanded (Mitri *et al.*, 2016). When colonies were grown under low-nutrient conditions, they remained more structured into areas dominated by different genotypes, but under these conditions, spatial expansion led to a loss of diversity (Mitri *et al.*, 2016).

The availability of water and nutrients in soil limits agriculture in many world areas. Understanding how microbes cope with extremes in soil water and nutrient availability may help improve microbiota-based agricultural practices, improving our capacity to address current and future challenges in food production while protecting the environment.

1.2 Santiago island as a model to understand environment-microbiota-plant interactions

Maize (*Zea mays* L.) is one of the most cultivated crops, widely consumed, and a multipurpose crop. It provides over 45% of food calories in Eastern and Southern Africa (ESA), 21% of food calories in West and Central Africa (WCA), occasionally reaching 50% in Lesotho, Malawi, and Zambia (Kumssa *et al.*, 2015). In West Africa, maize is critical for nutritional security. In 2020 the area harvested covered 15 million ha, largely in smallholder farming systems that produced more than 25 Mt of grains (FAOSTAT, 2020). Maize production is particularly challenging in the West region, due to low soil fertility, poor soil fertility management, and water deficits.

The Cabo Verde archipelago is part of West African countries, and maize and intercropping beans (*Vigna unguiculata, Phaseolus* spp., *Lablad purpureus,* and *Cajanus cajan*) are the major dryland crops, occupying over 80% of the arable land (Baptista, Ritsema, *et al.*, 2015). Maize production in Cabo Verde is mainly subsistence-based and rainfed. Climate change and altered weather variability threaten Cabo Verde's maize production due to its dependence on favourable weather conditions, such as sufficient rainfall and well-distributed rains (Monteiro et al., 2020).

In Cabo Verde, the dynamics of dominant climatic and meteorological factors do not favour the occurrence of rainfall. Typically, rainfall occurs between August to September (60% to 80%), and the amount of rain varies between islands depending on the prevalence of factors favourable to its occurrence, combined with topography and altitude (Sanchez-Moreno *et al.*, 2014a). For example, the mountainous island receives the highest amount of rainfall, especially Fogo (495 mm/year), Santiago (321 mm/year), Brava (268 mm/year) and Santo Antão (237 mm/year). The islands of Sal and Boavista receive a minimum amount of rain (60 mm and 68 mm per year, respectively), while the other islands receive an intermediate amount of rain per year: São Vicente (93 mm), São Nicolau (142 mm) and Maio (150 mm).

On Santiago Island, as in the Cabo Verde archipelago, there are two seasons: rainy and dry. The rainy season is generally marked by short but high-

intensity rainfall. In the rainy season, the temperature and humidity of the air are higher, and the temperature amplitude is low. The dry season lasts approximately nine months (November to July) and is marked by a significant absence of precipitation associated with high evaporation capacity (Varela-Lopes and Molion, 2014). This asymmetric spatial and temporal rainfall distribution affects agricultural productivity and soil health (Sanchez-Moreno *et al.*, 2014b). Due to the well-defined dry season with a small amount of rainfall and strong winds throughout the year, the soils are generally incipient and poorly developed (Moreno *et al.*, 2014). The soils in Santiago Island are relatively thin, ranging from 10 to 20 cm, and are dominated by coarse materials, which are poor in organic matter and nitrogen (Baptista, et al., 2015). Usually, at depth, they are rich in silts and clays. On this island, the soils are unstable due to permanent erosion caused by wind and water, aggravated by human action.

1.3 The plant microbiota

Complex multi-kingdom microbial communities colonize plants and confer numerous beneficial functions to the host, such as nutrient acquisition, pathogen protection or immune system modulation (Hiruma *et al.*, 2016; Castrillo *et al.*, 2017; Vogel *et al.*, 2021). Plants, in particular, produce exudates from the roots, which penetrate the rhizosphere, the soil portion surrounding the roots modulating the microbiota (Moitinho *et al.*, 2018; Mavrodi *et al.*, 2021). Reciprocally, the population of soil microbes that live in close association with the root, in the rhizosphere, on the root surface, or inside the root (the endosphere), engage with the plant in mutualistic interactions that benefit plant development.

In nature, plants are colonized by microbial communities consisting of hundreds to thousands of species of bacteria, archaea, fungi, oomycetes, viruses, protists, and other microbial eukaryotes, collectively known as plant microbiota (Hacquard *et al.*, 2017; Walters *et al.*, 2018). This microbiota originates mainly from the soil and air but is also transferred via vertical transmission (from parent to progeny) (Truyens *et al.*, 2015; Wagner *et al.*, 2016). In soil, the microbiota is mainly modulated by abiotic factors such as nutrients, moisture, soil structure, temperature and pH (Figure 2). It is broadly recognized that plants grown in different soils host distinct microbiota (Lundberg *et al.*, 2012; Vorholt, 2012) and that the different plant compartments, rhizosphere,

rhizoplane, endosphere, and phyllosphere, accommodate diverse microbial populations (Bulgarelli *et al.*, 2012; Leff *et al.*, 2015; Zarraonaindia *et al.*, 2015).

The rhizosphere is consistently enriched in bacterial microbiota members belonging to the phyla Actinobacteria, Acidobacteria, Firmicutes, Cyanobacteria, Bacteroidetes, Verrucomicrobia, and Proteobacteria. The rhizoplane, the external surface of the root, and also the root endosphere, sequentially filter out a vast majority of the microbes present in the rhizosphere, resulting in a microbiota community different and less diverse in these compartments (Ofek-Lalzar *et al.*, 2014; Chen *et al.*, 2020).

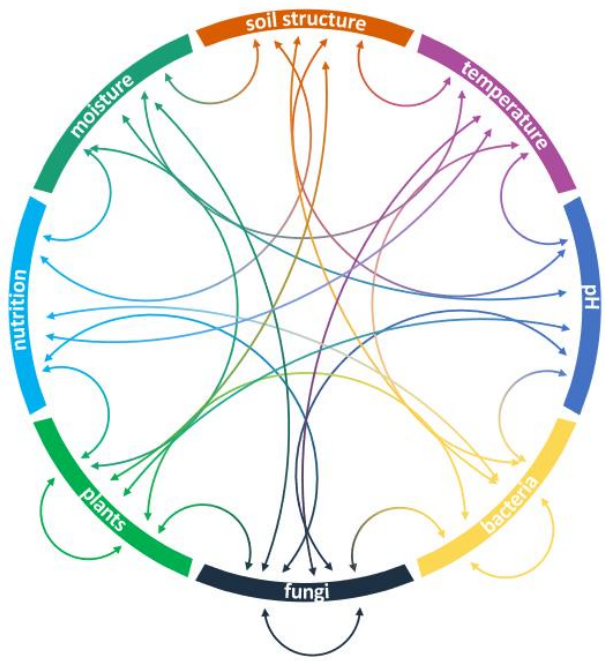


Figure 2: Schematic representation of the network of abiotic and biotic factors influencing the microbiota configuration in plant and at the soil interface. Arrows indicate interactions among abiotic and biotic factors. Bacteria, fungi and plants can influence each other in a community (indicated by the looped arrows in the figure)

In *Arabidopsis thaliana,* Actinobacteria, Bacteroidetes, Firmicutes, and Betaproteobacteria are enriched in these two root compartments (Bulgarelli *et al.*, 2012; Lundberg *et al.*, 2012), while Acidobacteria and Cyanobacteria members are depleted (Finkel *et al.*, 2019; Hou *et al.*, 2021). Microbes belonging to other groups of bacteria and kingdoms, also colonize and constitute the plant microbiota. For instance, symbiotic rhizobia colonizes legume root nodules and are hosted in specialized structures, the so-called symbiosomes. Arbuscular mycorrhizal fungi (AMF) form mutualistic symbioses with the host plants. They

grow into the root and form tree-shaped structures (arbuscules) inside the root cortex cells (Chiu and Paszkowski, 2019).

Not only plant compartments in close interaction with the soil are colonised by microbes, but also the phyllosphere, the aboveground part of the plant, accommodates a distinct microbiota population (Müller *et al.*, 2016; Chaudhry *et al.*, 2021). The bacterial microbiota in the leaves is dominated by Alphaproteobacteria and Gammaproteobacteria, where *Sphingomonas* and *Methylobacterium* (both belonging to Alphaproteobacteria) and *Pseudomonas* (Gammaproteobacteria) are the most dominant genera identified in several plant species such as maize, soybean, clover, and *Arabidopsis thaliana* (Delmotte *et al.*, 2009; Wallace *et al.*, 2018; Vogel *et al.*, 2021). Despite the striking structural differences observed in the different plant compartments, large-scale genome sequencing and re-colonization of germ-free plant experiments revealed a functional overlap between the root- and leaf-associated microbiota (Bai *et al.*, 2015).

1.4 Maize as a model to explore microbiota-leaf interactions

Maize is an annual grass that displays a complex root system architecture and a simple shoot architecture (Yu *et al.*, 2016; X., Liu *et al.*, 2021). In maize, shoot architecture, the three-dimensional structural arrangement of the above-ground plant body consists of a single erect stem, called culm, nodes, internodes and leaves (Figure 3A). Leaves are a predominant component of the shoot system, and they are uniquely simple and orderly structures (Strable, 2021). Each leaf is divided into three major regions, an upper strap-like blade separated by a joint-like ligule or auricles from the lower cylindrical sheath (Figure 3B).

The leaves of maize plants perform C4 photosynthesis and support the plant prior to reproductive development (Cao *et al.*, 2020). Apart from its impact on important agronomic traits such as maize grain productivity (Tian *et al.*, 2019), maize leaves have a long history as a model system from several studies ranging from epigenetics to development. In the field of cellular biology, maize leaves have been an excellent model to study the dynamics of growth regulatory networks (Nelissen *et al.*, 2015; de Vos *et al.*, 2020). This is because leaf development in maize is linear with cell division at the leaf basis, followed by cell

expansion and maturation (Pick *et al.*, 2011). Furthermore, the growth zone is relatively large allowing easy access to tissues at different positions (Nelissen *et al.*, 2018).

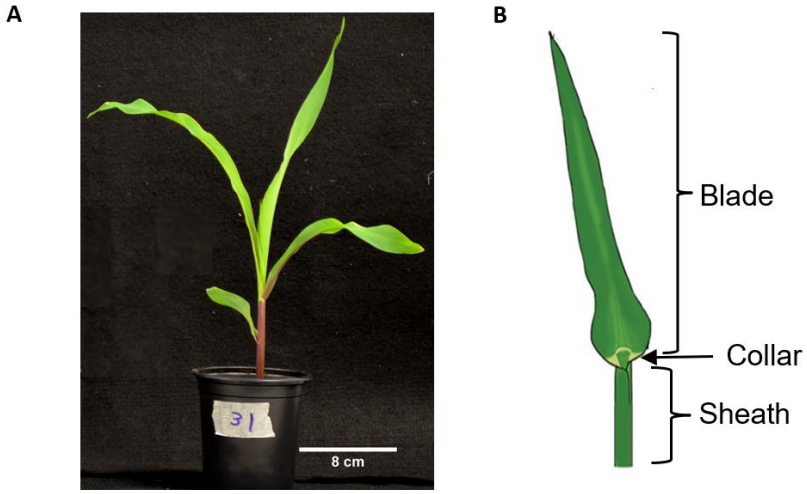


Figure 3: Maize leaf structure. A) Four week-old maize seedling growing in pot; (B) Identification of the different parts of the maize leaf (adapted from Richardson et al., 2021).

1.4.1 Leaf development

Organ development is a dynamic process of progressive specialization of cells and tissues that determines organ growth and the rate at which developmental processes occur (Kalve *et al.*, 2014). In plants, leaf development follows a general program that is flexible and adjusts according to species, developmental stages, and in response to environmental factors (Poethig, 2013; Bar *et al.*, 2015; Dubois *et al.*, 2017). This program is sequential and involves four main stages: (1) initiation, (2) morphogenesis, (3) expansion and (4) senescence (Dengler and Tsukaya, 2001).

Leaf initiation occurs at the shoot apical meristem (SAM) as simple bulges for dicotyledons and half-ring shapes for monocotyledons, which later grow along three-dimensional axes: adaxial-abaxial, proximal-distal, and mediolateral axes (Satterlee and Scanlon, 2019; Conklin *et al.*, 2019). Differences in leaf size and shape between different groups of plants arise at this stage.

The morphogenesis stage involves developmental processes in which the shape of the leaf develops as growth progresses (Ha *et al.*, 2007; Ha *et al.*, 2010). The beginning of the morphogenesis stage involves an alteration of growth direction so that the leaf begins to grow in the transversal plane and along its longitudinal axis (Tadege *et al.*, 2011). Usually, in dicotyledons, blade formation is restricted to the distal portion of the primordium, delimiting the more proximal petiole and leaf base regions (Ichihashi *et al.*, 2011). Conversely, monocotyledons follow an entirely different pattern of morphogenesis. After leaf initiation, leaf growth is mainly vertical, resulting in a short ensheathing leaf base and a strap-shaped distal region (Nelissen *et al.*, 2012; Johnston *et al.*, 2015).

The third stage of leaf development, the expansion, encompasses a more extended period and represents an increase in leaf surface area and volume (Pantin *et al.*, 2012). At this stage, the expansion occurs at similar rates across the leaf, but cells at the tip of the leaf are the first to cease expanding while the cells at the base continue the expansion until the leaf is fully expanded (Donnelly *et al.*, 1999; Beemster *et al.*, 2005).

Senescence is the last stage of leaf development preceding its death (Woo *et al.*, 2016). During senescence, leaf cells undergo significant changes in cell structure, metabolism and gene expression (Breeze *et al.*, 2011; Otegui, 2018). In this stage, the catabolism of chloroplast and macromolecules recycle nutrients that are supplied to sink sources such as roots, young leaves and seeds (Ishida *et al.*, 2014; Lornac *et al.*, 2020).

These four development stages are tightly regulated at the molecular level (Li *et al.*, 2020; Vercruysse *et al.*, 2020). However, the final size of the leaf is often adjusted based on the environmental conditions (Urano *et al.*, 2022; Clauw *et al.*, 2015). In general, during the early stages of leaf development, initiation and morphogenesis, leaves respond less to environmental perturbations (Zhao and Traas, 2020). At these early stages, leaf primordia are well protected within the bud, and the distribution of resources within the plants, such as water and nutrients, are channelled to the young leaves, even during times of stress. As leaves undergo the expansion stage and senescence, they become more vulnerable to environmental cues (Parent *et al.*, 2010). For instance, under moderate water scarcity, plants respond by reducing the leaf size through the

repression of cell proliferation and elongation (Baerenfaller *et al.*, 2012; Dubois *et al.*, 2017), while nutrient scarcity is known to accelerate leaf senescence (Lornac *et al.*, 2020).

1.4.2 Regulation of leaf growth

Leaf development originates from a pool of pluripotent cells located at the SAM (Barton, 2010). Functionally, SAM is divided into three zones, a central zone (CZ) containing pluripotent stem cells that organize and renew the meristem, a peripheral zone (PZ) from which leaves initiate, and a rib zone (RZ) that provides cells for the stem (Aichinger *et al.*, 2012; Tsuda and Hake, 2016).

At early developmental stages, the founder cells are recruited from PZ at the site of the incipient leaf primordium, and leaf initiation starts at auxin maxima generated by the auxin efflux carrier PINFORMED1 (PIN1) (Reinhardt *et al.*, 2003; Heisler *et al.*, 2005; Johnston *et al.*, 2015). Auxin maxima trigger activation of the auxin-responsive transcription factors MONOPTEROS (MP; ARF5), which down-regulates the *SHOOTMERISTEMLESS* (*STM*) gene (Chung *et al.*, 2019). STM encodes a class *I KNOTTED-LIKE HOMEOBOX* (*KNOX1*) transcription factor responsible for maintaining stem cell fate in the SAM (Yanai *et al.*, 2005; Jasinski *et al.*, 2005).

At early leaf initiation, the leaf primordium grows predominantly in the distal direction. During this process, cells in the primordium develop a polarity gradient along the adaxial-abaxial axis (Bowman and Floyd, 2008; Qi *et al.*, 2017). Cell identity in the adaxial domain depends on the expression of *REVOLUTA* (*REV*), *PHABULOSA* (*PHB*), and *PHAVOLUTA* (*PHV*) genes, which encode class III homeodomain-leucine zipper (HD-ZIPIII) proteins (McConnell *et al.*, 2001; Emery *et al.*, 2003; Caggiano *et al.*, 2017; Yu *et al.*, 2017). Abaxial domain identity depends on the *KANADI1* (*KAN1*) gene expression and members of the YABBY gene family (Kerstetter *et al.*, 2001; Eshed *et al.*, 2001; Stahle *et al.*, 2009; Sarojam *et al.*, 2010; Yu *et al.*, 2017). These two classes of genes produce signals that suppress each other's expression creating an asymmetric distribution of cells and tissue types that determine the final leaf shape (Reinhart *et al.*, 2013; Huang *et al.*, 2014; Xie *et al.*, 2015).

Subsequently, the adaxial-abaxial polarity establishment promotes leaf blade outgrowth through the mediolateral axis (Xu *et al.*, 2003). Expression of *WUSCHEL-RELATED HOMEOBOX1* (*WOX1*) and *PRESSED FLOWER* (*PRS*) in the middle leaf domain located between the adaxial and abaxial domains affects leaf blade outgrowth and margin-specific development (Nakata *et al.*, 2012; Guan *et al.*, 2017).

Finally, along with the termination of marginal meristem activity, cell proliferation and expansion occur at the entire blade, a process called intercalary growth, which leads to both distal and lateral leaf expansion (Nakata and Okada, 2013). Two classes of miRNA/transcription factor modules play dominant and antagonistic roles in sustaining cell proliferation and expansion during intercalary growth (Palatnik *et al.*, 2003; Liu *et al.*, 2009; Rodriguez *et al.*, 2010; Wang *et al.*, 2011). The module miR319, also called miRJAW in *Arabidopsis* and its target genes *TEOSINTE BRANCHED1/CYCLOIDEA/PROLIFERATING CELL FACTOR* (*TCP*) transcription factors promote the switch from cell proliferation to cell expansion by regulating the expression of miR396 (Cubas *et al.*, 1999; Ori *et al.*, 2007; Li *et al.*, 2012). The activation of miR396 by TCPs transcriptionally repressed several *GROWTH-REGULATING FACTORS* (*GRFs*) genes (Liu *et al.*, 2009; Wang *et al.*, 2011), that formed the second module with miR396. The GRFs are a family of transcription factors that delay the transition from proliferation to expansion in leaf development by modulating the levels of cyclins (Lee *et al.*, 2009; Debernardi *et al.*, 2014).

Along with regulation provided by miRNA-transcription factor modules, hormonal regulation also plays a role in the regulation of intercalary growth. Gibberellins (GAs) and brassinosteroids (BRs) contribute to leaf growth by enhancing cell proliferation and expansion (Achard *et al.*, 2009; Tong *et al.*, 2014). The overexpression of GA biosynthetic enzymes or constitutive activation of GA signalling leads to larger leaves (Eriksson *et al.*, 2000; Oikawa *et al.*, 2004; Fagoaga *et al.*, 2007; Voorend *et al.*, 2016). The positive effect of GAs on cell proliferation relies on the repression of cell-cycle inhibitors, such as KIP-RELATED PROTEIN 2 (KRP2) and SIAMESE (Achard *et al.*, 2009). On the other hand, blocking GA signalling decreases leaf size (Huang *et al.*, 1998; Achard *et al.*, 2009).

The overexpression of *BRASSINOSTEROID INSENSITIVE1*, encoding the BR receptor, or *DWARF4,* a BR biosynthetic gene, results in larger leaves (Choe *et al.*, 2001; Gonzalez *et al.*, 2010; Zhiponova *et al.*, 2013). Mutants deficient in BR biosynthesis or signalling display reduced leaf size and lower cyclinB (*CYCB*) expression, suggesting that BR controls the exit from mitosis (Zhiponova *et al.*, 2013). Moreover, BR signalling represses *PEAPOD1* (*PPD1*) and *PPD2*, which encode transcription factors that limit the proliferation of meristemoids (stem cell-like cells) (Gonzalez et al., 2015).

1.5 External cues affecting leaf growth

In nature, plants integrate different mechanisms to ensure the adaptative growth of roots and leaves in response to environmental changes (Martins *et al.*, 2017; Rao *et al.*, 2021; Peng *et al.*, 2021). This response occurs through complex regulatory networks that integrate internal signals (genetically determined) and external cues, such as nutrient availability, soil water content, light, pathogens, and beneficial bacteria (De Wit *et al.*, 2016; Vega *et al.*, 2019; Durán *et al.*, 2018; Gu *et al.*, 2020a). Recent evidence suggests that these external cues may differentially affect the growth of roots and leaves. For instance, magnesium (Mg) deficiency impaired root and leaf growth in potato (*Solanum tuberosum* L.) (Koch *et al.*, 2020), while iron and ammonium induce adventitious root formation in leaf cuttings of *Petunia hybrida* (Hilo *et al.*, 2017). However, in most of these studies, the plant-associated microbial community, the plant microbiota, that also contributes to host nutrition, development and growth (Hiruma *et al.*, 2016; Castrillo *et al.*, 2017; Harbort *et al.*, 2020) was not taken into consideration. Therefore, little is known about how plant organs ensure their growth when the microbiota is present and how the plants integrate the microbiota role in their responses to environmental stressors.

1.5.1 Nutrient-dependent leaf development (abiotic cues)

Plant development results from a combined action among nutrient availability, energy status, cell growth and expansion (Dobrenel *et al.*, 2016). Nutritionally, development is a demanding process where nutrient sensors often control developmental decisions (Liu *et al.*, 2017; Oldroyd and Leyser, 2020). Signals derived from different nutrients induce the expression of a wide variety of genes,

including transporter-, transceptor- and assimilation-associated genes, genes involved in phytohormone synthesis and response, and genes encoding regulatory proteins such as transcription factors and protein kinases (Hu *et al.*, 2019; Sakakibara *et al.*, 2006). For instance, nitrate (NO_3^-) has been extensively characterized in plants as a nutrient signal that modulates development by triggering complex transcriptional networks and by affecting phytohormone biosynthesis (Hu *et al.*, 2019; Ötvös *et al.*, 2021). The biosynthesis of cytokinins (CKs) is highly responsive to NO_3^- provision (Takei *et al.*, 2001; Takei *et al.*, 2002; Takei *et al.*, 2004). In Arabidopsis, NO_3^- specifically upregulates the expression of *ISOPENTENYLTRANSFERASE 3* (*IPT3*) that mediated the synthesis of isopentenyl (iP)-type CK in the root phloem. Subsequently, CYP735A (encoding a cytochrome P450 monooxygenase) converts iP-type CK into trans-zeatin (tZ)-type CK, the active zeatin form transported through the xylem to promote leaf expansion (Ko *et al.*, 2014; Zhang *et al.*, 2014). This tZ precursor also mediates the adaptation of SAM size and organogenesis rate to the availability of mineral nutrients by modulating the expression of WUSCHEL (Landrein *et al.*, 2018; Osugi *et al.*, 2017; F., Zhang *et al.*, 2017), a homeodomain transcription factor that regulates stem cell fates in SAM (Ma *et al.*, 2019; Plong *et al.*, 2021).

The four-stages of leaf development are tightly coordinated by miRNAs (Pulido and Laufs, 2010). Nutrient availability has been shown to alter the expression of more than 80 miRNA families (Chien *et al.*, 2017). For instance, nitrogen (N) supply alters the expression of miR160 and miR167 (Liang *et al.*, 2012). These two miRNAs regulate leaf initiation by targeting different members of AUXIN RESPONSE FAMILY (ARFs) required for the local auxin peaks in the meristems, and regular leaf pattern along the stem (phyllotaxis) (Mallory *et al.*, 2005).

Recently, the TARGET OF RAPAMYCIN (TOR) has been characterized as one of the essential pathways involved in the integration of nutritional signals and growth decisions (Dobrenel *et al.*, 2016; Busche *et al.*, 2021). TOR activity is induced by sugars such as sucrose and glucose by a largely unknown mechanism, and TOR activates meristems by inducing WUSCHEL transcription factors in the SAM (Xiong *et al.,* 2013; Pfeiffer *et al.*, 2016). This activity is impaired under nutrient scarcity or defects in nutrient assimilation. For example,

a decrease in sulphur assimilation due to mutations in the sulphite reductase gene diminishes TOR activity and growth through a reduction in sugar accumulation (Dong *et al.*, 2017).

1.5.2 Microbiota influence plant development (biotic factors)

The plant-associated microbiota can directly or indirectly provide benefits to plant growth and health through several mechanisms (Castrillo *et al.*, 2017; Vogel *et al.*, 2021a; Harbort *et al.*, 2020). The direct mechanisms are based on the production of molecules that either facilitate resource acquisition (Abadi and Sepehri, 2016; Kumar *et al.*, 2016; Verma *et al.*, 2021) or modulate plant hormone levels (Shi *et al.*, 2010; Spaepen *et al.*, 2014a). The indirect mechanisms involve the decrease of detrimental effects of plant-pathogen interactions (Ritpitakphong *et al.*, 2016; Helfrich *et al.*, 2018).

The mechanisms of plant growth promotion driven by the microbiota effect on nutrient acquisition have been extensively studied in the case of the rhizosphere microorganisms (Pérez-De-Luque *et al.*, 2017; Gu *et al.*, 2020b; Lidbury *et al.*, 2020). For example, nitrogen-fixing rhizobia and mycorrhizal fungi have been extensively reported as influencing plant nutrient status by providing plants with nutrients, increasing nutrient availability, or enhancing nutrient acquisition capacity in the soil (Smith and Smith, 2011; Udvardi and Poole, 2013; Schwember *et al.*, 2019; Chiu and Paszkowski, 2019). In addition, members of the microbiota other than rhizobia, increase the availability and uptake of phosphorus (P), zinc (Zn), and iron (Fe) (Goteti *et al.*, 2013; Vaid *et al.*, 2014; Pandey and Gupta, 2020; Chen *et al.*, 2021a) by releasing phosphatase enzymes or organic chelates in the rhizosphere (Sharma *et al.*, 2003; Richardson and Simpson, 2011; Jain *et al.*, 2020). For instance, strains belonging to the genera *Pseudomonas, Burkholderia, Paraburkholderia, Novosphingobium,* and *Ochrobactrum* increase leaf biomass and leaf nutrient uptake by affecting soil urease and acid phosphatase activities (Chen *et al.*, 2021b). This example reveals that, although these bacterial activities take place at the soil level, they have an impact on leaf growth and nutrient content.

Surveys of phyllosphere community composition have reported the presence of nitrogen-fixing bacteria and phosphate solubilizer members, but mechanistic

studies of plant-phyllosphere nutrient dynamics are generally lacking (Abadi *et al.*, 2020; Arun *et al.*, 2020; Abadi *et al.*, 2021).

Hormones such as auxins and cytokinins (CKs) produced by rhizobacteria, alter the root system architecture, thereby indirectly enhancing nutrient acquisition by the roots (Spaepen *et al.*, 2007; Spaepen *et al.*, 2014b; Bailly *et al.*, 2014; Zamioudis *et al.*, 2013). In most plants, auxins and CKs are crucial to plant development, interacting and overlapping functions in response to nutrients (Schaller *et al.*, 2015). There are many microorganisms isolated from the phyllosphere that are capable of synthesizing auxin, and these organisms either promote plant growth or cause plant diseases (Spaepen *et al.*, 2007). Microorganisms that produce auxin in the phyllosphere have only recently been studied, and evidence suggests they can increase plant productivity (Romero *et al.*, 2016). For instance, phylloplane bacteria isolated from *Jatropha curcas* significantly increased leaf length in maize seedlings, and the auxin indole acetic acid was significantly correlated with variation identified in these seedlings (Dubey *et al.*, 2017).

Although plants have a range of biochemical, molecular, and physiological mechanisms to cope with adverse environmental conditions, they often rely on their microbiota for survival and growth under abiotic stresses (Hartman and Tringe, 2019; Santos-Medellín *et al.*, 2021). An enzyme implicated in plant-growth promotion under various stress conditions is 1-aminocyclopropane-1-carboxylate (ACC) deaminase, which converts the ethylene precursor ACC to α-ketobutyrate and ammonia. Microorganisms with ACC deaminase activity divert ACC, reducing ethylene production via ACC oxidase by the plant, thereby alleviating ethylene-mediated inhibition of plant growth in response to various stress factors (Glick, 2014).

Furthermore, diverse members of microbiota associated with plants promote plant growth indirectly by protecting them from biotic stress (Mendes *et al.*, 2013; Berendsen *et al.*, 2012). For instance, a study in *Arabidopsis thaliana* shows that the root-associated bacteria modulate interkingdom interactions between bacterial, fungal, and oomycetal community, resulting in a balanced plant-microbiota interaction that promotes plant growth and survival against root-derived filamentous eukaryotes (Durán *et al.*, 2018). The mechanisms by which

plant-associated microorganisms protect against plant pathogens include competition for niches and nutrients, antibiosis, production of lytic enzymes, inhibition of pathogen virulence, and induction of plant-mediated resistance (Raaijmakers and Mazzola, 2012; Pieterse *et al.*, 2014). For example, the MYB72 transcription factor, known to be involved in the induction of systemic acquired resistance, has also been identified as crucial in plant adaptation to iron deficiency (Palmer *et al.,* 2013; Stringlis *et al*., 2018; Zamioudis *et al*., 2014). MYB72 regulates genes responsible for the biosynthesis of coumarins, which are prominent compounds in root exudates, playing an important role in iron uptake and assimilation (Fourcroy *et al.*, 2016; Harbort *et al.,* 2020; Schmid *et al*., 2014).

1.6 Methods to study microbiome-plant interaction mechanisms

1.6.1 Culture-independent approaches

In recent years, advances in high-throughput sequencing and significant cost reductions have enabled culture-independent studies of microbial communities in a wide range of plant hosts and environments (Bulgarelli *et al.*, 2012; Lundberg *et al.*, 2012; Edwards *et al.*, 2015; Delgado-Baquerizo *et al.*, 2018). Culture-independent methods use molecular strategies, especially those based on survey analysis of genes identified after polymerase chain reaction (PCR) (DeLong and Pace, 2001). By using this approach, a better understanding was achieved of the relative abundance of different phylogenetic groups within a community and their metabolic and physiological potential (Fitzpatrick *et al.*, 2018; Castellano-Hinojosa and Strauss, 2021).

The most common culture-independent methods to study plant-microbe interactions are based on microbial marker-gene amplicon profiling and microbial whole-genome shotgun metagenomics (Liu et al., 2021). The major marker genes used in amplicon sequencing include 18S rRNA and internal transcribed spacers (ITS) for eukaryotes and 16S ribosome RNA (rRNA) for prokaryotes (Turnbaugh *et al.*, 2007; Gilbert *et al.*, 2014). The 16S rRNA gene encodes the small subunit ribosomal RNA molecules and comprises nine variable regions (V1-V9) distributed throughout the highly conserved 16S sequence (Johnson *et al.*, 2019). Its extensive use as a microbial barcode gene is mainly due to two basic characteristics: 1) universal presence across bacteria and archaea kingdom, 2)

existence of conserved regions for which universal primers can be designed, and variable regions that allow differentiation (Raina *et al.*, 2019). However, this technique can only reach genus-level resolution.

Metagenomics captures the genetic sequence of a microbial community and not only extends taxonomic resolution to the species- or strain-level but also provides potential functional information (Sessitsch *et al.*, 2012; Bulgarelli *et al.*, 2015). For instance, a culture-independent analysis of a collection of plant-associated bacterial genomes isolated from different plant hosts identified bacterial genes involved in adaptation to plants, including genes associated with plant colonization, microbe-microbe competition, and host-microbe interactions (Levy *et al.*, 2017). While amplicon sequences are designed mainly for the purpose of studying the phylogenetic relationships of species, the species composition, and the biodiversity of the microbial community (Finkel *et al.*, 2020; Salas-González *et al.*, 2021; Hou *et al.*, 2021), metagenomic sequences also conduct in-depth research on genes and functions of a microbial community, such as pathway analysis, KEGG, and gene ontology (GO) (Kanehisa *et al.*, 2014; Najar *et al.*, 2020).

Currently, several sequencing platforms and standardized protocols are available for amplicon and metagenomic sequencing. In the field of 16S rRNA amplicon sequencing, Illumina sequencing platforms (MiSeq and HiSeq) have been widely used (Levy and Boone, 2019). This approach permits analysis of a sub-region of the 16S rRNA gene and taxonomically assigns the reads at the genus level (Delgado-Baquerizo *et al.*, 2021). Besides producing very high-quality sequences, Illumina sequencing platforms produce short read lengths (~300bp). This characteristic limited the use of this approach in the metagenomic analysis since the short reads cannot reliably reconstruct long reads because of uncertainties in mapping reads. Due to that, long-read sequencing platforms have become increasingly popular for metagenomic analysis. Recently, Oxford Nanopore Technologies (ONT) sequencing platforms have been combined with short-read techniques to generate high-quality genomes for metagenomic analysis (Daims *et al.*, 2015). On average, ONT generates 30- to 50-kb reads, although suffering from low raw read accuracy and higher per-read costs, as compared to the Illumina platform.

1.6.2 Culture-dependent approaches

Currently, research in plant-microbe interactions is reaching new dimensions by breaking down the complexity of the regulatory mechanisms between plants and microbes (Finkel *et al.*, 2020; Salas-González *et al.*, 2021). This has been mainly accomplished using culture-dependent approaches that allow the generation and study of representative taxonomic and functional collections of plant-associated microbes isolated from different habitats (Bai *et al.*, 2015; Levy *et al.*, 2017; Zhang *et al.*, 2019).

Since a large portion of the plant microbiome is unculturable, the generation of a comprehensive bacterial culture collection combines different approaches to increase the diversity of the collection (Bai et al., 2015; Zhang et al., 2019). These approaches include testing different growth media with various substrate compositions, modification of existing culture media and application of different cultivation methods (Bai *et al.*, 2015; Durán *et al.*, 2018).

The selection of the growth media is important to guarantee bacterial growth and collection diversity. For example, to isolate microbial members with specific culture requirements, several approaches based on 16S rRNA gene sequence or the metagenome can help to identify appropriate selective media. By using the 16S rRNA gene sequence, Oberhardt et al. (2015) developed a tool to predict the selective media required for bacterial isolation. Another tool developed to assist in media selection is BugBase database (https://bugbase.cs.umn.edu/), also based on 16S rRNA. BugBase helps to predict bacterial characteristics, such as aerobic or anaerobic growth, thus providing guidance regarding the types of cultivation conditions likely required (Ward *et al.*, 2017). Metagenomic data or bacterial genomes present a second source of information about bacterial carbon utilization and antibiotic resistance genes, which aids the design of isolation media to enrich specific bacteria (Lombard *et al.*, 2014; Jia *et al.*, 2017). Guided by metagenomic data, Kwak *et al.* (2018) used a marine broth medium containing kanamycin to successfully isolate 22 flavobacteria strains that confer wilt resistance in tomatoes (Kwak *et al.*, 2018).

The type of bacterial isolation system is also important for the culture collection. A taxonomically comprehensive microbial culture collection can be

built by using mainly two isolation procedures: colony picking and limiting dilution (Bai *et al.*, 2015; Durán *et al.*, 2018). In colony picking, microbial members are isolated from individual colonies grown on selective solid media based on the colony morphology, shape and colour (Lagier *et al.*, 2016; Lagier *et al.*, 2018). For instance, to study the role of salicylic acid role in root microbiome assembly, Lebeis *et al.*, (2015) used different dilutions of solid media to isolate the microbes that they later used to test the hypothesis of the involvement of this plant hormone in microbiome assembly.

In limiting dilution methods, the microbes are isolated in 96-well cell culture plates by successive dilutions in liquid media of the root and leaf microbiota samples (Zhang *et al.*, 2021). In this method, the bacterial suspension isolated from plant tissues is diluted and separated in the individual wells, assuming that microbial members are separated following a Poisson distribution (Goodman *et al.*, 2011). For example, Bai *et al.* (2015) established a pipeline for high-throughput bacterial cultivation and identification using two methods (Bai *et al.*, 2015). They first isolated bacteria from healthy plant roots or leaves using colony picking from agar plates and limiting dilutions in liquid media in 96-well plates. Then, they performed a two-step barcoding sequencing for taxonomic classification. With this strategy, Bai *et al.* (2015) succeeded in characterizing 7943 bacterial isolates from roots and leaves of *Arabidopsis thaliana.* The combination of colony picking and limiting dilution allows the cultivation of more than 50% of the bacterial taxa present in the Arabidopsis root and leaf microbiomes (Durán *et al.*, 2018).

Future cultivation efforts should focus on strategies to isolate ‘difficult-to-grow’ taxa, such as Acidobacteria, Choroflexi, Delta-proteobacteria and anaerobic taxa, which ubiquitously exist in root and leaf microbiota but are under-represented in culture collections (Lebeis *et al.*, 2015; Bai *et al.*, 2015; Durán *et al.*, 2018; Niu *et al.*, 2017). These taxa are particularly interesting given the fact that they are very abundant in the plant natural community (Bulgarelli *et al.*, 2012; Finkel *et al.*, 2019).

1.6.3 Reductionist Microbial Synthetic Communities (SynCom) approach

A particularly important aspect of plant microbiome research is understanding the mechanisms underlying the assembly and function of host-associated microbial communities (Durán *et al.*, 2018; Salas-González *et al.*, 2021; Hou *et al.*, 2021). Addressing such a fundamental principle is challenging due to the complexity and dynamics of plant-microbe interactions. These complex interactions can be simplified by constructing synthetic microbial communities (SynCom).

A synthetic community, or constructed microbiome, is composed of selected strains of microorganisms, which can be applied to plants for studying various aspects of plant-microbe interactions (Castrillo *et al.*, 2017; Durán *et al.*, 2018; Finkel *et al.*, 2019). There are many advantages to this system, including the fact that its components can be studied in detail under controlled and reproducible conditions, allowing causal links to be established between genotypes and phenotypes (Salas-González *et al.*, 2021; Finkel *et al.*, 2020). For instance, in a recent study that used a SynCom of 41 taxonomically diverse bacteria isolated from roots and shoots of *Arabidopsis*, the authors showed that the SynCom had the capacity to change the composition of the root endodermal diffusion barriers (Salas-González *et al.*, 2021). Another advantage of the SynCom approach is that organisms can be added, eliminated, or substituted at the strain level. In addition, changes can be implemented not only at the organism level but also at the functional level, where individual functions can be specifically removed or added. For example, by inoculating seedlings with a 185-member bacterial SynCom, Finkel *et al.* (2020) identified interactions between microorganisms that determine root development. Remarkably, the authors identified a single bacterial genus (*Variovorax*) involved in reversing the severe inhibition of root growth induced by the SynCom. By removing an auxin-degradation operon, they could demonstrate that *Variovorax* manipulates the plant hormone levels.

These examples could be studied due to the lower complexity of the SynCom as compared with natural populations, although it also represents one of the limitations of this approach. Simplified experimental systems do not represent the full breadth of environmental heterogeneity, where many varying factors act

simultaneously, but they are a prerequisite for experimentation as a means to uncover causality through target manipulation. Even if these reduced-complexity systems do not exactly mimic nature, they allow for the recapitulation of microbiome-mediated phenotypes that can be further tested with alternative communities and/or be used to search for structure/function correlations in environmental samples to demonstrate general principles beyond a specific experimental setup.

2 Hypothesis, objectives, and thesis outline

In Cabo Verde, maize (*Zea mays* L.) is the major dryland crop, and its production suffers from two major constraints: water scarcity and nutrient-deficient soils. Under these conditions, plant development is heavily impaired. Recent evidence has shown that soil microbiota has a critical role in promoting plant growth in challenging environments. Importantly, the plant microbiota, the population of microbes that live on and in the plant, is essential for maintaining plant health in nature. However, how these two elements, the plant development regulation and the microbiota, coordinate to improve plant growth remains poorly understood.

In this project, we planned a holistic approach, combining the characterization of a collection of agricultural soils, soil microbiome composition, plant microbiota and nutrition to disentangle mechanisms mediating the development of the model plant *Zea mays* (maize) grown under realistic conditions.

Since a wide range of microbiota that inhabit plant tissues are largely derived from soil, we aimed to define the principles underpinning microbiota adaptation to changing environmental conditions in the soil and in association with the plant, as a way to help design microbial synthetic communities with a predictable effect on the plant host, as this can open new avenues in the development of more efficient and sustainable agricultural systems. To achieve our aims, we have identified three main objectives:

a) **To characterise the climate and soil properties of maize producer's farms**
b) **To assess how the soil mineral content and microbiota modulation affect plant performance**

c) To elucidate microbiota functions that alter maize development under nutrient-deficiency conditions

To address the above objectives, I have organized my thesis work into four additional chapters as described below:

Chapter II – The role of soil mineral content and soil microbiota in maize productivity

In this chapter, I used the Santiago Island (Cabo Verde) as a model to identify environmental factors and agricultural practices that modulate the soil mineral content and soil free-living microbiota and to which extent this modulation affects plant performance. I have identified 24 farmers that were almost fully dedicated to maize production. These farmers are distributed across 23 regions on the Island, representing eight of the nine municipalities of Santiago and extending from the north to the south and from the east to the west. Soil samples from each farm were analyzed in two independent experiments to determine the nutrient and the microbial composition. Climatic and soil properties data were added to the analysis to aid characterized the environment on the island. Finally, fieldwork was performed in two different locations on Santiago Island to understand the effect of variations in soil mineral content and microbiota on maize productivity.

Chapter III – The bacterial communities inhabiting individual leaves may contribute to their developmental program

In this chapter, I established maize leaf- and root-derived microbiota culture collections representing the majority of bacterial species that are reproducibly detectable by culture-independent community sequencing. To expand the genome-based understanding of plant-microbe interactions, I have also analysed the genome drafts of the maize collection isolates. Finally, I used defined bacterial communities and a gnotobiotic maize plant system to analyse the capacity of the isolates to colonize and resemble the natural microbiota pattern *in planta*.

Chapter IV – Individual leaf microbiota modulates the growth-defense trade-off to promote leaf growth in maize

In this chapter, I studied the SynCom effect on maize growth at the single-leaf resolution, going through leaves at different developmental stages. I found several transcription regulators and genes expressed at the maize leaf level that were affected by the microbiome.

Chapter V – Conclusion and future perspectives

In this final chapter, I discuss the main findings of the work described in the previous chapters and examine future steps to complement or follow up on this research.

3 References

Bai, Y., Müller, D.B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., Dombrowski, N., Munch, P.C., Spaepen, S., Remus-Emsermann, M., Huttel, B., McHardy, A.C., Vorholt, J.A., Schulze-Lefert, P. (2015) Functional overlap of the Arabidopsis leaf and root microbiota. *Nature*, **528**, 364–369.

Baptista, I., Fleskens, L., Ritsema, C.J., Querido, A., Tavares, J., Ferreira, A.J.D., Reis, E.A., Gomes, S. and Varela, A. (2015) Soil and Water Conservation Strategies in Cape Verde (Cabo Verde in Portuguese) and Their Impacts on Livelihoods : An Overview from the Ribeira Seca Watershed.

Baptista, I., Ritsema, C. and Geissen, V. (2015) Effect of Integrated Water-Nutrient Management Strategies on Soil Erosion Mediated Nutrient Loss and Crop Productivity in Cabo Verde Drylands P. C. Abhilash, ed. *PLoS One*, **10**, e0134244.

Bar, M., Ben-Herzel, O., Kohay, H., Shtein, I. and Ori, N. (2015) CLAUSA restricts tomato leaf morphogenesis and GOBLET expression. *The Plant Journal*, **83**, 888–902.

Barnard, R.L., Osborne, C.A. and Firestone, M.K. (2015) Changing precipitation pattern alters soil microbial community response to wet-up under a Mediterranean-type climate. *ISME Journal*.

Beemster, G.T.S., Veylder, L. de, Vercruysse, S., West, G., Rombaut, D., Van Hummelen, P., Galichet, Arnaud G., Gruissem, W., Inzé, D., Vuylsteke, M. . (2005) Genome-Wide Analysis of Gene Expression Profiles Associated with Cell Cycle Transitions in Growing Organs of Arabidopsis. *Plant Physiol*, **138**, 734–743.

Benbi, D.K. and Nieder, R. (2003) *Handbook of processes and modeling in the soil-plant system* R. Nieder and D. Benbi, eds., New York : Food Products Press : Haworth Reference Press.

Blum, W.E.H., Zechmeister-Boltenstern, S. and Keiblinger, K.M. (2019) Does Soil Contribute to the Human Gut Microbiome? *Microorganisms 2019, Vol. 7, Page 287*, **7**, 287.

Bocking, C.R. and Blyth, M.G. (2018) Oxygen uptake and denitrification in soil aggregates. *Acta Mech*, **229**, 595–612.

Borrelli, P., Robinson, D.A., Panagos, P., Lugato, E., Yang, J.E., Alewell, C., Wuepper, D., Montanarella, L. and Ballabio, C. (2020) Land use and climate change impacts on global soil erosion by water (2015-2070). *Proc Natl Acad Sci U S A*, **117**, 21994–22001.

Breeze, E., Harrison, E., McHattie, S., et al. (2011) High-Resolution Temporal Profiling of Transcripts during Arabidopsis Leaf Senescence Reveals a Distinct Chronology of Processes and Regulation. *Plant Cell*, **23**, 873–894.

Bulgarelli, D., Garrido-Oter, R., Münch, P.C., Weiman, A., Dröge, J., Pan, Y., McHardy, A.C. and Schulze-Lefert, P. (2015) Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe*, **17**, 392–403.

Bulgarelli, D., Rott, M., Schlaeppi, K., van Temaat, E., Ahmadinejad, N., Assenza, F., Rauf P., Huettel, B., Reinhardt, R., Schmelzer, E., Peplies, J., Gloeckner, F.O., Amann, R., Eickhorst, T., Schulze-Lefert, P. (2012) Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature*, **488**, 91–5.

Canarini, A., Kiær, L.P. and Dijkstra, F.A. (2017) Soil carbon loss regulated by drought intensity and available substrate: A meta-analysis. *Soil Biol Biochem*.

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

Cao, Y., Zeng, H., Ku, L., Ren, Z., Han, Y., Su, H., Dou, D., Liu, H., Dong, Y., Zhu, F., Li, T., Zhao, Q., Chen, Y. (2020) ZmIBH1-1 regulates plant architecture in maize. *J Exp Bot*, **71**, 2943–2955.

Castrillo, G., Teixeira, P.J.P.L., Paredes, S.H., Law, T.F., de Lorenzo, L., Feltcher, M.E., Finkel, O.M., Breakfield, N.W., Mieczkoswi, P., Jones, C.D., Paz-Ares, J., Dangl, J.L. (2017) Root microbiota drive direct integration of phosphate stress and immunity. *Nature*, **543**, 513–518.

Chaudhry, V., Runge, P., Sengupta, P., Doehlemann, G., Parker, J.E. and Kemen, E. (2021) Shaping the leaf microbiota: plant–microbe–microbe interactions. *J Exp Bot*, **72**, 36–56.

Chen, T., Nomura, K., Wang, X., Sohrabi, R., Xu, J., Lingya, Y., Paasch, B.C., Ma, L., Kremer, J., Cheng, Y., Zhang, L., Wang, N., Wang, E., Xin, X, He, S.Y. (2020) A plant genetic network for preventing dysbiosis in the phyllosphere. *Nature 2020 580:7805*, **580**, 653–657.

Chiu, C.H. and Paszkowski, U. (2019) Mechanisms and Impact of Symbiotic Phosphate Acquisition. *Cold Spring Harb Perspect Biol*, **11**, a034603.

Conklin, P.A., Strable, J., Li, S. and Scanlon, M.J. (2019) On the mechanisms of development in monocot and eudicot leaves. *New Phytologist*, **221**, 706–724.

Daims, H., Lebedeva, E. v., Pjevac, P., Han, P., Herbold, C., Albersten, M., Jehmlich, N., Palatinszky, M., Vierheilig, J., Bulaev, A., Kirkegaard, R.H., von Bergen, M., Rattei, T., Bendinger, B., Nielsen, P.H., Wagner, M. (2015) Complete nitrification by Nitrospira bacteria. *Nature 2015 528:7583*, **528**, 504–509.

Dechesne, A., Wang, G., Gülez, G., Or, D. and Smets, B.F. (2010) Hydration-controlled bacterial motility and dispersal on surfaces. *Proc Natl Acad Sci U S A*.

Delgado-Baquerizo, M., Eldridge, D.J., Liu, Y.R., Sokoya, B., Wang, J.T., Hu, H.W., He, J.Z., Bastida, F., Moreno, J.L., Bamigboye, A.R., Blanco-Pastor, J.L., Cano-Díaz, C., Illán, J.G., Makhalanyane, T.P., Siebe, C., Trivedi, P., Zaady, E., Verma, J.P., Wang, L., Wang, J., Grebenc, T., Peñaloza'Bojacá, G.F., Nahberger, T.U., Teixido, A.L., Zhou, X.Q., Berdugo, M., Duran, J., Rodríguez, A., Zhou, X., Alfaro, F., Abades, S., Plaza, C., Rey, A., Singh, B.K., Tedersoo, L., Fierer, N. (2021) Global homogenization of the structure and function in the soil microbiome of urban greenspaces. *Sci Adv*, **7**.

Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-González, A., Eldridge, D.J., Bardgett, R.D., Maestre, F.T., Singh, B.K. and Fierer, N. (2018) A global atlas of the dominant bacteria found in soil. *Science (1979)*, **359**, 320–325.

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., Mering, C. von and Vorholt, J.A. (2009) Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proc Natl Acad Sci U S A*, **106**, 16428–16433.

DeLong, E.F. and Pace, N.R. (2001) Environmental Diversity of Bacteria and Archaea. *Syst Biol*, **50**, 470–478.

Dengler, N.G. and Tsukaya, H. (2001) Leaf Morphogenesis in Dicotyledons: Current Issues. *https://doi.org/10.1086/320145*, **162**, 459–464.

Donnelly, P.M., Bonetta, D., Tsukaya, H., Dengler, R.E. and Dengler, N.G. (1999) Cell Cycling and Cell Enlargement in Developing Leaves of Arabidopsis. *Dev Biol*, **215**, 407–419.

Dubois, M., Claeys, H., Broeck, L. van den and Inzé, D. (2017) Time of day determines Arabidopsis transcriptome and growth dynamics under mild drought. *Plant Cell Environ*, **40**, 180–189.

Durán, P., Thiergart, T., Garrido-Oter, R., Agler, M., Kemen, E., Schulze-Lefert, P. and Hacquard, S. (2018) Microbial Interkingdom Interactions in Roots Promote Arabidopsis Survival. *Cell*, **175**, 973-983.e14.

Ebrahimi, A. and Or, D. (2015) Hydration and diffusion processes shape microbial community organization and function in model soil aggregates. *Water Resour Res*, **51**, 9804–9827.

Ebrahimi, A. and Or, D. (2016) Microbial community dynamics in soil aggregates shape biogeochemical gas fluxes from soil profiles - upscaling an aggregate biophysical model. *Glob Chang Biol*, **22**, 3141–3156.

Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N.K., Bhatnagar, S., Eisen, J.A., Sundaresan, V. and Jeffery, L.D. (2015) Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci U S A*, **112**, E911–E920.

FAOSTAT (2020) https://www.fao.org/faostat/en/#data/QCL/visualize. Date accessed: 11/06/2022.

Fierer, N., Lauber, C.L., Ramirez, K.S., Zaneveld, J., Bradford, M.A. and Knight, R. (2012) Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME J*, **6**, 1007–17.

Finkel, O.M., Salas-González, I., Castrillo, G., et al. (2020) A single bacterial genus maintains root growth in a complex microbiome. *Nature 2020 587:7832*, **587**, 103–108.

Finkel, O.M., Salas-González, I., Castrillo, G., Spaepen, S., Law, T.F., Teixeira, P.J.P.L., Jones, C.D. and Dangl, J.L. (2019) The effects of soil phosphorus content on plant microbiota are driven by the plant phosphate starvation response. *PLoS Biol*, **17**, e3000534.

Fox, A., Ikoyi, I., Torres-Sallan, G., Lanigan, G., Schmalenberger, A., Wakelin, S. and Creamer, R. (2018) The influence of aggregate size fraction and horizon position on microbial community composition. *Applied Soil Ecology*.

Gilbert, J.A., Jansson, J.K. and Knight, R. (2014) The Earth Microbiome project: Successes and aspirations. *BMC Biol*, **12**, 1–4.

Goodman, A.L., Kallstrom, G., Faith, J.J., Reyes, A., Moore, A., Dantas, G. and Gordon, J.I. (2011) Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc Natl Acad Sci U S A*, **108**, 6252–6257.

Ha, C.M., Ji, H.J., Hong, G.N. and Fletcher, J.C. (2007) BLADE-ON-PETIOLE1 and 2 control Arabidopsis lateral organ fate through regulation of LOB domain and adaxial-abaxial polarity genes. *Plant Cell*, **19**, 1809–1825.

Ha, C.M., Jun, J.H. and Fletcher, J.C. (2010) Control of Arabidopsis Leaf Morphogenesis Through Regulation of the YABBY and KNOX Families of Transcription Factors. *Genetics*, **186**, 197–206.

Hacquard, S., Spaepen, S., Garrido-Oter, R. and Schulze-Lefert, P. (2017) Interplay Between Innate Immunity and the Plant Microbiota. *Annu Rev Phytopathol*, **55**, 565–589.

Hiruma, K., Gerlach, N., Sacristán, S., Nakano, R.T., Hacquard, S., Kracher, B., Neumann, U., Ramírez, D., Bucher, M., O'Connell, R.J., Schulze-Lefert, P. (2016) Root Endophyte Colletotrichum tofieldiae Confers Plant Fitness Benefits that Are Phosphate Status Dependent. *Cell*, **165**, 464–74.

Hou, S., Thiergart, T., Vannier, N., Mesny, F., Ziegler, J., Pickel, B. and Hacquard, S. (2021) A microbiota–root–shoot circuit favours Arabidopsis growth over defence under suboptimal light. *Nature Plants 2021 7:8*, **7**, 1078–1092.

Ichihashi, Y., Kawade, K., Usami, T., Horiguchi, G., Takahashi, T. and Tsukaya, H. (2011) Key Proliferative Activity in the Junction between the Leaf Blade and Leaf Petiole of Arabidopsis. *Plant Physiol*, **157**, 1151–1162.

Ishida, H., Izumi, M., Wada, S. and Makino, A. (2014) Roles of autophagy in chloroplast recycling. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, **1837**, 512–521.

Jia, B., Raphenya, A.R., Alcock, B., Waglechner, N., Guo, P., Tsang, K.K., Lago, B.A., Dave, B.M., Pereira, S., Sharma, A.N., Doshi, S., Courtot, M., Lo, R, Williams, L.E., Frye, J.G., Elsayegh, T., Sardar, D., Westman, E.L., Pawlowski, A.C., Johnson, T.A., Brinkman, F.S.L., Wright, G.D., McArthur, A.G. (2017) CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res*, **45**, D566–D573.

Johnson J.S., Spakowicz D.J., Hong B.Y., Petersen L.M., Demkowicz P., Chen L., Leopold S.R., Hanson B.M., Agresta H.O., Gerstein M., Sodergren E., Weinstock G.M. (2019) Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications 2019 10:1*, **10**, 1–11.

Johnston, R., Leiboff, S. and Scanlon, M.J. (2015) Ontogeny of the sheathing leaf base in maize (Zea mays). *New Phytologist*, **205**, 306–315.

Kalve, S., Vos, D. de and Beemster, G.T.S. (2014) Leaf development: A cellular perspective. *Front Plant Sci*, **5**, 362.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M. and Tanabe, M. (2014) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res*, **42**, D199–D205.

Kumssa, D.B., Joy, E.J.M., Ander, E.L., Watts, M.J., Young, S.D., Walker, S. and Broadley, M.R. (2015) Dietary calcium and zinc deficiency risks are decreasing but remain prevalent. *Sci Rep*.

Kwak M.J., Kong H.G., Choi K., Kwon S.K., Song J.Y., Lee J., Lee P.A., Choi S.Y., Seo M., Lee H.J., Jung E.J., Park H., Roy N., Kim H., Lee M.M., Rubin E.M., Lee S.W., Kim J.F. (2018) Rhizosphere microbiome structure alters to enable wilt resistance in tomato. *Nature Biotechnology 2018 36:11*, **36**, 1100–1109.

Lagier, J.C., Dubourg, G., Million, Cadoret F., Bilen M., Fenollar F., Levasseur A., Rolain J.M., Fournier P.E., Raoult D. (2018) Culturing the human microbiota and culturomics. *Nature Reviews Microbiology 2018 16:9*, **16**, 540–550.

Lagier, J.C., Khelaifia, S., Alou, M.T., Ndongo S., Dione N., Hugon P., Caputo A., Cadoret F., Traore S.I., Seck E.H., Dubourg G., Durand G., Mourembou G., Guilhot E., Togo A., Bellali S., Bachar D., Cassir N., Bittar F., Delerce J., Mailhe M., Ricaboni D., Bilen M., Dangui Nieko N.P., Dia Badiane N.M., Valles C., Mouelhi D., Diop K., Million M., Musso D., Abrahão J., Azhar E.I., Bibi F., Yasir M., Diallo A., Sokhna C., Djossou F., Vitton V., Robert C., Rolain J.M., La Scola B., Fournier P.E., Levasseur A., Raoult, D. (2016) Culture of previously uncultured members of the human gut microbiota by culturomics. *Nature Microbiology 2016 1:12*, **1**, 1–8.

Lebeis, S.L., Paredes, S.H., Lundberg, D.S., Glavina, T. and Jones, C.D. (2015) Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. , **349**, 1678–1681.

Leff, J.W., Tredici, P. del, Friedman, W.E. and Fierer, N. (2015) Spatial structuring of bacterial communities within individual Ginkgo biloba trees. *Environ Microbiol*, **17**, 2352–2361.

Levy, A., Salas Gonzalez, I., Mittelviefhaus, M., Clingenpeel S., Herrera Paredes S., Miao J., Wang K., Devescovi G., Stillman K., Monteiro F., Rangel Alvarez B., Lundberg D.S., Lu T.Y., Lebeis S., Jin Z., McDonald M., Klein A.P., Feltcher M.E., Rio T.G., Grant S.R., Doty S.L., Ley R.E., Zhao B., Venturi V., Pelletier D.A., Vorholt J.A., Tringe S.G., Woyke T., Dangl J.L. (2017) Genomic features of bacterial adaptation to plants. *Nat Genet*, **50**, 138–150.

Levy, S.E. and Boone, B.E. (2019) Next-Generation Sequencing Strategies. *Cold Spring Harb Perspect Med*, **9**.

Liu, X., Chen, J. and Zhang, X. (2021) Genetic regulation of shoot architecture in cucumber. *Horticulture Research 2021 8:1*, **8**, 1–14.

Liu, Y.X., Qin, Y., Chen, T., Lu, M., Qian, X., Guo, X. and Bai, Y. (2021) A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell*, **12**, 315–330.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M. and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*, **42**.

Lornac, A., Havé, M., Chardon, F., Soulay, F., Clément, G., Avice, J.C. and Masclaux-Daubresse, C. (2020) Autophagy Controls Sulphur Metabolism in the Rosette Leaves of Arabidopsis and Facilitates S Remobilization to the Seeds. *Cells*, **9**.

Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone S., Gehring J., Malfatti S., Tremblay J., Engelbrektson A., Kunin V., Del Rio T.G., Edgar R.C., Eickhorst T., Ley R.E., Hugenholtz P., Tringe S.G., Dangl J.L. (2012) Defining the core Arabidopsis thaliana root microbiome. *Nature*, **488**, 86–90.

Mavrodi, O. v., McWilliams, J.R., Peter, J.O., Berim A., Hassan K.A., Elbourne L.D.H., LeTourneau M.K., Gang D.R., Paulsen I.T., Weller D.M., Thomashow L.S., Flynt A.S., Mavrodi D.V. (2021) Root Exudates Alter the Expression of Diverse Metabolic, Transport, Regulatory, and Stress Response Genes in Rhizosphere Pseudomonas. *Front Microbiol*, **12**, 651282.

Mitri, S., Clarke, E. and Foster, K.R. (2016) Resource limitation drives spatial organization in microbial groups. *ISME J*, **10**, 1471–82.

Moitinho, M.A., Bononi, L., Souza, D.T., Melo, I.S. and Taketani, R.G. (2018) Bacterial Succession Decreases Network Complexity During Plant Material Decomposition in Mangroves. *Microb Ecol*, **76**, 954–963.

Moreno, A., Ramos, T.B., Gonçalves, M.C. and Pereira, L.S. (2014) Estimating soil hydraulic properties from limited data to improve irrigation management in agricultural soils of Santiago Island, Cape Verde. *Irrigation and Drainage*, **63**, 405–415.

Müller, D.B., Vogel, C., Bai, Y. and Vorholt, J.A. (2016) The Plant Microbiota: Systems-Level Insights and Perspectives. *http://dx.doi.org/10.1146/annurev-genet-120215-034952*, **50**, 211–234.

Najar, I.N., Sherpa, M.T., Das, S. and Thakur, N. (2020) Bacterial diversity and functional metagenomics expounding the diversity of xenobiotics, stress, defense and CRISPR gene ontology providing eco-efficiency to Himalayan Hot Springs. *Funct Integr Genomics*, **20**, 479–496.

Nelissen, H., Eeckhout, D., Demuynck, K., Persiau G., Walton A., van Bel M., Vervoort M., Candaele J., De Block J., Aesaert S., Van Lijsebettens M., Goormachtig S., Vandepoele K., Van Leene J., Muszynski M., Gevaert K., Inzé D., De Jaeger G. (2015) Dynamic Changes in ANGUSTIFOLIA3 Complex Composition Reveal a Growth Regulatory Mechanism in the Maize Leaf. *Plant Cell*, **27**, 1605–1619.

Nelissen, H., Rymen, B., Jikumaru, Y., Demuynck, K., Lijsebettens, M. van, Kamiya, Y., Inzé, D. and Beemster, G.T.S. (2012) A Local Maximum in Gibberellin Levels Regulates Maize Leaf Growth by Spatial Control of Cell Division. *Current Biology*, **22**, 1183–1187.

Nelissen, H., Sun, X.H., Rymen, B., Nelissen H., Sun X.H., Rymen B., Jikumaru Y., Kojima M., Takebayashi Y., Abbeloos R., Demuynck K., Storme V., Vuylsteke M., De Block J., Herman D., Coppens F., Maere S., Kamiya Y., Sakakibara H., Beemster G.T.S., Inzé D. (2018) The reduction in maize leaf growth under mild drought affects the transition between cell division and cell expansion and cannot be restored by elevated gibberellic acid levels. *Plant Biotechnol J*, **16**, 615–627.

Niu, B., Paulson, J.N., Zheng, X. and Kolter, R. (2017) Simplified and representative bacterial community of maize roots. *Proc Natl Acad Sci U S A*, **114**, E2450–E2459.

Oberhardt, M.A., Zarecki, R., Gronow, S., Lang, E., Klenk, H.P., Gophna, U. and Ruppin, E. (2015) Harnessing the landscape of microbial culture media to predict new organism–media pairings. *Nature Communications 2015 6:1*, **6**, 1–14.

Ofek-Lalzar, M., Sela, N., Goldman-Voronov, M., Green, S.J., Hadar, Y. and Minz, D. (2014) Niche and host-associated functional signatures of the root surface microbiome. *Nature Communications 2014 5:1*, **5**, 1–9.

Otegui, M.S. (2018) Vacuolar degradation of chloroplast components: autophagy and beyond. *J Exp Bot*, **69**, 741–750.

Pantin, F., Simonneau, T. and Muller, B. (2012) Coming of leaf age: control of growth by hydraulics and metabolics during leaf ontogeny. *New Phytologist*, **196**, 349–366.

Pick, T.R., Bräutigam, A., Schlüter, U., Denton A.K., Colmsee C., Scholz U., Fahnenstich H., Pieruschka R., Rascher U., Sonnewald U., Weber A.P. (2011) Systems Analysis of a Maize Leaf Developmental Gradient Redefines the Current C4 Model and Provides Candidates for Regulation. *Plant Cell*, **23**, 4208–4220.

Poethig, R.S. (2013) Vegetative phase change and shoot maturation in plants. *Curr Top Dev Biol*, **105**, 125–152.

Raina, V., Nayak, T., Ray, L., Kumari, K. and Suar, M. (2019) A Polyphasic Taxonomic Approach for Designation and Description of Novel Microbial Species. *Microbial Diversity in the Genomic Era*, 137–152.

Richardson, A.E., Cheng, J., Johnston, R., Kennaway R., Conlon B.R., Rebocho A.B., Kong H., Scanlon M.J., Hake S., Coen E. (2021) Evolution of the grass leaf by primordium extension and petiole-lamina remodeling. *Science (1979)*, **374**, 1377–1381.

Roesch, L.F.W., Fulthorpe, R.R., Riva, A., Casella G., Hadwin A.K., Kent A.D., Daroub S.H., Camargo F.A., Farmerie W.G., Triplett E.W. (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. *The ISME Journal 2007 1:4*, **1**, 283–290.

Salas-González, I., Reyt, G., Flis, P., Custódio V., Gopaulchan D., Bakhoum N., Dew T.P., Suresh K., Franke R.B., Dangl J.L., Salt D.E., Castrillo G. (2021) Coordination between microbiota and root endodermis supports plant mineral nutrient homeostasis. *Science (1979)*, **371**.

Sanchez-Moreno, J.F., Mannaerts, C.M. and Jetten, V. (2014a) Influence of topography on rainfall variability in Santiago Island, Cape Verde. *International Journal of Climatology*, **34**, 1081–1097.

Sanchez-Moreno, J.F., Mannaerts, C.M. and Jetten, V. (2014b) Rainfall erosivity mapping for Santiago Island, Cape Verde. *Geoderma*, **217–218**, 74–82.

Satterlee, J.W. and Scanlon, M.J. (2019) Coordination of Leaf Development Across Developmental Axes. *Plants 2019, Vol. 8, Page 433*, **8**, 433.

Schimel, J.P. (2018) Life in dry soils: Effects of drought on soil microbial communities and processes. *Annu Rev Ecol Evol Syst*.

Sessitsch, A., Hardoim, P., Döring, J., Weilharter A., Krause A., Woyke T., Mitter B., Hauberg-Lotte L., Friedrich F., Rahalkar M., Hurek T., Sarkar A., Bodrossy L., van Overbeek L., Brar D., van Elsas J.D., Reinhold-Hurek. (2012) Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. *Mol Plant Microbe Interact*, **25**, 28–36.

Strable, J. (2021) Developmental genetics of maize vegetative shoot architecture. *Molecular Breeding 2021 41:3*, **41**, 1–31.

Tadege, M., Lin, H., Bedair, M., Berbel A., Wen J., Rojas C.M., Niu L., Tang Y., Sumner L., Ratet P., McHale N.A., Madueño F., Mysore K.S. (2011) STENOFOLIA regulates blade outgrowth and leaf vascular patterning in Medicago truncatula and Nicotiana sylvestris. *Plant Cell*, **23**, 2125–2142.

Tian, J., Wang, C., Xia, J., Wu L., Xu G., Wu W., Li D., Qin W., Han X., Chen Q., Jin W., Tian F. (2019) Teosinte ligule allele narrows plant architecture and enhances high-density maize yields. *Science (1979)*, **365**, 658–664.

Truyens, S., Weyens, N., Cuypers, A. and Vangronsveld, J. (2015) Bacterial seed endophytes: genera, vertical transmission and interaction with plants. *Environ Microbiol Rep*, **7**, 40–50.

Turnbaugh, P.J., Ley, R.E., Hamady, M., Fraser-Liggett, C.M., Knight, R. and Gordon, J.I. (2007) The Human Microbiome Project. *Nature 2007 449:7164*, **449**, 804–810.

Varela-Lopes, G.E. and Molion, L.C.B. (2014) Precipitation Patterns in Cape Verde Islands: Santiago Island Case Study. *Atmospheric and Climate Sciences*, **04**, 854–865.

Velez, P., Espinosa-Asuar, L., Figueroa, M., Gasca-Pineda, J., Aguirre-von-Wobeser, E., Eguiarte, L.E., Hernandez-Monroy, A. and Souza, V. (2018) Nutrient Dependent Cross-Kingdom Interactions: Fungi and Bacteria From an Oligotrophic Desert Oasis . *Frontiers in Microbiology* , **9**, 1755.

Vetsigian, K., Jajoo, R. and Kishony, R. (2011) Structure and evolution of Streptomyces interaction networks in soil and in silico. *PLoS Biol*, **9**, e1001184.

Vogel, C.M., Potthoff, D.B., Schäfer, M., Barandun, N. and Vorholt, J.A. (2021) Protective role of the Arabidopsis leaf microbiota against a bacterial pathogen. *Nature Microbiology 2021 6:12*, **6**, 1537–1548.

Vorholt, J.A. (2012) Microbial life in the phyllosphere. *Nature Reviews Microbiology 2012 10:12*, **10**, 828–840.

Vos, D. de, Nelissen, H., AbdElgawad, H., Prinsen, E., Broeckhove, J., Inzé, D. and Beemster, G.T.S. (2020) How grass keeps growing: an integrated analysis of hormonal crosstalk in the maize leaf growth zone. *New Phytologist*, **225**, 2513–2525.

Wagner, M.R., Lundberg, D.S., Rio, T.G., Tringe, S.G., Dangl, J.L. and Mitchell-olds, T. (2016) Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat Commun*, **7**, 1–15.

Wallace, J.G., Kremling, K.A., Kovar, L.L. and Buckler, E.S. (2018) Quantitative genetics of the maize leaf microbiome. *Phytobiomes J*, **2**, 208–224.

Walters, W.A., Jin, Z., Youngblut, N., Wallace J.G., Sutter J., Zhang W., González-Peña A., Peiffer J., Koren O., Shi Q., Knight R., Glavina Del Rio T., Tringe S.G., Buckler E.S., Dangl J.L., Ley R.E. (2018) Large-scale replicated field study of maize rhizosphere identifies heritable microbes. *Proc Natl Acad Sci U S A*, **115**, 7368–7373.

Ward, T., Larson, J., Meulemans, J., Hillman, B., Lynch, J., Sidiropoulos, D., Spear, J.R., Caporaso, G., Blekhman, R., Knight, R., Fink, R., Knights, D. (2017) BugBase predicts organism-level microbiome phenotypes. *bioRxiv*.

Woo, H.R., Koo, H.J., Kim, J., Jeong H., Yang J.O., Lee I.H., Jun J.H., Choi S.H., Park S.J., Kang B., Kim Y.W., Phee B.K., Kim J.H., Seo C., Park C., Kim S.C., Park S., Lee B., Lee S., Hwang D., Nam H.G., Lim P.O. (2016) Programming of Plant Leaf Senescence with Temporal and Inter-Organellar Coordination of Transcriptome in Arabidopsis. *Plant Physiol*, **171**, 452–467.

Yu, P., Gutjahr, C., Li, C. and Hochholdinger, F. (2016) Genetic Control of Lateral Root Formation in Cereals. *Trends Plant Sci*, **21**, 951–961.

Zarraonaindia, I., Owens, S.M., Weisenhorn, P., West K., Hampton-Marcell J., Lax S., Bokulich N.A., Mills D.A., Martin G., Taghavi S., van der Lelie D., Gilbert J.A. (2015) The soil microbiome influences grapevine-associated microbiota. *mBio*, **6**.

Zhang, J., Liu, Y.X., Guo, X., Qin, Y., Garrido-Oter, R., Schulze-Lefert, P. and Bai, Y. (2021) High-throughput cultivation and identification of bacteria from the plant root microbiota. *Nature Protocols 2021 16:2*, **16**, 988–1012.

Zhang, J., Liu, Y.X., Zhang, N., Hu B., Jin T., Xu H., Qin Y., Yan P., Zhang X., Guo X., Hui J., Cao S., Wang X., Wang C., Wang H., Qu B., Fan G., Yuan L., Garrido-Oter R., Chu C., Bai Y. (2019) NRT1.1B is associated with root microbiota composition and nitrogen use in field-grown rice. *Nat Biotechnol*.

Zhang, Z., Qu, Y., Li, S., Feng K., Wang S., Cai W., Liang Y., Li H., Xu M., Yin H., Deng Y. (2017) Soil bacterial quantification approaches coupling with relative abundances reflecting the changes of taxa. *Scientific Reports 2017 7:1*, **7**, 1–11.

Zhu, G., Wang, S., Wang, C., Zhou L., Zhao S., Li Y., Li F., Jetten M.S.M., Lu Y., Schwark L. (2019) Resuscitation of anammox bacteria after >10,000 years of dormancy. *ISME Journal*, **13**, 1098–1109.

Chapter II

The role of soil mineral content and soil microbiota in maize productivity

Abstract

Maize holds significant importance as a primary dryland crop, but its productivity faces substantial challenges due to climate change and soil degradation. To cope with this challenge, several agronomic management are applied that drive multiple changes in structural and functional ecosystem attributes such as nutrient cycling, microbial community and plant productivity. Here, we characterize the agricultural practices and environmental characteristics that affect maize productivity in Santiago island (Cabo Verde). Using Inductively Coupled Plasma Mass Spectrometry (ICP-MS), we showed that soil mineral content is a good integrator of water- and nutrient-related factors. And since the composition of soil microbiota is often modulated by the availability of nutrients in the soil, we explore the soil free-living bacterial community in Santiago island. We observed strong structuring of the free-living bacteria according to the soil mineral content but not with the climatic conditions. Finally, we explore how the variation in soil mineral content and free-living bacteria could affect maize productivity in the experimental field independent of climate factors. We found that these variations or the combinations between them contribute to the maize differences in productivity in Santiago island.

1 Introduction

As a result of climate change, the amount of greenhouse gases in the atmosphere is increasing, temperatures are rising, precipitation patterns are being modified, and extreme weather events are occurring more frequently (Stott, 2016). These changes accounted for 25% of average yield losses for low-latitude maize (Rosenzweig *et al.*, 2014). To tackle this reduction of maize productivity, several strategies have been developed, such as altering the planting and harvesting time, sowing crops with a short life cycle, applying crop rotation and irrigation techniques, and introducing variation in cropping systems (Duku *et al.*, 2018; Marcinkowski and Piniewski, 2018; Teixeira *et al.*, 2018). One direct benefit of using these strategies to manage the soil is their positive impact on the soil microbiota. This effect has been described as a promising strategy to sustainably increase crop yield. For example, crop rotation balances the degradation and composition of nutrients in the soil and reduces the risk of plant diseases and animal pests infesting crops (van Bruggen *et al.*, 2016). Since different plants favour different microbiota in the soil, mainly through root exudates (Moitinho *et al.*, 2018; Mavrodi *et al.*, 2021), crop rotation also contributes to enhancing the diversity of microbial communities. Similarly, intercropping in which different crops are planted together and roots of different plant species interact directly, positively affecting soil fertility and microbial diversity and activity (Lian *et al.*, 2019; Zheng *et al.*, 2018; Zhou *et al.*, 2011). Although the beneficial effects of these practices are generally known, more knowledge is needed to understand the role of microorganisms in plant productivity. It will facilitate the development of new microbial-based strategies to increase plant productivity in agricultural settings.

Taking this into consideration, in this chapter, we have characterized the agricultural practices and environmental conditions that affect maize productivity in Santiago island. Using a survey applied to maize producers in Santiago island, we confirmed that maize production in Santiago island follows traditional practices that rely on natural resources (soil nutrient composition and precipitation). Indeed, by analysing climate and soil properties available data for the different producers, we demonstrated that water- and nutrient-related factors are the most variable factors across the different regions of the island.

Further, we selected a group of farms that captured the majority of climate and soil properties identified in the survey to collect soil samples for soil mineral content and microbial analysis. The results from these studies demonstrated that soil mineral content and soil microbiota differs between region. To better understand whether these variations affect maize productivity, we selected two field sites with similar climate characteristics. We imposed short-term changes in soil mineral content with an expected effect on the soil microbiota through two agricultural practices: fertilization and cropping system. We quantified the height of maize plants as a proxy for plant productivity. We found that the field site locations and the agricultural practices explained small variations observed in maize height, allowing us to reason that soil mineral content, soil microbiota or interactions between both affect maize height in the field.

2 Materials and methods

2.1 Main drivers affecting maize production in Cabo Verde - Santiago Island as a model

2.1.1 Census of agricultural practices and environmental conditions in Santiago Island

To identify the agricultural practices used in maize production in Santiago island, Cabo Verde, we conducted an interview-based survey crosswise of the whole island territory. We interviewed 122 farmers across 23 geographically diverse regions.

Table 1: The survey questionnaire designed for the farmers

General Information	Environmental and local characteristics	Farm and crop information
1- **Farmer's name**	1- Weather conditions at the day of interview (sunny, rainy or cloudy)	1- Farm size
2- Geographical coordinates of the field	2- Type of soil (Agricultural or non-agricultural)	2- Ownership period
	3- Landform (hill, valley, mountain, plain, plateau)	3- Soil nutrition management (manures, chemical fertilizers)
	4- Land exposed to the sun (Yes or Not)	4- In case of chemical fertilizers: Formulation and amount applied in the field
		5- Previous crop planted
		6- Cropping system used (monocrop or intercrop)
		7- Name of the species intercropped with maize
		8- Grain yield in the last season
		9- Seed sources
		10- Amount of seeds used per unit area per season

This selection of farms covers eight of the nine Santiago municipalities and represents 0.6% of the total maize production areas on the island. The survey questionnaire was divided into three major groups of questions: 1) general information, 2) environmental and local characteristics, and 3) farm and crop information (Table 1). The survey was conducted from August to November 2017.

2.1.2 Characterisation of the environmental conditions of the selected farmers

a) Data collection

To characterize the environmental conditions and soil properties in Santiago island, environmental and soil parameters were retrieved from publicly available sources with global coverage. The data were obtained via the function extract from the raster v3.5-2 package in R (Hijmans, 2014). Briefly, the global raster file was downloaded from Worldclim and Soilgrids databases for climate and soil edaphic data, respectively. The climate data include monthly minimum, maximum and average temperature; mean monthly precipitation, solar radiation, wind speed, and water vapour pressure (Fick and Hijmans, 2017). Additionally, we collect the aridity index (mean annual precipitation divided by mean annual potential evapotranspiration) calculated from Worldclim data and downloaded from Consultative Group for International Agricultural Research Consortium for Spatial Information (CGIAR-CSI) Globality-Aridity database (Zomer *et al.*, 2008).

The edaphic data included chemical and physical properties of the soil, such as nitrogen and carbon concentration, bulk density, soil texture fractions (percentage of sand, clay, and silt), coarse fragments, cation exchange capacity, and soil pH (Hengl *et al.*, 2017).

These variables were selected since they are often used to monitor soil health and land degradation (Shojaei *et al.*, 2020).

b) Computational and statistical analysis

To determine the correlation among climate and edaphic parameters, we first constructed independent matrices for the data of climate and edaphic factors. Next, we standardized the data by MinMax scaler or Robust scaler.

Then, we input the aforementioned independent matrices to calculate the Pearson correlation coefficients between the climate factors and another one between the edaphic factors. We subjected this correlation dissimilarity matrix to hierarchical clustering using ward.D method, implemented in the rcorr function from the Hmisc v4.6.0 R package (Oksanen, 2007a).

For clustering analysis of the environmental datasets, we used the chibi.heatmap function from the ohchibi R package. Clusters of environmental factors with similar patterns were delimited via dendrogram cutting through the function cutree from the stats R package (Oksanen, 2007a).

For each environmental and soil edaphic factor, we conducted a Random forest analysis (Breiman, 2001) to determine the importance of each factor in classifying the Santiago regions according to the characteristics of the environmental and edaphic data. The importance of each factor was determined on the 20% of training data through the function ExtraTreesClassifier from the Scikit-Learn package in Python version 3.8 (Anon, 2016).

2.1.3 Characterization of farmers' soils

c) Soil sampling

The soil sampling was undertaken across the 24 selected farms in two seasons. In the first season, September to November 2017, the sampling campaign was conducted during the maize production period in Cabo Verde. For the second season, April 2019, soil samples were collected after maize harvesting.

For the soil samples collection, each field was divided into six equal parts (subplots). Within each subplot, six samples were collected following a typical zig-zag pattern with approximately 3 m of separation between the harvesting points and avoiding field edges, manure piles, roads, and pathways. Soil samples were then collected from the top 10 cm layer using a handheld soil auger and pooled to produce a unique composite sample per subplot. Thus, for each soil, we collected 6 composite samples. All tools used for soil collection were previously disinfected with 70% ethanol, and the operators wore gloves the whole time. Soil samples were transported in clean plastic bags to Instituto

Nacional de Investigação e Desenvolvimento Agrário (INIDA), where they were air-dried in sterile aluminium trays at room temperature for 2 weeks and then sieved using a 2 mm sieve. Soil samples were stored at 4 ºC until use.

d) Soil nutrient and soil pH analysis

The phyto-availability of a selection of nutrients (P, K, S, Ca, Mg, Fe, Mn, Cu, Zn, Mo, B, Ni, Na, Si, Ti, Li, Be, Al, V, Cr, Co, As, Se, Rb, Sr, Ag, Cd, Cs, Ba, Tl, Pb, U) was determined in soil using a diethylenetriaminepentaacetic acid (DTPA)-based extraction method followed by inductively coupled plasma mass spectrometer (ICP-MS) analysis. Briefly, five grams of soil were weighed in polycarbonate centrifuge tubes and then extracted at room temperature with 10 mL of 5 mM DTPA, 0.1 M Triethanolamine, and 10 mM $CaCl_2$ buffer for 2h with agitation in a shaker. The resulting soil solutions were then centrifuged at 3000 rpm, for 10 min in a centrifuge Eppendorf 5810R, and the supernatants were filtered through a 0.22 µm syringe filter unit (Millipore, Cork, Ireland) into a universal tube (Sarstedt, Germany). 1 mL of each soil supernatant was diluted with 9 mL of 2% HNO_3 (PrimarPlus – Trace Analysis Grade (TAG), Fisher Scientific), and analyzed using ICP-MS, (Thermo Fisher X-Series^{II}, Thermo Fisher Scientific Inc.) equipped with Cetac ASX-520 covered autosampler (Omaha) with 4x60-place sample racks. The equipment was set up in the reaction mode using hydrogen as a reaction gas. Samples were introduced to ICP-MS at a 1 ml min^{-1} rate through a concentric glass venturi nebulizer and Peltier-cooled (3 ºC) spray chamber (Thermo Fisher Scientific Inc.). Internal standards were introduced to the sample stream via a T-piece and included Sc (50 ng/mL), Rh (10 ng/mL) and Ir (5 ng/mL) in 2% TAG HNO_3. The calibration standards were prepared from single-element standards solutions. Sample concentrations were calculated using the external calibration method within the instrument software.

To measure the pH of the collection of soils, each soil sample was mixed in 12.5 mL of water with agitation for 30 min, and the pH of the solubilized fraction was measured (Fisherbrand™, Thermo Fisher Scientific Inc.). We used 6 biological replicates per soil.

2.1.4 Soil free-living bacteria – a 16S rRNA barcode analyses

a) DNA extraction

To determine the composition of bacterial communities in the collection of soils, approximately 5 g of each soil was transferred into 50-mL falcon tubes containing 20 mL of sterile distilled water. To remove large plant debris and big soil particles, the samples were shaken thoroughly and filtered using 100-µm nylon mesh cell strainers (Fisherbrand™, Thermo Fisher Scientific Inc.) into a new sterile 50-mL tube. The filtered solutions were centrifuged at 4,000 rpm for 20 min in a centrifuge Eppendorf 5810R, and most of the supernatants were discarded. The remaining 1-2 mL of supernatants were used to dissolve the pellets, and the resulting suspensions were transferred to sterile 1.5-mL Eppendorf tubes. Samples were centrifuged again at 13,000 rpm for 20 min in a benchtop centrifuge, and the supernatants were discarded. The remaining pellets were stored at -80 ºC until DNA extraction. For the DNA extraction, we used 96-well-format MoBio PowerSoil Kit (MOBIO Laboratories; Qiagen) following the manufacturer's instructions. Before starting the extraction, the individual eppendorf with samples were randomized by placing them in a plastic bag and shaking them several times. Samples were taken individually from the bag and loaded into the DNA extraction plates. This random distribution was maintained throughout library preparation and sequencing.

b) Amplification and sequencing of 16S rRNA

For the 16S rRNA sequencing, the V3-V4 region of the bacterial 16S rRNA gene was amplified using the primers 338F (5′-ACTCCTACGGGAGGCAGCA-3′) and 806R (5′-GGACTACHVGGGTWTCTAAT-3′). Each sample was amplified in triplicate, and the PCR reactions were performed using three sets of combinations of unique primers per plate. The following conditions were used for PCR amplification: 5 µL Kapa Enhancer, 5 µL Kapa Buffer A, 1.25 µL of 5µM 338F, 1.25 µL of 5µM 806R, 0.375 µL mixed plastid and mitochondrial rRNA gene-blocking peptide nucleic acids (PNAs; 1:1 mix of 100 µM plastid PNA and 100 µM mitochondrial PNA), 0.5 µL Kapa dNTPs, 0.2 µL Kapa Robust Taq, 8 µL dH2O and 8 µL DNA; temperature cycling: 95 °C for 1 min, 24 cycles of 95 °C for 15 s, 78 °C (PNA) for 10 s, 50 °C for 30 s, 72 °C for 30 s and 4 °C until use. PCR products of triplicate were combined and purified using AMPure

XP magnetic beads (Beckman Coulter) to remove primer dimers. Ten microliter of purified DNA was subjected to a second PCR using the following conditions: 12 µL Kapa HiFi Readymix, 0.375 µL mixed (1:1) of 100 µM pPNA and 100 µM mPNA; 2.5 µL of 5 µM index primers; temperature cycling: 95 ºC for 1 min, 9 cycles of 95 ºC for 15 s, 78 ºC for 10 s, 60 ºC for 30 s, 72 ºC for 35 s and 4 °C until use. Subsequently, amplicons were pooled together in equal amounts and then diluted to 10 pM for sequencing. Sequencing was performed on an Illumina MiSeq instrument using a 600-cycle V3 chemistry kit in the DeepSeq facility at the University of Nottingham. DNA sequence data, abundance matrix, metadata, and taxonomy will be available upon publication.

c) Amplicon sequence data processing

For the 16S rRNA analysis, sequence data were processed with cutadapt (Martin, 2011), DADA2, and R packages. Raw reads were demultiplexed and trimmed with cutadapt. The resulting sequences were then denoised and collapsed into amplicon sequence variants (ASVs) using DADA2 v.1.10.1 (Callahan *et al.*, 2016). Briefly, the paired reads were filtered by removing sequences containing uncalled bases, truncating reads after bases with a quality score of 2 and reads with no more than 3 expected errors using the FilterAndTrim function. We then learned the error for the forward and reverse reads separately with the learnErrors function. These error rates were then used to infer amplicon sequence variants (ASVs) with the function dada on the reverse and forward reads individually. Afterwards, we merged the forward and reverse sequence with the mergePairs function. The merged ASVs were used to build an ASV sequencing table using the makeSequenceTable function. Then, we removed chimaeras from the sequence table with the removeBimeraDenova function. The resulting ASVs were mapped to SILVA 132 database (Quast *et al.*, 2013). We filtered ASVs that were assigned to chloroplasts, mitochondria, oomycetes, archaea or did not have a known kingdom assignment. After the filtering, the remaining ASVs, with more than 1000 reads per sample were used to create the raw count abundance tables. The resulting abundance tables were processed and analysed with functions from the ohchibi package (https://github.com/isaisg/ohchibi).

2.1.5 Data analysis

To generate a matrix of bacterial relative abundance averages across replicates in each region, we created a group variable that merged the region variable and the latitude coordinate (Lat). Then, we averaged all replicates in each farmer's field using this group variable. Finally, we estimated the Bray-Curtis dissimilarity using the matrix constructed above. These average distances were hierarchically clustered (hclust in R; method = "complete"). The relative abundance of bacterial phyla was depicted using a stacked bar representation encoded in the function chibi.phylogram from the ohchibi package (https://github.com/isaisg/ohchibi).

To compare alpha diversity across Santiago Island regions, we calculated the Shannon diversity index using the diversity function from the vegan package v2.5-5 (Oksanen, 2007b). We used ANOVA to test for differences in alpha diversity between the different regions. Beta diversity analyses (Principal coordinate analysis, Pco, and canonical analysis of principal coordinates, CAP) were based on Bray-Curtis dissimilarity matrices calculated from the rarefied abundance table. PCo using the function prcomp in R. We used the capscale function from the vegan R package version v2.5-5 (Oksanen, 2007b) to compute CAP. To analyse the full dataset (all the factors in the dataset), we constrained by region while conditioning for the year. Additionally, we estimated the variance explained by soil regions and agricultural management by performing PERMANOVA via the function adonis from the vegan v2-5.5 R package (Oksanen, 2007b).

2.2 Modulation of ionome and microbiome effect on maize productivity

2.2.1 Experimental plots – edaphic conditions

a) Soil bulk density and texture

The soil samples used to determine the soil bulk density were collected between June and September 2017, before applying the treatments to the field experiments. Within each experimental location (São Domingos and Tarrafal), samples were taken from 5 different sites distributed across the whole experimental area. The soil bulk density *(ρ_b)* was determined, in two soil

horizons: 0 cm to 20 cm, and from 20 cm to 40 cm of depth, using an adapted excavation method (Grossman and Reinsch, 2002). Briefly, we removed the vegetation in each individual site (50 x 50 cm) using a hoe, and approximately 2 cm of the topsoil was discarded. To determine the soil weight contained in a given volume in the first horizon (0 cm to 20 cm), we excavated in each site a soil cube of 20 x 20 x 20 cm. The soils derived from these cubes were collected in plastic bags individually. Then, cubes were covered with plastic, and we estimated the volume of soil extracted by filling the cube with a known volume of water. Using the same procedure, we estimate the weight of soil in a given volume for the second soil horizon (20 cm to 40 cm). The soils derived from both horizons were weighted in the lab. To estimate the total soil dry weight, 100 g for each horizon was dried in an oven at 110 ºC and then weighed again. Subsequently, the bulk density was calculated using the following formula (Bilskie and Campbell Scientific, 2001):

$$\rho_b = \frac{M_s}{V}$$

where, M_s is the total soil dry weight (g) and V is the total volume of soil (L).

The soil texture (proportion of sand, silt and clay) was determined in the soil samples collected from the two soil horizons (0 cm to 20 cm and 20 cm to 40 cm) using a certified protocol (ISO 11277:2009) at the INIDA laboratory, Cabo Verde. A total of 20 g of each soil was treated individually with 5% sodium hexametaphosphate with agitation to disperse the soil particles. Then, for each individual soil, the sand fractions were obtained by sieving the samples using a 0.05 mm sieve. The resulting solution containing the silt and clay fractions was separated using a Robinson's pipette. Each fraction was oven-dried at 105 ºC, and the proportion of sand, silt and clay was calculated for each soil.

b) Estimate soil available water content

To estimate the soil available water content, namely the water available to plants in the two experimental fields, we first calculated the field capacity and permanent wilting point (Rai *et al.*, 2017) using the soil texture data from material and methods 2.2.1a. Field capacity (θ_f) was calculated using the following equation: $\theta_f = 0.555P_{clay} + 0.187P_{silt} + 0.027P_{sand}$. Where P_{clay} is the percentage of clay, P_{silt} is the percentage of silt and P_{sand} is the percentage of

sand. Afterwards, we applied the results from θ_f to calculate the permanent wilting point (θ_w) (Briggs and Shantz, 1912): $\theta_w = \theta_f/1.84$. Finally, the soil available water content (θaw) was calculated using the following equation (Rai *et al.*, 2017): $\theta_{aw} = \theta_f - \theta_w$.

c) Soil sampling for elemental and soil pH analyses

To determine the effect of the different treatments on the elemental profile and pH of the soils, we collected samples from the individual blocks in the field experiments (Figure 1). In all cases, soil samples were collected from the first 10 cm of soil, using a clean handheld soil auger at three time points: (i) before starting the experiments, (ii) at the plant tasseling stage, and (iii) at the seed harvest. Five samples were collected from each individual block and then pooled to produce a unique composite sample. In the case of time points ii and iii, the samples were collected at approximately 5 cm distance from the maize stalks. The soil samples were transported to the lab in a clean plastic bag, air-dried in clean plastic trays for two weeks, and sieved using a 2 mm sieve. Soil samples were kept at 4 ºC until use. The analysis of the soil elements and pH were conducted as described in Materials and Methods 2.1.3d.

2.2.2 Set-up of field experiment

a) Irrigation scheduling strategy

In order to define the amount of water required to irrigate the crop, the appropriate irrigation interval, and the irrigation time to provide the required amount of water to the crop, we have estimated the total available water in the root zone and maize water requirement (Rai *et al.*, 2017; Allen *et al.*, 1998). Briefly, the amount of water required to irrigate maize in the two experimental fields was estimated using the total available water (TAW) formula (Rai *et al.*, 2017). Next, we determined the maize water requirement (ET_c) using crop coefficient (K_c) and reference crop evaporation (ET_0) (Allen *et al.*, 2005). The K_c values were obtained from FAO-56 document (Allen *et al.*, 1998) and meteorological datasets from São Domingos and Tarrafal were used to calculate reference crop evaporation (ET_0) using Class A pan method (Snyder *et al.*, 2005). To calculate the interval of irrigation in days we divided the results from TAW by the results of ET_C x 10.

Once the crop water requirement and the irrigation interval were defined, we estimated the irrigation time to provide the required amount of water to the crop. First, we calculated the total volume of water required to irrigate an entire block in each field: Total volume = area (m^2) x (ET_C (mm/day)/1000). Next, we calculated the volume of irrigation to compensate for the plant water requirement in the irrigation interval: irrigation volume = Total volume x irrigation interval. Finally, the irrigation time (time) was calculated using the following formula: time (min) = (irrigation volume/Q)/60. The flow rate (Q) was inferred through the relation of loss of pressure (h_f) and Q using the Colebrook equation, Darcy-Weisbach equation and the definition of flow rate (Munson *et al.*, 1995):

$$Q = -2\,\frac{\pi\, D^2\,\sqrt{h_f D g}}{\sqrt{8L}}\; log_{10}\left(\frac{\frac{\varepsilon}{D}}{3.7} + \frac{2.51\sqrt{8L}}{\frac{4D}{\nu}\sqrt{h_f D g}}\right)$$

Where Q is the volumetric flow rate ($m^3\ s^{-1}$), D and L is the pipe diameter (m) and pipe length, respectively in each experimental field. The h_f is the pressure or head loss (m), g the gravitational force ($m\ s^{-2}$), ε is the absolute wall roughness (mm) and v the water viscosity ($m^2\ s^{-1}$).

b) Experimental design

To evaluate the effect of the agricultural practices (fertilization and intercropping) and the soil microbiota in maize productivity, two field experiments were conducted simultaneously in 2017 and 2019. The experiments took place at two different locations in Santiago Island, the semi-arid region (331 mm of accumulated rainfall per year) of São Domingos (15º1'7.6728''N, 23º32'54.3372''W), and the arid region (196 mm of accumulated rainfall) of Tarrafal (15º15'17.5468''N, 23º44'38.9256''W). In each case, a split-plot design with completely randomized blocks was used in triplicate (Figure 1). The dimensions of a single randomized block were 6.3 m length × 3.2 m width, comprising 5 rows per plot, with a total area of 20.16 m^2. Each block was separated from each other at a distance of 1 m.

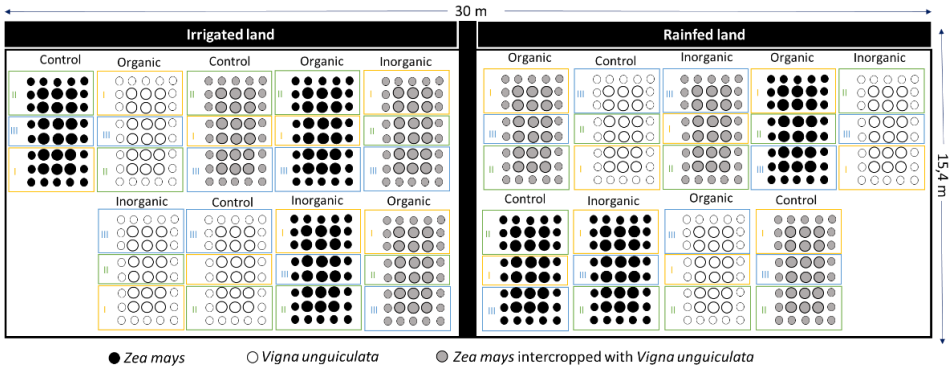


Figure 1: Field experimental design. In each plot, the central area indicates the crops that were monitored (larger circles), while the surrounding area (small circles) represents the "buffer" zone (the same crop grown under the same treatments but not used for sampling).

Following the experimental design in Figure 1, we sowed maize (*Zea mays*) and cowpea (*Vigna unguiculata*) seeds alone (monocrop) or together (intercrop) with 70 cm inter-row and 80 cm intra-row spacing. In each field, the randomized blocks were additionally exposed to two treatments, control (monocrops and intercrop that never received fertilization) and inorganic fertilization (monocrops and intercrop that received 120 kg/ha $(NH4)_2SO_4$ twenty days after sowing) (Figure 1). No fungicides or herbicides were used during the experiments. Automatic Em5b data loggers with moisture sensors (Decagon Devices) were installed in each experimental field, two per field, at a depth of 5 cm to register the volumetric moisture content in the soil every hour.

2.2.3 Characterization of soil free-living bacterial composition

The soil samples extracted from each individual block were used for the microbiome characterization. The DNA extraction, 16S rRNA amplification, sequencing, and data analysis were performed as above.

2.2.4 Maize height quantification

To measure the plant height, 18 fully-developed maize plants were selected from the central area of each block avoiding the edges. Then, for each individual plant, the distance from the soil surface to the base of the plant tassel was determined using a measuring tape.

3 Results

3.1 Maize production in Cabo Verde relies on natural resources

Maize is the most abundant cereal in the world. As a staple food, its production in some areas is very high. For example, in the USA, maize production in 2016 was 11 tons/ha. Unfortunately, this is not the case for Cabo Verde, which shows low yields with an average of 0.178 tons/ha in 2016, 62 times less than the USA, and it is one of the lowest maize producers among the major producers in Africa (Figure 2A). Therefore, new maize production strategies are needed to increase maize yield in Cabo Verde.

To identify factors influencing maize production in Cabo Verde with the potential to be used in strategies to improve maize yield, we conducted an interview-based survey in the major agricultural island in Cabo Verde, Santiago island. We interviewed 122 farmers across Santiago island territory, covering 24 geographically diverse regions representing eight Santiago municipalities.

During the personal interviews in each selected location, each farmer was asked to list, for example, the agricultural practices used, the soil fertility management practices applied and crops used, and the yield from the previous year (Figure 2B).

The analysis of the survey data showed that maize production on the island of Santiago is mainly traditional, a rainfed agricultural practice where, in 90% of the cases, the farmers avoid the use of commercial chemical fertilizers and primarily produce maize in intercrop with beans (Figure 2B). These results indicate that maize production in Santiago island follows traditional practices, relying on natural soil resources and being vulnerable to environmental fluctuations.

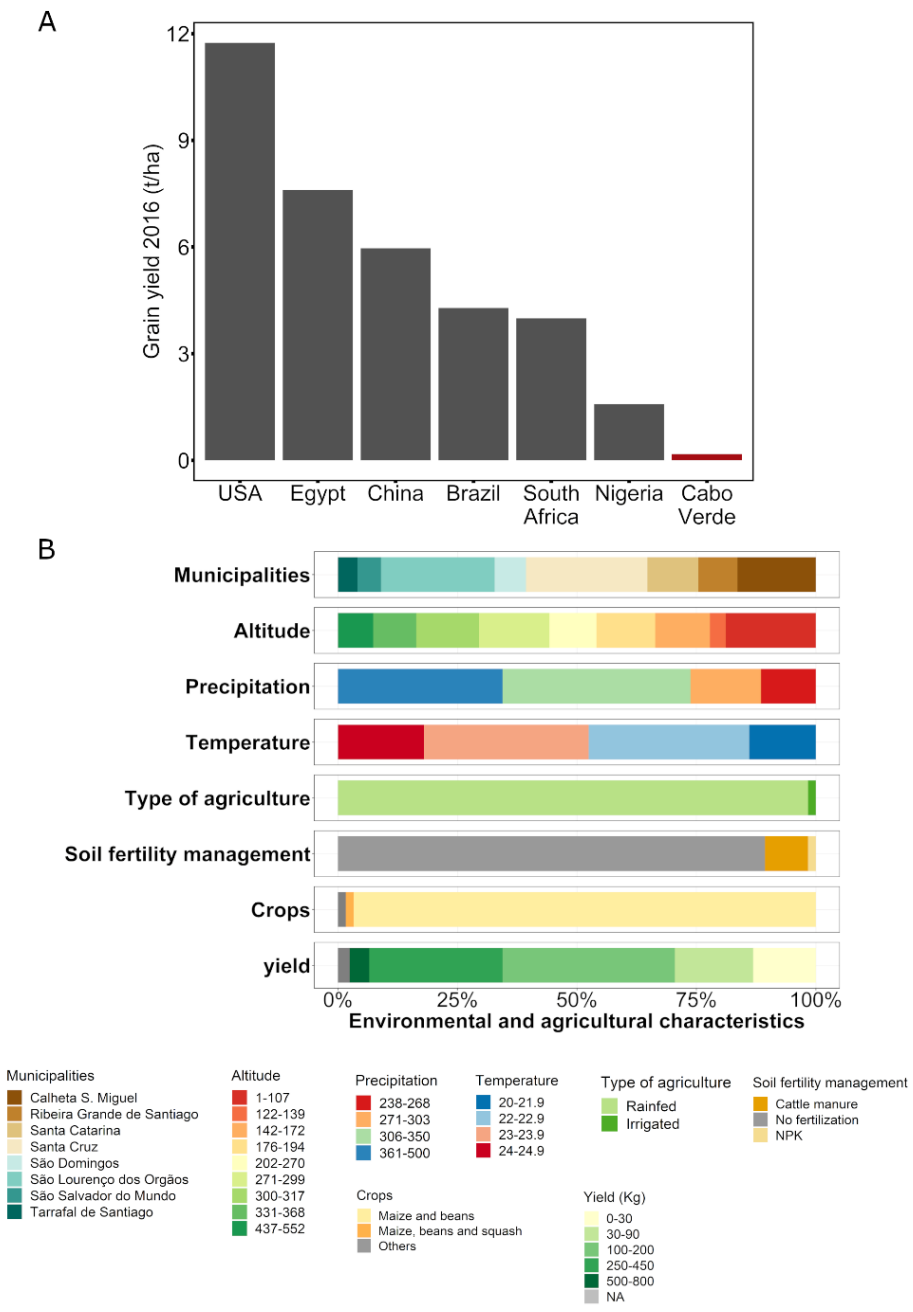


Figure 2: Maize production in Cabo Verde relies on natural resources and is vulnerable to environmental fluctuations. A) Bar graph of grain yield productivity in 2016 for three major producers in the world and Africa continent compared to Cabo Verde productivity. Data collected from FAOSTAT. **B)** Bar graph of geographic location, environmental conditions and agricultural

practices identified in the interview-based survey in Santiago Island. Different colours represent different factors identified in these interviews.

3.2 Water- and nutrient-related factors in the soil are the most variable in Santiago island

To have a deeper understanding of soil characteristics and environmental fluctuations that might affect maize yield in Santiago island, we downloaded from Soilgrids and Worldclim databases, information about some parameters with potential effect on crop performance. From soilgrids we retrieved data on soil pH, soil texture, bulk density, cation exchange, carbon and nitrogen composition, while from Worldclim we downloaded data on wind speed, aridity (Berdugo *et al.*, 2020), solar radiation, level of rainfall, water vapour pressure (López *et al.*, 2021), and minimum, maximum and average temperature (Singh *et al.*, 2021).

To identify the association between the parameters used, we have independently clustered pairwise correlation matrixes of the data from the two datasets. As expected, we found correlations among the soil properties parameters, some of which with high correlations (Figure 3A). We noticed that cluster 1 (Cl1) contains mainly parameters related to the soil physical properties, such as its texture and bulk density. The remaining clusters (Cl2, Cl3, and Cl4) included parameters related to the soil chemical properties, such as cation exchange capacity, soil organic carbon, and nitrogen composition (Figure 3A). Regarding the climatic parameters, we identified four clusters of highly correlated factors (Figure 3B). Precipitation and solar radiation are clustered in Cl1 and Cl2, whereas Cl3 contains all temperature and water vapour pressure parameters. Finally, Cl4 comprised wind speed and aridity factors (Figure 3C). All these results indicate that the parameters studied are redundant and that a certain level of duplicity exists in the information.

To remove redundancy in the environmental and soil properties parameters, we performed a Random Forest analysis to identify the most important factors in each cluster. Random Forest (RF) is a machine-learning algorithm that creates a collection of classification trees with binary divisions (Breiman, 2001).

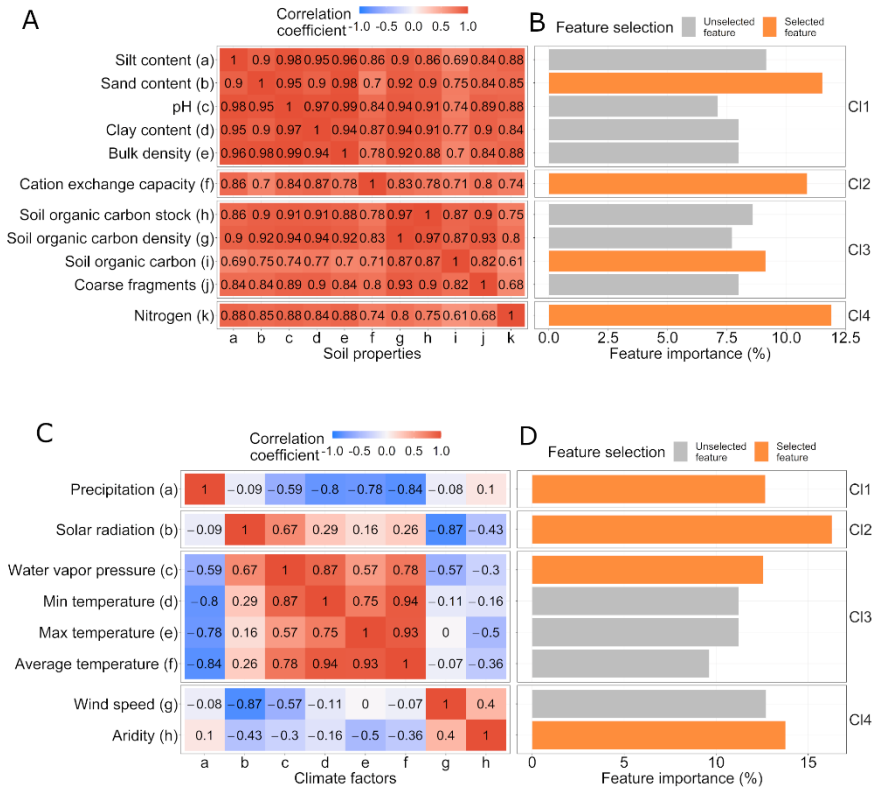


Figure 3: Climate and soil properties parameters identified in Santiago island regions. A) Heatmap showing the Pearson's correlation coefficient from each pairwise comparison between parameters identified from the Worldclim database. **B)** Bar graph showing the climate factor importance related to the regions in Santiago Island. In the figure, panels A and B have been hierarchically clustered according to the correlation coefficients from panel A to define clusters of highly correlated climate factors. **C)** Heatmap showing the Pearson's correlation coefficient from each pairwise comparison between the soil parameters extract from the Soilgrids database. **D)** Bar graph showing the soil properties' importance related to the regions in Santiago Island. Panels C and D have been hierarchically clustered according to the correlation coefficients from panel C to define the clusters of highly correlated soil factors within each cluster, factors with the highest importance value were selected as relevant to characterize the regions in Santiago island (orange bars in panel **B** and **D**).

In RF, the importance of each variable is determined by evaluating the decrease in prediction accuracy when the variable is randomly permuted (Wei *et al.*, 2010). Regarding the soil-related parameters, our RF analysis found that sand content, cation exchange capacity, soil organic carbon, and nitrogen

composition were the most important soil properties factors across Santiago regions (Figure 3B). By performing the same analysis for climate factors, we observed that precipitation, solar radiation, water vapour pressure, and aridity were the most important factors (Figure 3D).

These findings indicate that these factors are the ones that best integrate the fluctuations found across the different regions in Santiago island (Figure 4). They can be grouped into two groups: water-related (sand content, precipitation, solar radiation, water vapour pressure, and aridity) and nutrient-related (cation exchange, soil organic carbon, and nitrogen concentration) factors.

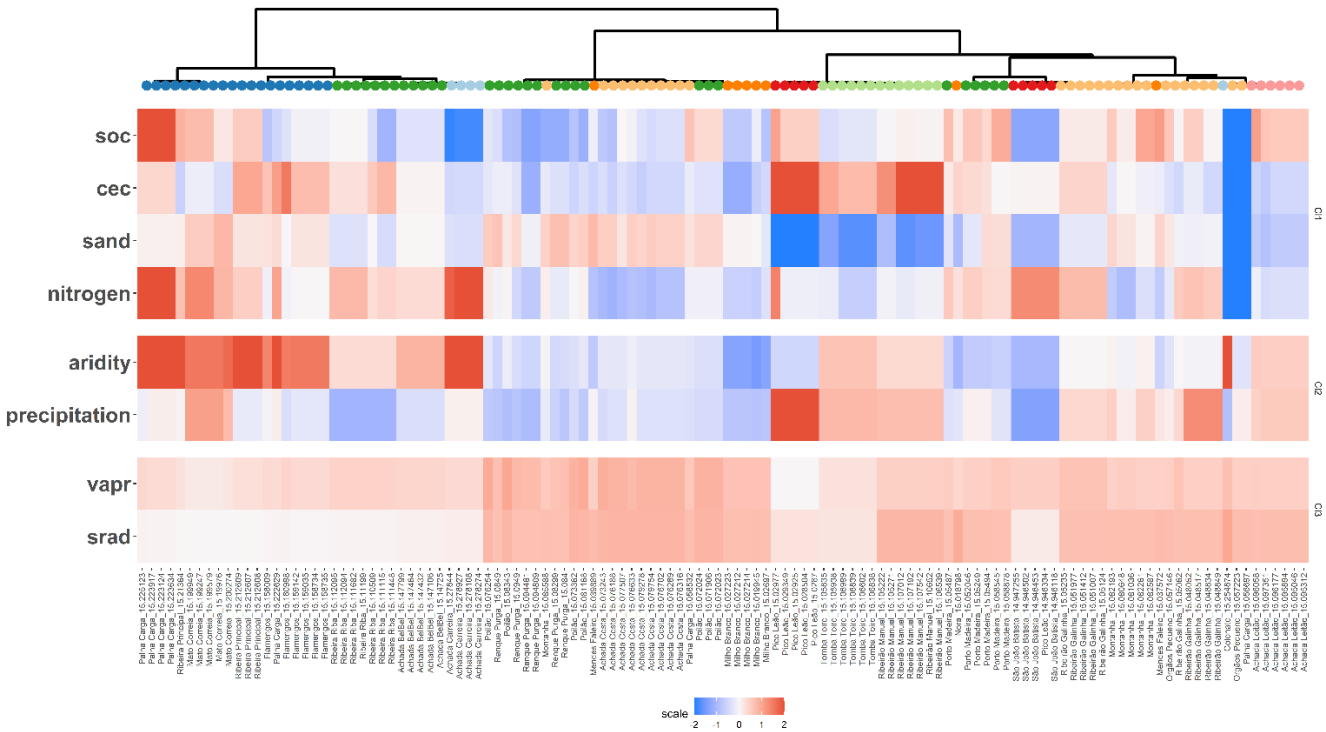


Figure 4: Water- and nutrient-related parameters integrate fluctuations found in the different regions in Santiago island. Heatmap showing the environmental and soil properties scaled values for selected factors that fully captured variability in Santiago regions. The values have been clustered according to the regions and their environmental and soil properties characteristics. Colours on the horizontal dendrogram tips represent the different municipalities in Santiago island, and each circle represent a farm.

3.3 Soil mineral nutrient content differs across Santiago island regions.

In a complex network such as the soil, the availability of nutrients for plants is influenced by other abiotic and biotic components of the network through intricated pathways that indirectly influence plant productivity and growth (Delgado-Baquerizo *et al.*, 2020; Delgado-Baquerizo *et al.*, 2013). Therefore, the amount of mineral nutrients available in the soil could function as an integrator of the effect of other components of the soil network, such as climate and soil parameters previously identified in this work. We used Inductively Coupled Plasma Mass Spectrometry (ICP-MS) to analyse the soil mineral content from 24 farms representing the environmental and soil variability found in Santiago Island (Figure 5). Principal component analysis (PCA) of the mineral nutrient profiles of the different soils showed that the mineral nutrient composition differs among municipalities (Figure 6A). We identified geographical location (municipalities) (Adonis: df=7, $R^2 = 0.313$, $p < 0.001$) and soil fertility management (Adonis: df=2, $R^2 = 0.073$, $p < 0.001$) as the main drivers of the observed variations in mineral nutrient accumulation.

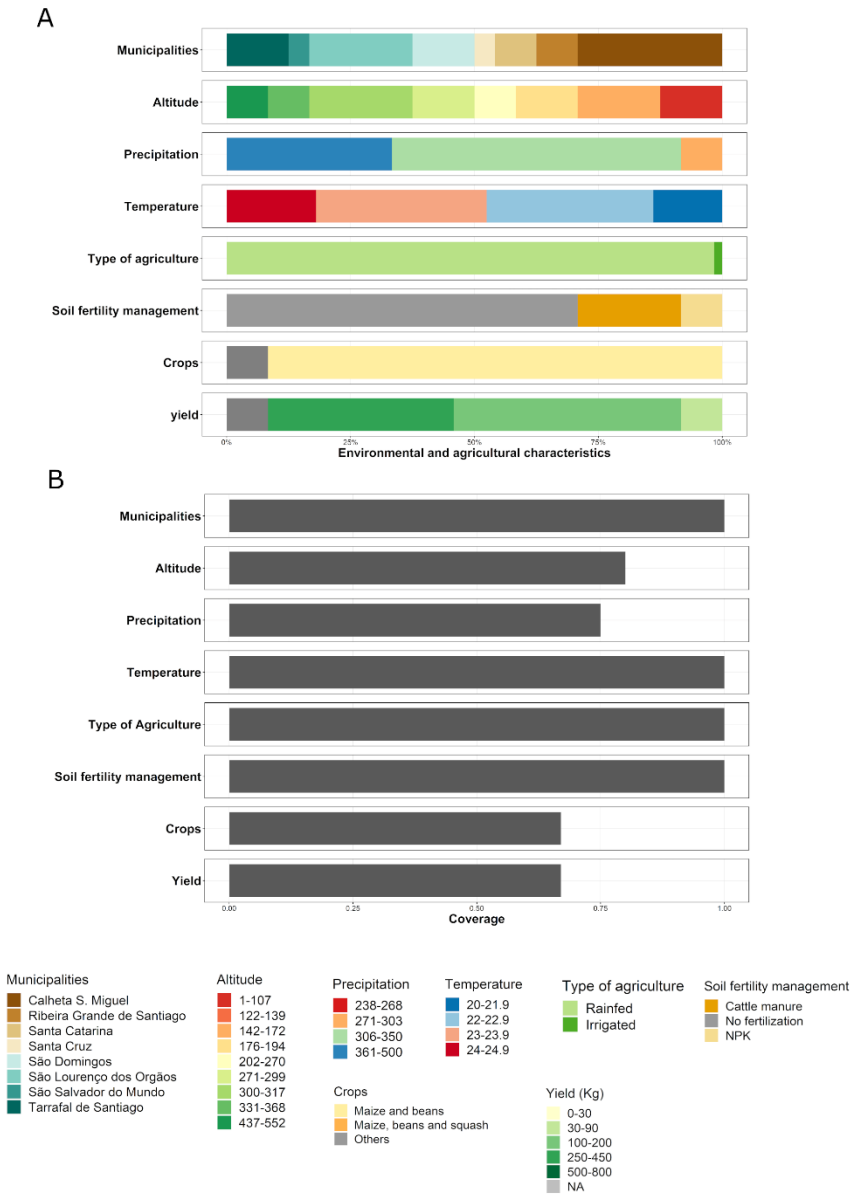


Figure 5: Selected soils cover all the parameters influencing maize yield identified in Santiago island. A) Bar graphs showing the geographic location, environmental, and agricultural practices in the selected farms. **B)** Bar graphs showing the coverage, in the selected farms, of the climate and soil parameters identified during the interviews.

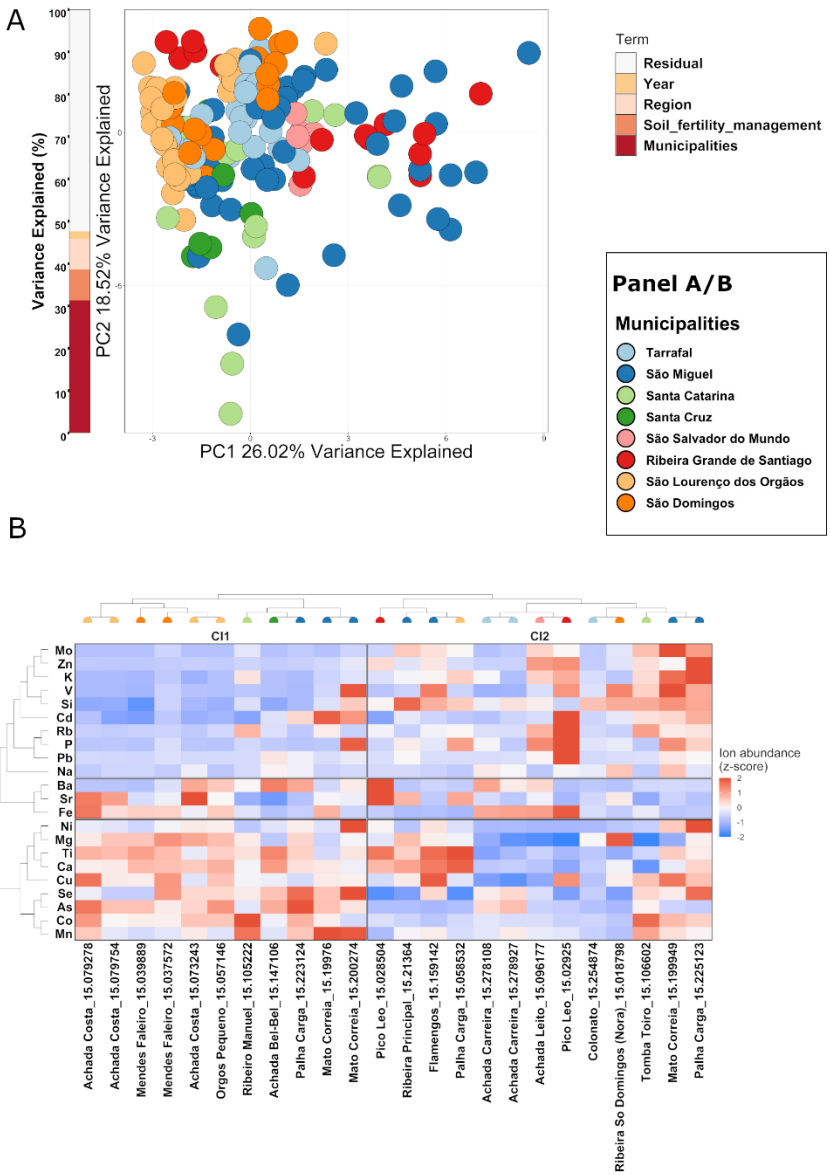


Figure 6: The soils collected from the selected farms in Santiago island differ in mineral content. A) Principal component analysis (PCA) based on Euclidean distance between mineral content across soils in Santiago Island. The bar graph to the left of the PCA represents the percentage of variance explained by statistically significant ($p<0.05$) variables in a PERMANOVA model. **B)** Heatmap showing the standardized mineral nutrient content in Santiago Island. Colours represent the different Municipalities on Santiago Island.

We also found that the mineral nutrient content in soils grouped in two main clusters of regions in Santiago Island (Figure 6B). This indicates that, generally, the soils on Santiago Island are poor, and there are no significant variations in mineral content throughout the island.

To further understand which ions are driving the differences between the two clusters of farms identified (Figure 7), we compared 22 ions between the different farms. We identified 16 mineral nutrients that varied significantly between the two clusters, 50% of them (eight nutrients) being essential nutrients for plants. For example, K, P, Zn, and Mo are present at lower concentrations in farms belonging to cluster 1 compared to farms belonging to cluster 2 (negative estimate values) (Figure 7). On the contrary, Mn, Ca, Mg, and Ni were enriched in farms belonging to cluster 1 (positive estimate values).

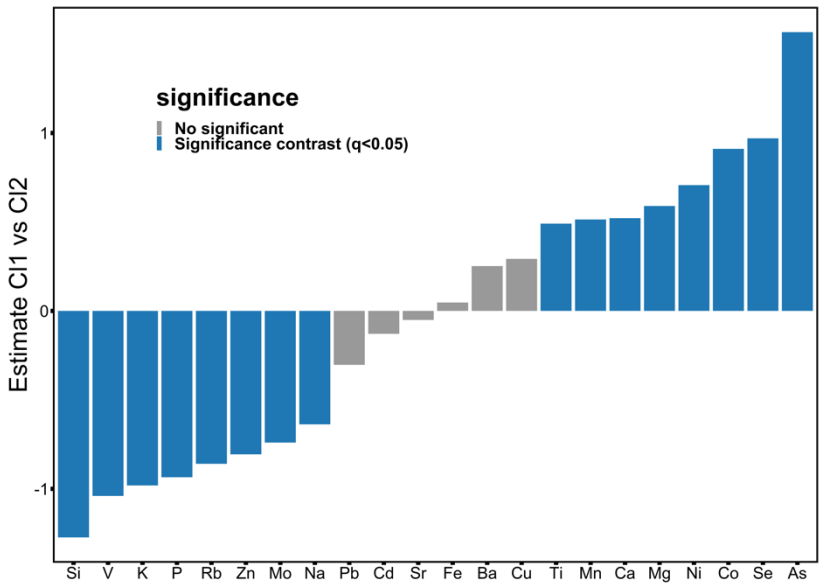


Figure 7: Mineral nutrient concentration drives differences between farms in Santiago island. Bar graph showing the estimated concentrations of ions in cluster 1 with respect to the concentrations of ions in cluster 2. Values that are significantly different (*q*-value < 0.05) are highlighted in blue.

Finally, to evaluate whether the climate and soil parameters identified here are associated with a mineral nutrient concentration in the soils, we correlated matrices of dissimilarities of climate and soil properties parameters (soil organic

carbon, cation exchange, sand content, nitrogen, aridity, precipitation, water vapour pressure, and solar radiation) and soil mineral content using a mantel test. We found a significant correlation between the climate and soil parameters with soil mineral content (Figure 8). This observation suggests that elements of the soil network that are controlled by the variability of the accumulation of mineral nutrients in the soil might also be controlled by the variation of climatic factors and other soil parameters. Thus, variations in soil mineral content also integrate fluctuation in all the previously identified climate and soil properties factors.

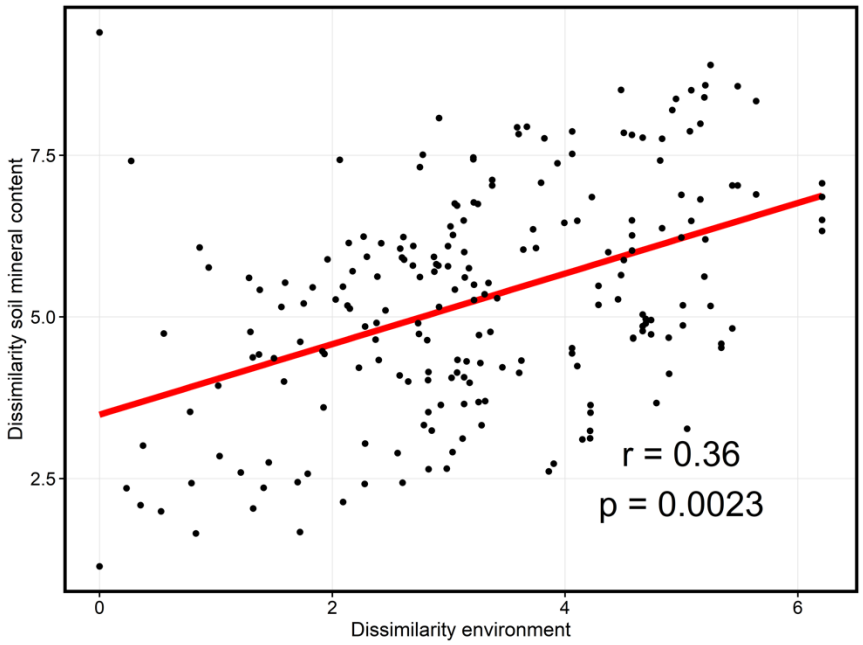


Figure 8: Soil mineral content is a good integrator of climate and soil properties. Correlation plot between Euclidean distance matrices calculated for soil mineral content and climate and soil properties in Santiago island. The r and p-value were calculated using Mantel tests.

3.4 Census of soil free-living bacterial composition in Santiago island

Since soil health and plant productivity depend on the interactions among abiotic and biotic factors within the soil matrix (Joswig *et al.*, 2021), we also explored the soil bacterial composition in Santiago island. To study the composition of soil-free-living bacterial communities, we analysed six replicates

per soil from each farm in two independent years (2017 and 2019). We generated a bacterial community profile for each sample via PCR amplification of the 16S ribosomal RNA (rRNA) gene targeting the V3-V4 regions, followed by Illumina sequencing. A total of 1,652,136 high-quality sequences were obtained from 152 samples. Then, we analysed high-quality reads with DADA2 and R and removed chimeric and organelle sequences to produce 18,951 amplicon sequence variants (ASVs).

We measured the α-diversity (within-sample diversity) and β-diversity (between-sample diversity) across the different farms. We estimated α-diversity by calculating the Shannon diversity index, and α-diversity inspection showed a quantitatively similar index between the different farms (ANOVA, *p*-value > 0.05) (Figure 9A), indicating that the farms have similar microbial diversity. As we expected, we found that the composition of soil bacterial communities differs across municipalities. Canonical analysis of principal coordinate (CAP) revealed significant differences between Santiago Island municipalities (Figure 9B). This pattern was corroborated by a permutational multivariate analysis of variance (PERMANOVA) with Bray-Curtis distances (adonis function in the R library vegan), which indicated that the factor municipalities explained major variation in bacterial community composition (adonis: coefficient of variation (R^2) = 0.13, *p*-value < 0.001) (Figure 9B).

To further explore soil microbial composition in each farm, we analysed the bacterial composition at the phylum level. The most abundant phyla detected were Acidobacteria, Actinobacteria and Proteobacteria (Figure 10A). Similarly to the previously described content in soil mineral nutrients, we found that the bacterial community also defined two main clusters of regions in Santiago Island (Figure 10A). Cluster 1 comprises farms belonging mainly to Ribeira Grande de Santiago, Tarrafal, Santa Catarina, São Domingos, São Lourenço dos Órgãos, and Santa Cruz. While cluster 2 comprises farms mainly belonging to São Miguel and São Salvador do Mundo (Figure 10A).

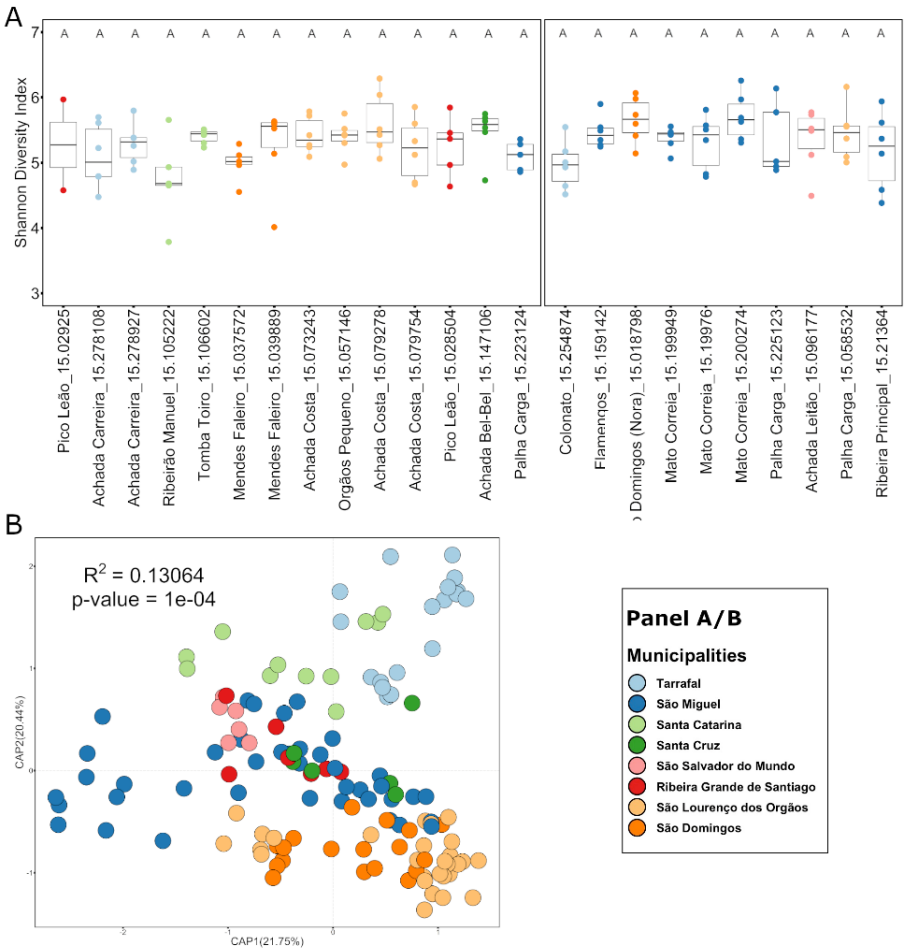


Figure 9: Free-living bacterial communities differ among municipalities on Santiago Island. **A)** Bacterial alpha diversity measured in soil across the 24 fields based on the Shannon index. Letters represent *post hoc* results based on the ANOVA model. **B)** Canonical analysis of principal component (CAP) based on Bray-Curtis distances between bacterial composition in soil samples harvested across 24 fields and two years (2017 and 2019). The R^2 and *p*-value inside the plot are derived from the PERMANOVA model and correspond to municipalities term. In graphs A and B, different colours represent different municipalities.

We applied a generalized linear model (GLM) to the count datasets to define which taxa at the genus level were different between farms. By contrasting cluster 1 against cluster 2 we could detect 30 unique genera, that were differentially abundant or depleted in the farms. Cluster 1 was enriched mainly in bacteria genera belonging to Proteobacteria, Gemmatimonadetes,

and Verrumicrobia. Whereas cluster 2 was enriched in genus belonging to Acidobacteria and Chloroflexi (Figure 10B). These data suggest that these groups of bacteria are driving the differences between the two groups of farms identified on the island. Therefore, the soil microbiome is a parameter to consider as part of strategies to increase maize yield in Cabo Verde.

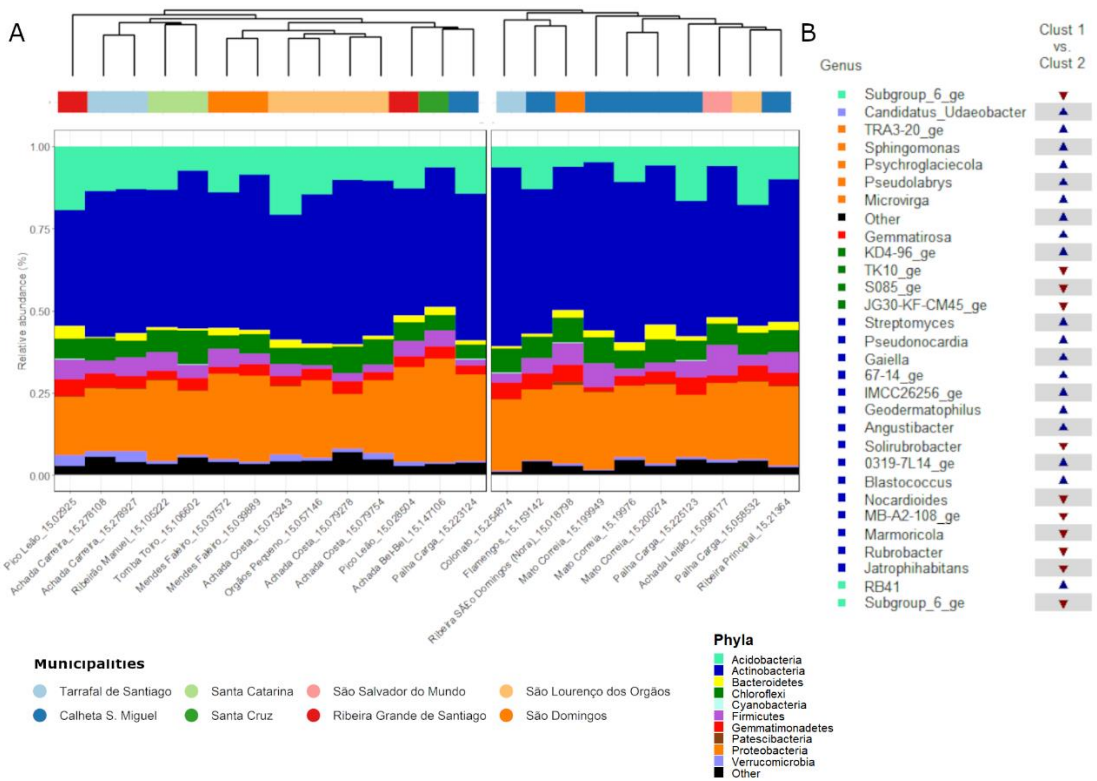


Figure 10: Free-living bacterial community structure in 24 maize fields on Santiago Island. A) Phyllum-level distribution of the soil microbiota in the 24 fields. The Bray-Curtis similarity-based dendrogram shows averaged bacterial composition for each farmer's field across two years (2017 and 2019). The total number of processed samples was 136, and only those with over 1,000 reads were used to calculate the average Bray-Curtis distance. Municipalities are indicated with coloured squares. B) Differential abundance analysis (phyla level) between soils collected from farms belonging to cluster 1 (Cl1) *versus* cluster 2 (Cl2). Triangles depict statistically significant differences (q-value < 0.05), with blue colour indicating genus enrichment in the farms belonging to Cl1, and red representing genus enrichment in farms from Cl2.

3.5 Field sites differ in soil mineral content and microbiome

The interaction among abiotic and biotic factors widely affects plant productivity (Pandey *et al.*, 2017). To disentangle the contribution of soil mineral content and soil microbiota to plant productivity in Santiago island, we conducted a field experiment in two geographically distant regions, Tarrafal and São Domingos (Figure 11A). The field sites were separated by 33.49 km, and showed similar climate characteristics (Figure 11B - Figure 11D). By selecting regions with similar climate characteristics, we aimed to also disentangle the climate effect and/or interactions with soil properties (physical, hydraulic, chemical and biological) on plant productivity.

To describe the soil properties in the two field sites, we analyse bulk density, soil texture, soil field capacity, soil permanent wilting point, soil pH, soil mineral content, and soil bacterial composition. Similarly to climate factors, we found that soil physical properties (bulk density and texture) and soil hydraulic properties (field capacity and permanent wilting point) were similar between regions (Kruskal-Wallis, $p>0.05$) (Figure 12A - Figure 12C). We examined differences in soil mineral content in the different soils using Inductively Coupled Plasma Mass Spectrometry (ICP-MS). Principal component analysis (PCA) of the mineral nutrient profiles showed that the fields are different in terms of their mineral nutrient composition (adonis: degrees of freedom (df) = 1; coefficient of determination (R^2) = 0.57; $P<0.001$) (Figure 12E).

To understand which ions are driving the differences between the two field sites, we compared the 22 ions previously described for the soil collection. We identified 13 mineral nutrients that varied significantly between the two fields (Figure 13A), where 46% of them (six nutrients) being essential to plants. For example, K and Fe present lower concentrations in São Domingos field site as compared to Tarrafal field sites (negative estimate values) (Figure 13A). On the contrary, Mo, Ca, Mg, and Ni were enriched in São Domingos field site (positive estimate values).

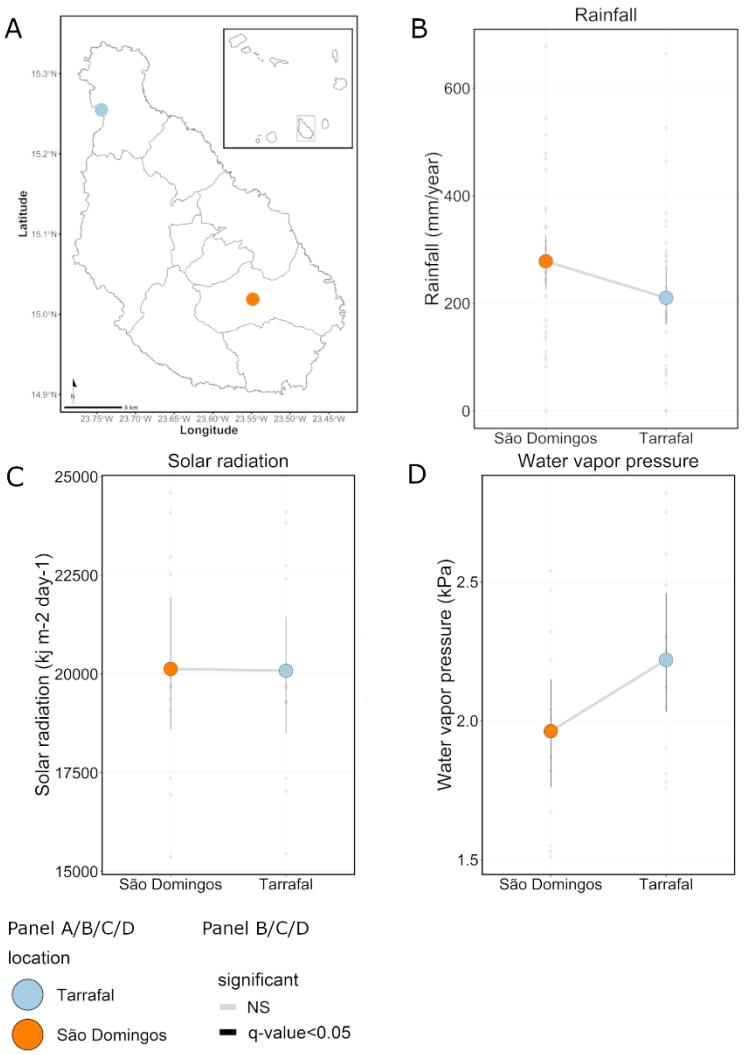


Figure 11: Geographical distant regions selected in Santiago island with similar climate characteristics. A) Map of the study region in Santiago Island, Cabo Verde. **B)** Average annual rainfall from 1984 to 2009 reported in the work by Neves, *et al.* (Neves *et al.*, 2017). **C)** Solar radiation and **D)** Water vapor pressure downloaded from Worlclim database. The line in grey in plot **B, C** and **D** denote no significance between regions (Kruskal-Wallis, $p<0.05$). The line connecting both points is the difference between São Domingos and Tarrafal.

We then investigated the bacterial composition in the two soils. In line with previous publications, we found that bacterial communities' composition differs among soils. Principal coordinates analysis (PCo) showed that the different soils

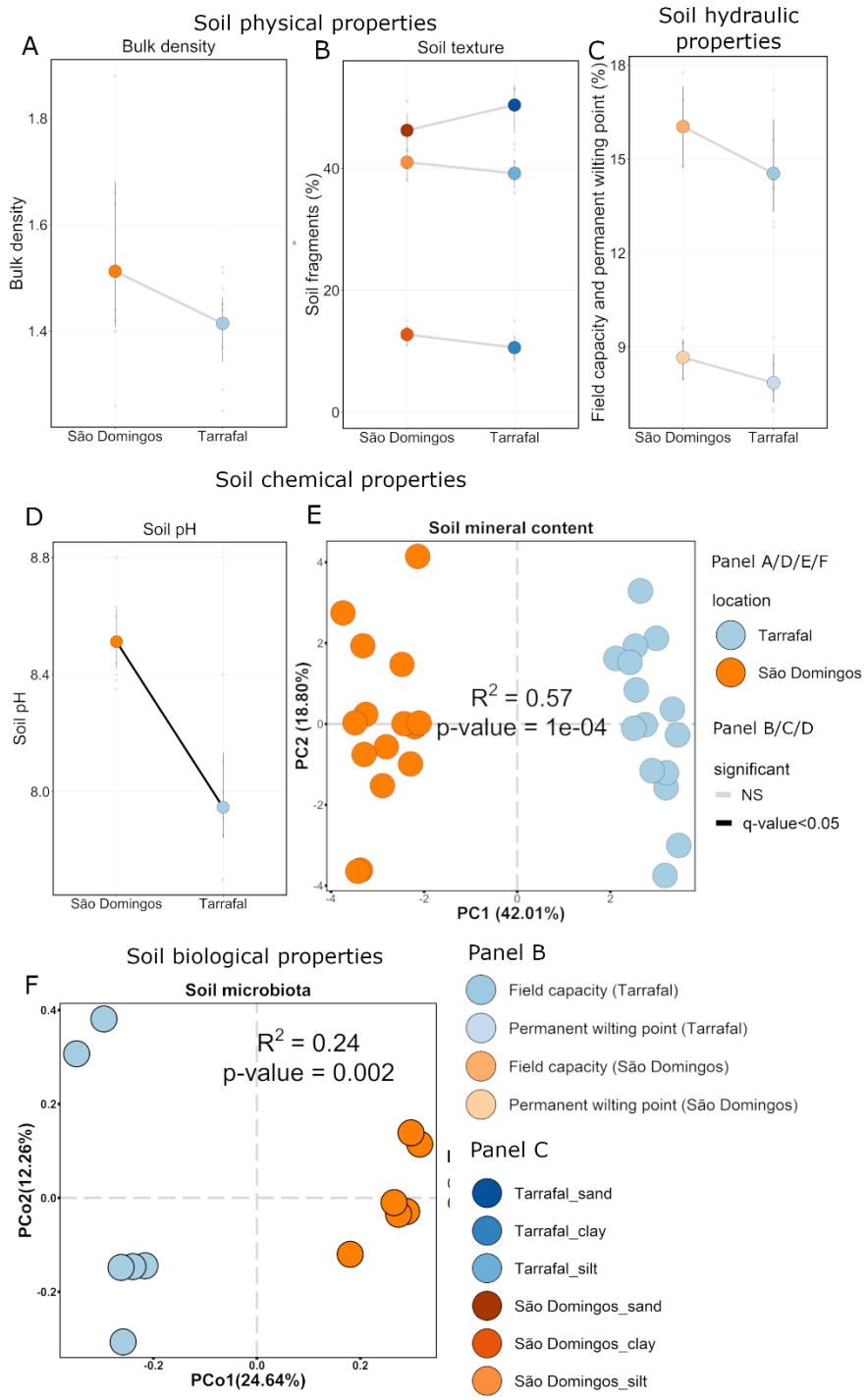


Figure 12: Abiotic and biotic soil properties variation in the two experimental field. A) Point plot based on measured bulk density in the experimental fields. The line connecting both

point is the difference between Sao Domingos and Tarrafal. **B)** Sand, clay and silt proportion measured in the soils. **C)** Calculated field capacity and permanent wilting point using the texture results**. D)** Measured soil pH for the experimental fields. **E)** Principal component analysis (PCA) based on Euclidean distances between soil mineral content in the two experimental fields. . The R^2 and p-value inside the plot are derived from the PERMANOVA model and correspond to the location term. **F)** Principal coordinate analysis (PCo) based on Bray-Curtis distances between bacterial composition in Sao Domingos and Tarrafal soil samples. The R^2 and p-value inside the plot are derived from the PERMANOVA model and correspond to the location term. The soil was collected in 2017 from 5 different 50x50 cm squares distributed across the experimental field.

carried different bacterial communities (adonis: degrees of freedom (df) = 1; coefficient of determination (R^2) = 0.24; P<0.01) (Figure 12F). To identify bacterial community genera driving this variation between the two field sites, we applied a generalized linear model (GLM) and contrasted São Domingos microbial count dataset against Tarrafal microbial count dataset. We found 44 unique genera that were differentially abundant or depleted in the field sites (Figure 13B). For these 44 genera, a total of 33 genera (75%) were enriched in São Domingos field site while the remaining 11 genera (25%) were enriched in Tarrafal field site (Figure 13B). Despite these significant differences in the enrichment patterns, both field sites were enriched in bacteria genera belonging to Acidobacteria, Actinobacteria, Chloroflexi, Proteobacteria, Gemmatimonadetes, and Verrumicrobia (Figure 13B).

Taken together, these results demonstrate that our field site selection represents similar climate characteristics and different levels of variability in mineral content and bacterial composition. Thus, these two locations offer opportunities for the study of soil mineral content and bacterial composition variation in maize productivity while minimising the influence of the climate factors.

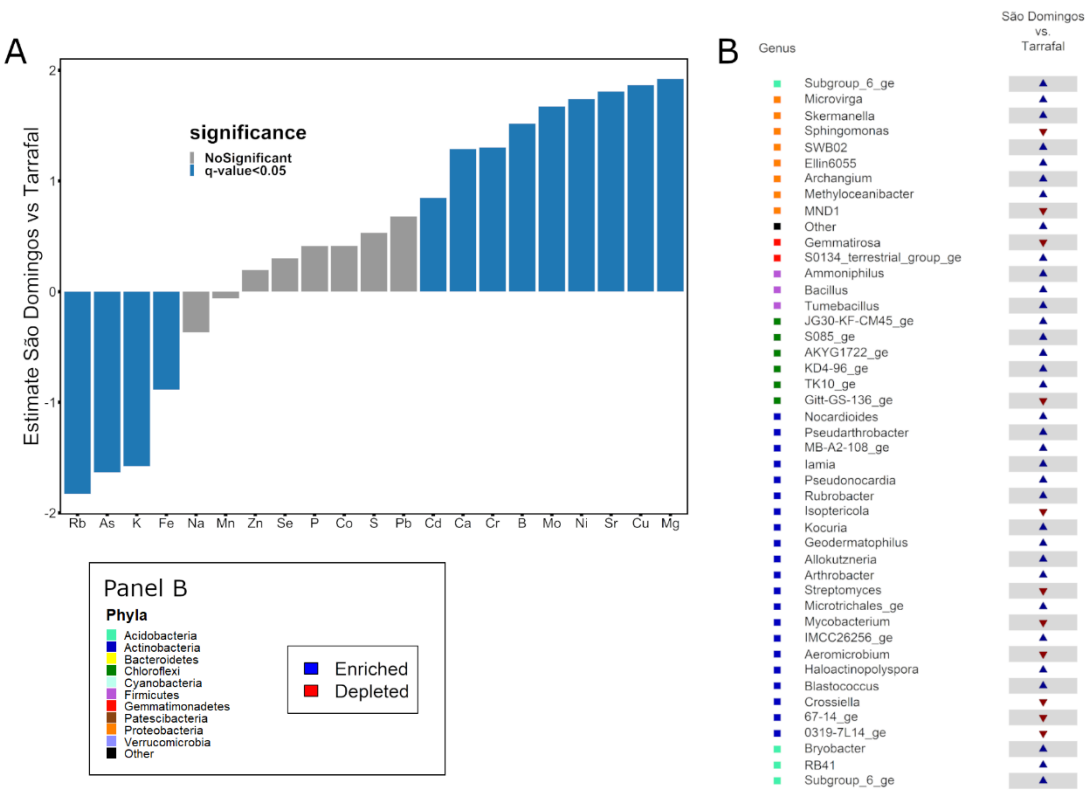


Figure 13: Mineral nutrient concentration and bacterial genera variation in the field sites. A) Bar graph showing the estimated concentrations of ions in Sao Domingos *versus* Tarrafal. Values that are significantly different (*q*-value < 0.05) are highlighted in blue. **B)** Soil enrichment patterns in the field sites. Each row represents a unique genera. The squares are coloured by phyla. The triangles represented enriched (blue) and depleted (red) genera coloured by log2-fold derived from a fitted generalized linear model comparing Sao Domingos and Tarrafal soils. The plot depicted only genera with false-discovery-rate (FDR)-corrected q-value < 0.05.

3.6 Soil mineral content and microbiota variations effect on maize productivity

To link the soil mineral content and microbial variability to maize productivity, we conducted a field experiment in the two selected field sites in Santiago island. We investigated the within-site effects of the variation of these two factors by imposed short-term changes in soil mineral content and bacterial composition through two agricultural practices: fertilization and cropping system. Each field was divided into two subplots and six blocks each, where we mimic the soil fertility management and cropping system identified in our farmers surveys. One of the subplots where supply with ammonium sulphate, here designated as inorganic fertilizer, and the other subplot received no fertilizer. Then in each block, we grew landrace maize in monocrop and in intercrop with cowpea (*Vigna unguiculata*). As a control, we also grew cowpea in monocrop.

First, at the harvest stage we quantified the plant height, as a proxy for plant productivity, in each blocks (Figure 14). We found that in general maize plants grows higher in São Domingos field as compared to plants growth in Tarrafal (Figure 14). As we expected, plants grown with inorganic fertilizer produce higher plants in both soils (Figure 14). Variation partitioning using a linear mixed model conducted by location, treatment, and cropping system indicates that these factors accounted for a small variation in plant height (location, R^2=0.0072; treatment, R^2=0.011; cropping system, R^2=0.014). Interestingly, we found that a substantial proportion of variation in plant height (variation partitioning, R^2=0.96) was explained by other factors that were not captured by the location and the agricultural practices. These results suggest that complex chemical conditions, soil microbes or a combination of these factors can influence plant height.

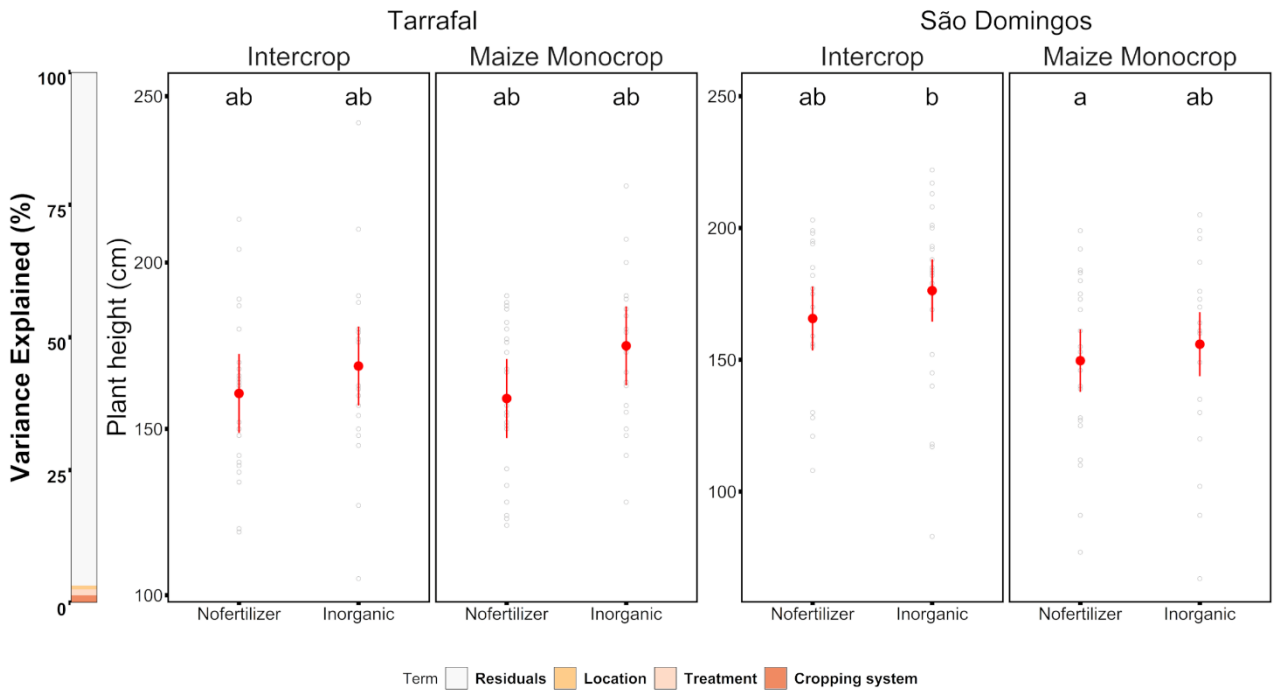


Figure 14: Plant height growth in the two experimental field. Pointrange plot showing the plant height in each block in the two experimental field. The letters indicate statistical significance corresponding to ANOVA with Tukey's HSD *post hoc* test (α = 0.05). The bar graph to the left of the pointrange plot depicts the percentage of variability explained by statistically significant ($p<0.05$) terms in the variance partitioning using linear mixed model.

4 Discussion

Dryland areas are particularly vulnerable to climate change and soil degradation (Middleton and Sternberg, 2013; Maestre *et al.*, 2016). This vulnerability drives multiple changes in diverse structural and functional ecosystem attributes such as nutrient cycling, microbial communities, and plant productivity (Pointing and Belnap, 2012). Here, we have characterized the climate, soil nutrient, and microbial composition in Santiago dryland farms. We have shown that soil mineral content is a good integrator of water-related (sand content, precipitation, solar radiation, water vapour pressure, and aridity) and nutrient-related (cation exchange, soil organic carbon, and nitrogen concentration) factors. This is consistent with previous study, that shows that continues increase in aridity in the global ecosystem abrupt decline soil nutrient composition (Berdugo *et al.*, 2020). Since the composition of the soil microbiota is often modulated by the availability of nutrients in the soil (Fierer *et al.*, 2012; Sheldrake *et al.*, 2018), we cannot exclude the possibility that the soil mineral content variation among the farms could indicate as well that the soil microbiota differs across farms. Further, we explore how these variations could affect maize productivity in the experimental field independent of climate factors. Our data suggest that these variations or the combination between them contribute for the maize differences in productivity.

In conclusion, this work characterizes a dryland ecosystem where the effect of the soil mineral content and bacterial composition on productivity was isolated from the climate factors. The trend identified here provides a framework for the future possibility of exploring to which extent variation in the soil environment could help plants cope with the constraints in dryland areas. Therefore, to reach this final goal, further studies are required to understand how the interactions of the external factors operate at the mechanistic level in plants.

5 References

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M. (1998) Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome*, **300**, D05109.

Allen, R.G., Pereira, L.S., Smith, M., Raes, D. and Wright, J.L. (2005) FAO-56 dual crop coefficient method for estimating evaporation from soil and application extensions. *Journal of Irrigation and Drainage Engineering*.

Anon (2016) The Python Language Reference Release 3.5.1 Guido van Rossum and the Python development team.

Berdugo, M., Delgado-Baquerizo, M., Soliveres, S., Hernández-Clemente R., Zhao Y., Gaitán J.J., Gross N., Saiz H., Maire V., Lehmann A., Rillig M.C., Solé R.V., Maestre F.T. (2020) Global ecosystem thresholds driven by aridity. *Science (1979)*, **367**, 787–790.

Bilskie, J. and Campbell Scientific (2001) Soil water status: content and potential.

Breiman, L. (2001) Random Forests. *Machine Learning 2001 45:1*, **45**, 5–32.

Briggs, L.J. and Shantz, H.L. (1912) The Wilting Coefficient and Its Indirect Determination. *Botanical Gazette*.

Bruggen, A.H.C. van, Gamliel, A. and Finckh, M.R. (2016) Plant disease management in organic farming systems. *Pest Manag Sci*, **72**, 30–44.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A. and Holmes, S.P. (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*.

Delgado-Baquerizo, M., Maestre, F.T., Gallardo, A., Bowker M.A., Wallenstein M.D., Quero J.L., Ochoa V., Gozalo B., García-Gómez M., Soliveres S., García-Palacios P., Berdugo M., Valencia E., Escolar C., Arredondo T., Barraza-Zepeda C., Bran D., Carreira J.A., Chaieb M., Conceição A.A., Derak M., Eldridge D.J., Escudero A., Espinosa C.I., Gaitán J., Gatica M.G., Gómez-González S., Guzman E., Gutiérrez J.R., Florentino A., Hepper E., Hernández R.M., Huber-Sannwald E., Jankju M., Liu J., Mau R.L., Miriti M., Monerris J., Naseri K., Noumi Z., Polo V., Prina A., Pucheta E., Ramírez E., Ramírez-Collantes D.A., Romão R., Tighe M., Torres D., Torres-Díaz C., Ungar E.D., Val J., Wamiti W., Wang D., Zaady E. (2013) Decoupling of soil nutrient cycles as a function of aridity in global drylands. *Nature 2013 502:7473*, **502**, 672–676.

Delgado-Baquerizo, M., Reich, P.B., Trivedi, C., Delgado-Baquerizo M., Reich P.B., Trivedi C., Eldridge D.J., Abades S., Alfaro F.D., Bastida F., Berhe A.A., Cutler N.A., Gallardo A., García-Velázquez L., Hart S.C., Hayes P.E., He J.Z., Hseu Z.Y., Hu H.W., Kirchmair M., Neuhauser S., Pérez C.A., Reed S.C., Santos F., Sullivan B.W., Trivedi P., Wang J.T., Weber-Grullon L., Williams M.A., Singh B.K. (2020) Multiple elements of soil biodiversity drive ecosystem functions across biomes. *Nature Ecology & Evolution 2020 4:2*, **4**, 210–220.

Duku, C., Zwart, S.J. and Hein, L. (2018) Impacts of climate change on cropping patterns in a tropical, sub-humid watershed. *PLoS One*, **13**, e0192642.

Fick, S.E. and Hijmans, R.J. (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, **37**, 4302–4315.

Fierer, N., Lauber, C.L., Ramirez, K.S., Zaneveld, J., Bradford, M.A. and Knight, R. (2012) Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME J*, **6**, 1007–17.

Grossman, R.B. and Reinsch, T.G. (2002) 2.1 Bulk Density and Linear Extensibility. In *Methods of soil analysis: Part 4physical methods,*. pp. 201–228.

Hengl, T., Jesus, J.M. de, Heuvelink, G.B.M., Ruiperez Gonzalez M., Kilibarda M., Blagotić A., Shangguan W., Wright M.N., Geng X., Bauer-Marschallinger B., Guevara M.A., Vargas R., MacMillan R.A., Batjes N.H., Leenaars J.G., Ribeiro E., Wheeler I., Mantel S., Kempen B. (2017) SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One*, **12**, e0169748.

Hijmans, R.J. (2014) Introduction to the ' raster ' package. *R: A Language and Environment for Statistical Computing*.

Joswig, J.S., Wirth, C., Schuman, M.C., Kattge J., Reu B., Wright I.J., Sippel S.D., Rüger N., Richter R., Schaepman M.E., van Bodegom PM, Cornelissen JHC, Díaz S, Hattingh WN, Kramer K, Lens F, Niinemets Ü, Reich PB, Reichstein M, Römermann C, Schrodt F, Anand M, Bahn M, Byun C, Campetella G, Cerabolini BEL, Craine JM, Gonzalez-Melo A, Gutiérrez AG, He T, Higuchi P, Jactel H, Kraft NJB, Minden V, Onipchenko V, Peñuelas J, Pillar VD, Sosinski Ê, Soudzilovskaia NA, Weiher E, Mahecha MD. (2021) Climatic and soil factors explain the two-dimensional spectrum of global plant trait variation. *Nature Ecology & Evolution 2021 6:1*, **6**, 36–50.

Lian, T., Mu, Y., Jin, J., Ma, Q., Cheng, Y., Cai, Z. and Nian, H. (2019) Impact of intercropping on the coupling between soil microbial community structure, activity, and nutrient-use efficiencies. *PeerJ*, **2019**, e6412.

López, J., Way, D.A. and Sadok, W. (2021) Systemic effects of rising atmospheric vapor pressure deficit on plant physiology and productivity. *Glob Chang Biol*, **27**, 1704–1720.

Maestre, F.T., Eldridge, D.J., Soliveres, S., Kéfi S, Delgado-Baquerizo M, Bowker MA, García-Palacios P, Gaitán J, Gallardo A, Lázaro R, Berdugo M. (2016) Structure and Functioning of Dryland Ecosystems in a Changing World. *https://doi.org/10.1146/annurev-ecolsys-121415-032311*, **47**, 215–237.

Marcinkowski, P. and Piniewski, M. (2018) Effect of climate change on sowing and harvest dates of spring barley and maize in Poland**. *Int. Agrophys*, **32**, 265–271.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*, **17**, 10–12.

Mavrodi, O. v., McWilliams, J.R., Peter, J.O., Berim A, Hassan KA, Elbourne LDH, LeTourneau MK, Gang DR, Paulsen IT, Weller DM, Thomashow LS, Flynt AS, Mavrodi DV. (2021) Root Exudates Alter the Expression of Diverse Metabolic, Transport, Regulatory, and Stress Response Genes in Rhizosphere Pseudomonas. *Front Microbiol*, **12**, 651282.

Middleton, N.J. and Sternberg, T. (2013) Climate hazards in drylands: A review. *Earth Sci Rev*, **126**, 48–57.

Moitinho, M.A., Bononi, L., Souza, D.T., Melo, I.S. and Taketani, R.G. (2018) Bacterial Succession Decreases Network Complexity During Plant Material Decomposition in Mangroves. *Microb Ecol*, **76**, 954–963.

Munson, B.R., Young, D.F. and Okiishi, T.H. (1995) Fundamentals of fluid mechanics. *Oceanographic Literature Review*, **10**, 831.

Neves, D.J.D., Silva, V. de P.R., Almeida, R.S.R., Sousa, F. de A.S. de and Silva, B.B. da (2017) General aspects of the climate in the Cabo Verde archipelago. *Ambiência*, **13**, 59–73.

Oksanen, J. (2007a) vegan : Community Ecology Package. R package version 1.8-5.

Oksanen, J. (2007b) vegan : Community Ecology Package. R package version 1.8-5.

Pandey, P., Irulappan, V., Bagavathiannan, M. v. and Senthil-Kumar, M. (2017) Impact of combined abiotic and biotic stresses on plant growth and avenues for crop improvement by exploiting physio-morphological traits. *Front Plant Sci*, **8**, 537.

Pointing, S.B. and Belnap, J. (2012) Microbial colonization and controls in dryland systems. *Nature Reviews Microbiology 2012 10:8*, **10**, 551–562.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res*, **41**.

Rai, R.K., Singh, V.P., Upadhyay, A., Rai, R.K., Singh, V.P. and Upadhyay, A. (2017) Soil Analysis. *Planning and Evaluation of Irrigation Projects*, 505–523.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane AC, Müller C, Arneth A, Boote KJ, Folberth C, Glotter M, Khabarov N, Neumann K, Piontek F, Pugh TA, Schmid E, Stehfest E, Yang

H, Jones JW. (2014) Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proc Natl Acad Sci U S A*, **111**, 3268–3273.

Sheldrake, M., Rosenstock, N.P., Mangan, S., Revillini D, Sayer EJ, Olsson PA, Verbruggen E, Tanner EVJ, Turner BL, Wright SJ. (2018) Responses of arbuscular mycorrhizal fungi to long-term inorganic and organic nutrient addition in a lowland tropical forest. *ISME Journal*.

Shojaei, S., Kalantari, Z. and Rodrigo-Comino, J. (2020) Prediction of factors affecting activation of soil erosion by mathematical modeling at pedon scale under laboratory conditions. *Scientific Reports 2020 10:1*, **10**, 1–12.

Singh, S.K., Reddy, V.R., Devi, M.J. and Timlin, D.J. (2021) Impact of water stress under ambient and elevated carbon dioxide across three temperature regimes on soybean canopy gas exchange and productivity. *Scientific Reports 2021 11:1*, **11**, 1–13.

Snyder, R.L., Orang, M., Matyac, S. and Grismer, M.E. (2005) Simplified estimation of reference evapotranspiration from pan evaporation data in California. *Journal of Irrigation and Drainage Engineering*.

Stott, P. (2016) How climate change affects extreme weather events: Research can increasingly determine the contribution of climate change to extreme events such as droughts. *Science (1979)*, **352**, 1517–1518.

Teixeira, E.I., Ruiter, J. de, Ausseil, A.G., Daigneault, A., Johnstone, P., Holmes, A., Tait, A. and Ewert, F. (2018) Adapting crop rotations to climate change in regional impact modelling assessments. *Sci Total Environ*, **616–617**, 785–795.

Wei, C.L., Rowe, G.T., Briones, E.E., Boetius A, Soltwedel T, Caley MJ, Soliman Y, Huettmann F, Qu F, Yu Z, Pitcher CR, Haedrich RL, Wicksten MK, Rex MA, Baguley JG, Sharma J, Danovaro R, MacDonald IR, Nunnally CC, Deming JW, Montagna P, Lévesque M, Weslawski JM, Wlodarska-Kowalczuk M, Ingole BS, Bett BJ, Billett DS, Yool A, Bluhm BA, Iken K, Narayanaswamy BE. (2010) Global Patterns and Predictions of Seafloor Biomass Using Random Forests. *PLoS One*, **5**, e15323.

Zheng, W., Gong, Q., Zhao, Z., Liu, J., Zhai, B., Wang, Z. and Li, Z. (2018) Changes in the soil bacterial community structure and enzyme activities after intercrop mulch with cover crop for eight years in an orchard. *Eur J Soil Biol*, **86**, 34–41.

Zhou, X., Yu, G. and Wu, F. (2011) Effects of intercropping cucumber with onion or garlic on soil enzyme activities, microbial communities and cucumber yield. *Eur J Soil Biol*, **47**, 279–287.

Zomer, R.J., Trabucco, A., Bossio, D.A. and Verchot, L. v. (2008) Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agric Ecosyst Environ*, **126**, 67–80.

Chapter III

The bacterial communities inhabiting individual leaves may contribute to their developmental program

This chapter was adapted from:

Custódio, V., Salas-González, I., Gopaulchan, D., Flis, P., Gao, Y., Amórós, R., Moreno, A., Oliveira, M.M., Salt, D. and Castrillo, G. (2023), Individual leaf microbiota tunes a genetic regulatory network to promote leaf growth in maize. (*In preparation)*.

Author's contribution

Custódio, V., Salas-González, I., and Castrillo, G. conceptualize the project and perform data curation. Custódio, V. and Salas-González, I., formally analyse the data. Custódio, V., Salas-González, I., Gopaulchan, D., Flis, P., Gao, Y., Amórós, R., Moreno, A., and Castrillo, G. perform the investigation. Custódio, V., and Castrillo, G wrote the original draft. Custódio, V., Oliveira, M.M., Salt, D. and Castrillo, G. Contributed to the text and corrected the different versions of the article.

Abstract

Plants integrate complex regulatory networks to ensure the adaptative growth of roots and leaves. These networks combine internal signals and external cues, such as nutrient availability and plant-colonising microorganisms – plant microbiota. The extent to which microbial communities, nutrient availability and internal signalling are coordinated to ensure adaptative growth of plant organs remains poorly understood. To study root and leaf responses to the different environmental cues, we grown four maize genotypes in three selected soils with different mineral nutrient contents and bacterial compositions. While root growth of the different genotypes was not affected by soil variability, leaves differentially integrated the soil effect. To understand the factors modulating the development of individual leaves, we investigated their mineral nutrient content and microbiome composition. We demonstrated that the soil effect on leaf development results from changes in the concentration of individual mineral elements impacting distinct leaf-associated bacterial communities.

1 Introduction

In nature, plant development results from the combined action of nutrient availability, energy status, cell growth and expansion, and microbial communities that colonize their organs (de Wit *et al.*, 2016; Durán *et al.*, 2018; Vega *et al.*, 2019). Analogous to other biological systems, including animals, plants harbour populations of microbes functionally adapted to colonize their different organs, enhancing their function and, consequently, the development of the entire organism. For example, seedlings of eucalyptus grown in the glasshouse under water deficit conditions that cause partial tree defoliation showed enhanced formation of new leaves when inoculated with two strains of *Pseudomonas* (Chaín *et al.*, 2020). Therefore, an ongoing question in studies of plant-microbe interactions is: what is the basis of crosstalk between plant and microbiota, and how do plants integrate the microbiota effect into their developmental program.

Given the fact that plants assemble distinct microbial communities in roots and leaves (Bai *et al.*, 2015), we hypothesized that root and individual leaves within a plant could represent differential ecological niches recruiting distinct bacteria-resident communities. These organ-specific communities of microbes could contribute to the growth plasticity of plant organs.

Here, we built a generalized framework to study the growth of roots and leaves in plants grown in three soils with different mineral nutrient content and bacterial composition. We assessed whether mineral content and bacterial composition could explain the leaf and root growth variations observed in these plants.

2 Materials and methods

2.1 Soil variability effect on maize ionome and microbiome composition

2.1.1 Soil collection

We collected two arable soils and a natural soil from three different geographical locations to evaluate their effect on plant growth. Arable soils were collected from the INIDA stations: (i) São Domingos, Cabo Verde, a semi-arid region (331 mm of accumulated rainfall) located in the central area of Santiago island (15º1'7.6728''N, 23º32'54.3372''W); and (ii) Tarrafal, Cabo Verde, an arid region (196 mm of accumulated rainfall) located in the northern coastal area of the same island (15º15'17.5468''N, 23º44'38.9256''W). The natural soil was collected from a farm located at Sutton Bonington Campus, University of Nottingham, UK (+52° 49′ 59.75″N, −1° 14′ 56.62″W). Before soil collection, all tools were first washed with water and then disinfected by immersion in 70 % ethanol for 10 minutes. Before collecting the soil samples, the upper soil layer containing local vegetation (5-10 cm deep, approximately) was discarded, and the soil collected was stored in clean plastic boxes and kept at 4 ºC until use.

2.1.2 Soil preparation

Collected soils were allowed to dry at room temperature for one week, in plastic trays, and then sifted using an ethanol-disinfected 2 mm sieve. Soils were mixed with autoclaved sand (pavior sand) in a proportion of 2:1 (v/v) to increase the soil drainage and decrease soil compaction. Sterile pots, with a square of autoclaved Miracloth tissue (Millipore) at the bottom, were weighted and filled with 300 g of each soil mixture and used to grow the maize plants.

2.1.3 Plant growth conditions – the nutrient effect

To determine the soil effect on plant leaf ionome and microbiome composition, we used a collection of maize (*Zea mays*) genotypes (B73, Matuba, CVSt and Zm523) expressing different levels of adaptation to nutrient and water content in soils (Thierfelder *et al.*, 2016; Chen *et al.*, 2012).

To determine the 50% of soil water content (SWC), we calculated the amount of water required for each soil. Briefly, five 9x9 cm pots with miracloth in the bottom were individually weighted using a bench scale (Mettler Toledo, Switzerland). Then, the pots were filled with 300 g of the soil mixture described above. In parallel, 100 g of the soil mixture was dried in an oven at 105 ºC for 24 h and used to determine the total soil dry weight per pot (m_{dry}). Pots were watered to saturation, covered with a plastic film to reduce water evaporation, and allowed to drain for 30 min. We weighted individual pots to determine the weight at full water capacity (PW_{FC}). The total amount of water retained per pot at full watered capacity was calculated using the following formula: m_{water} = (PW_{FC} – (pot and mesh weight)) - m_{dry}. Finally, we calculated the water weight per pot (PTW) to reach 50% SWC, using the following formula: PTW = (m_{water}/m_{dry} x (50/100)) x m_{dry} + (pot and mesh weight) + m_{dry} (Sapeta *et al.*, 2013; Bilskie and Campbell Scientific, 2001).

All seeds were surface-sterilized with 70 % bleach, 0.2 % Tween-20 for 8 min, followed by three rinses with sterile distilled water to eliminate any seed-borne microbe at the seed surface. Before sowing, seeds were placed in 9 cm Petri dishes with 10 mL of sterile water and incubated overnight at 30 ºC, in the dark, to embed and promote germination. Maize seeds were then sown in the soil mixtures prepared as described above using sterile tweezers, and as controls, we used unplanted pots (“bulk soil”).

Pots were randomised using a true random generator (random.org), and trays were reshuffled daily in the growth chamber. All pots, including controls, were watered daily with autoclaved Milli-Q water (18.2 MΩ cm; Merck Millipore water purification system) using a pump-action pressure spray bottle (Hozelock, Ltd, England).

Plants were grown in a growth chamber under 16-h light/8-h dark at 26 ºC day/24 ºC night, using a 500 µE m^{-2} s^{-1} photosynthetically active radiation regime. We repeated this experiment twice using three replicates per treatment.

2.2 Characterisation of plant phenotypes

2.2.1 Quantification of leaf length

At harvest time, images of the maize shoots were manually acquired using an EOS 4000D DSLR camera (Canon) with an 18-55 mm lens in the automatic mode. The flash function was kept off, and images were stored in JPEG format. Individual leaf length was determined from images using the Measure function in ImageJ software (Schneider *et al.*, 2012). After choosing 'Set the scale' in cm in each image, leaf length was determined by drawing a line from the leaf base to the leaf tip using the "freehand" function in ImageJ and measuring it using the Measure function.

2.2.2 Determination of root biomass

To determine the root dry weight, the total root system was collected and rinsed with water to remove the soil particles attached. Cleaned roots were lyophilized in an Alpha 2-4 LD freeze-dry system and then weighted on a Mettler four-decimal analytical scale.

2.2.3 Analysis of leaf mineral elemental profile

The elemental profiles of individual leaves were measured using Inductively Coupled Plasma Mass Spectrometry (ICP-MS). Individual leaves were collected from the plant at the V4-developmental stage (leaf number 4 fully expanded). The leaf material was washed three times with 18.2 MΩcm Milli-Q water (Merck Millipore) and cut into small pieces using a clean ceramic scalpel. Representative samples containing fragments from the different regions of the leaf were collected in a sterile 1.5-mL Eppendorf tube and stored at 4 ºC until further use. The leaf samples were placed in weighted Pyrex digestion tubes and dried at 88 ºC for 20-h. After cooling, 5-10 mg of leaf samples were weighted on a Mettler five-decimal analytical scale, and 1-3 mL (depending on the sample dry weight) of the concentrated trace metal grade nitric acid Primar Plus (Fisher Chemicals) was added to each tube. Prior to the digestion, 20 µg/L of Indium (In) was added to the nitric acid as an internal standard to assess putative errors in dilution or variations in sample introduction

and plasma stability in the ICP-MS instrument. The samples were then digested in DigiPREP MS dry block heaters (SCP Science; QMX Laboratories) for 4-h at 115 ˚C. After cooling down, the digests were diluted to 10-30 mL (depending on the volume of the nitric acid added) with 18.2 MΩcm Milli-Q Direct water. The elemental analysis was performed using an ICP-MS, PerkinElmer NexION 2000 equipped with Elemental Scientific Inc 4DXX FAST Dual Rinse autosampler, FAST valve and peristaltic pump. The instrument was fitted with a PFA-ST3 MicroFlow nebulizer, baffled cyclonic C3 high sensitivity glass spray chamber cooled to 2 ºC with PC3X Peltier heated/cooled inlet system, 2.0 mm i.d. quartz injector torch and a set of nickel cones. Twenty-four elements were monitored including the following stable isotopes: ^{7}Li, ^{11}B, ^{23}Na, ^{24}Mg, ^{31}P, ^{34}S, ^{39}K, ^{43}Ca, ^{48}Ti, ^{52}Cr, ^{55}Mn, ^{56}Fe, ^{59}Co, ^{60}Ni, ^{63}Cu, ^{66}Zn, ^{75}As, ^{82}Se, ^{85}Rb, ^{88}Sr, ^{98}Mo, ^{111}Cd, ^{208}Pb and ^{115}In. Helium was used as a collision gas in Kinetic Energy Discrimination mode (KED) at a flow rate of 4.5 mL/min while measuring Na, Mg, P, S, K, Ca, Ti, Cr, Mn, Fe, Ni, Cu, Zn, As, Se and Pb to exclude possible polyatomic interferences.

The remaining elements were measured in the standard mode. The instrument Syngistix™ software for ICP-MS v.2.3 (Perkin Elmer) automatically corrected any isobaric interferences. The ICP-MS measurements were performed in peak hopping scan mode with dwell times ranging from 25 to 50 ms depending on the element, 20 sweeps per reading and three replicates. The ICP-MS conditions were adjusted to an RF power of1600 Watts and an auxiliary gas flow rate of 1.20 L/min. Torch alignment, nebuliser gas flow and quadrupole ion deflector (QID) voltages (in standard and KED mode) were optimized before analysis for highest intensities and lowest interferences (oxides and doubly charged ions levels lower than 2.5 %) with NexION Setup Solution containing 1 µg/L of Be, Ce, Fe, ln, Li, Mg, Pb and U in 1 % nitric acid using a standard built-in software procedure. To correct for variation between and within ICP-MS analysis runs, liquid reference material was prepared using pooled digested samples and run after the instrument calibration and then after every nine samples in all ICP-MS sample sets. Equipment calibration was performed at the beginning of each analytical run using seven multi-element calibration standards (containing 2 µg/L In internal standard) prepared by diluting 1000

mg/L single element standards solutions (Inorganic Ventures; Essex Scientific Laboratory Supplies Ltd) with 10 % nitric acid. As a calibration blank, 10 % nitric acid containing 2 µg/L In internal standard was used, and it was run throughout the analysis. Sample concentrations were calculated using the external calibration method within the instrument software. Further data processing, including calculation of final elements concentrations, was performed in Microsoft Excel.

2.2.4 Analysis of soil mineral elemental profile

For the soil mineral elemental analysis, soil samples were harvested from the individual pots and placed in 50 mL Falcon tubes. The soil samples were first dried using a plastic weighing boat in the fume hood for approximately 72 h at room temperature. Five grams of soil was then weighted in 50 mL Falcon tubes with a four-decimal scale, and digested with 20 mL of 1 M NH_4HCO_3, 5 mM diamine-triamine-penta-acetic acid (DTPA), and 5 mL 18.2 MΩcm Milli-Q Direct water (Merck Millipore). The soil mixture were shaken for 1 h at 150 rpm in a rotary shaker (adapted from Soltanpour and Schwab, 1977). Each sample was gravity filtered through a quantitative filter paper (Whatman 42- WHA1442070) until obtaining approximately 5 mL of filtrate. From this filtrate, 0.5 mL were then open-air digested in Pyrex tubes using 1 mL of concentrated trace metal grade nitric acid Primar Plus (Fisher Chemicals) spiked with 20 µg/L indium internal standard for 4-h at 115 ºC in dry block heater (DigiPREP MS, SCP Science; QMX Laboratories, Essex, UK). Then, each sample was diluted up to 10 mL using Milli-Q water as before. The elemental analysis was performed using a PerkinElmer NexION 2000 ICP-MS equipped with an autosampler Elemental Scientific Inc., in the collision mode (He), as described for leaf samples. To correct for variation between and within ICP-MS analysis runs, liquid reference material was prepared using pooled digested samples and run in the same manner as leaf samples. Sample concentrations were calculated by the instrument software using the external calibration method. Further data processing to obtain the final concentration of the elements was performed in Microsoft Excel and included correction of the drift,

subtraction of the blank concentration, multiplication by the dilution factor and normalization to the soil dry weight.

2.2.5 Plant phenotypes and mineral nutrient profile analyses

The total leaf length and single-leaf length were compared across soils using the Kruskal-Wallis model. We inspected the parametric statistical assumptions (here and elsewhere) using Q-Q plots and Bartlett's tests (Kozak and Piepho, 2018). Differences between soils and leaves were indicated using compact letter display derived from Dunn's test implemented in the package FSA (Ogle *et al.*, 2020).

For the mineral nutrient analysis of leaf and soil elemental profiles, we created a matrix (samples x ion) in which each cell was filled with the calculated element concentration of each given sample. Afterwards, we applied a z-score transformation of each ion across the samples in the matrix. Next, we applied a principal component analysis (PCA) using the Euclidean distance between samples and the z-score matrix as input to compare the elemental profiles of leaves and soil. Additionally, we estimated the variance explained by soil origin, fractions and leaf number by performing PERMANOVA via the function adonis from the vegan v2-5.5 R package (Oksanen, 2007).

To visualize the mineral content in the individual leaves, the z-score matrix created above was hierarchically clustered (method ward.D2, function hclust), and we visualized the results using a heatmap. The rows in the heatmap were ordered according to the dendrogram order obtained from the clustered ions. The heat map was coloured based on the z-score.

2.3 Analysis of microbiome composition

2.3.1 DNA extraction

Bacterial composition was determined in individual leaves, roots, and soil samples from plants grown using the three soils described in material and methods 2.1. At the harvesting time, individual leaves were cut into small pieces using a sterile scalpel. A representative sample containing leaf fragments from the different regions of each leaf was collected in a sterile 2-mL Eppendorf tube. Leaf samples were rinsed three times with sterile distilled water to remove

weakly associated microbes and soil particles. The whole root system was also collected and rinsed three times with sterile distilled water in 500 mL glass bottles. Soil samples were collected in 50 mL falcon tubes containing sterile distilled water and then filtered using sterile 100 µm nylon mesh cell strainers (Fisherbrand™, Thermo Fisher Scientific Inc.) to remove plant debris and big soil particles. Then, soil samples were centrifuged at room temperature for 20 min at max speed in an Eppendorf 5810R centrifuge, and the supernatant was discarded. The pellet was dissolved in 1 mL of sterile water, and the resulting suspensions were transferred to sterile 1.5-mL Eppendorf tubes. Samples were centrifuged again at high speed in a benchtop centrifuge, and the supernatants were discarded. All samples were stored at -80 ºC until use.

Before DNA extraction, leaf and root samples were lyophilized using an Alpha 2-4 LD freeze-dry system. Individual root systems were cut into small pieces using flame-sterilized scissors, and a representative root sample was transferred to 2-mL Eppendorf tubes. Leaf and root samples were individually pulverised using a Tissue Lyser II (Qiagen) using 4 cycles of 60 seconds at the frequency of 30 s^{-1}. The DNA extraction was carried out using 96-well-format MoBio PowerSoil Kit (MOBIO Laboratories; Qiagen) following the manufacturer's instructions. Before starting the extraction, leaf, root and soil samples were manually randomised individually by thoroughly mixing in a plastic bag. Samples were taken individually from the bag and loaded into the DNA extraction plates. This random distribution was maintained throughout library preparation and sequencing.

2.3.2 16S rRNA amplification and sequencing

For the 16S rRNA sequencing, the V3-V4 region of the bacterial 16S rRNA gene was amplified using the primers 338F (5′-ACTCCTACGGGAGGCAGCA-3′) and 806R (5′-GGACTACHVGGGTWTCTAAT-3′). Each sample was amplified in triplicate, and the PCR reactions were performed using three sets of unique primers combinations per plate. The following conditions were used for PCR amplification: 5 µL Kapa Enhancer, 5 µL Kapa Buffer A, 1.25 µL of 5µM 338F, 1.25 µL of 5µM 806R, 0.375 µL mixed plastid and mitochondrial

rRNA gene-blocking peptide nucleic acids (PNAs; 1:1 mix of 100 µM plastid PNA and 100 µM mitochondrial PNA), 0.5 µL Kapa dNTPs, 0.2 µL Kapa Robust Taq, 8 µL dH_2O and 8 µL DNA; temperature cycling: 95 ºC for 1 min, 24 cycles of 95 ºC for 15 s, 78 ºC (PNA) for 10 s, 50 ºC for 30 s, 72 ºC for 30 s and 4 ºC until use. PCR products of triplicate were combined and purified using AMPure XP magnetic beads (Beckman Coulter) to remove primer dimers. Then, 10 µL of purified DNA underwent a second PCR using the following conditions: 12 µL Kapa HiFi Readymix, 0.375 µL mixed (1:1) of 100 µM pPNA and 100 µM mPNA; 2.5 µL of 5 µM index primers; temperature cycling: 95 ºC for 1 min, 9 cycles of 95 ºC for 15 s, 78 ºC for 10 s, 60 ºC for 30 s, 72 ºC for 35 s and 4 ºC until use. Subsequently, amplicons were pooled together in equal amounts and then diluted to 10 pM for sequencing. Sequencing was performed on an Illumina MiSeq instrument using a 600-cycle V3 chemistry kit in the DeepSeq facility at the University of Nottingham. DNA sequence data, the abundance matrix, metadata and taxonomy will be available upon publication.

2.3.3 16S rRNA amplicon sequence data processing

For the 16S rRNA analysis, sequence data were processed with cutadapt (Martin, 2011), DADA2 and R packages. Raw reads were demultiplexed and trimmed with cutadapt. The resulting sequences were then denoised and collapsed into amplicon sequence variants (ASVs), using DADA2 v.1.10.1 (Callahan *et al.*, 2016). Briefly, the paired reads were filtered by removing sequences containing uncalled bases, truncating reads after bases with a quality score of 2 and reads with no more than 3 expected errors using the FilterAndTrim function. We then learn the error for the forward and reverse reads separately using the learnErrors function. These error rates were then used to infer amplicon sequence variants (ASVs) with the function dada on the reverse and forward reads individually. Thereafter, we merged the forward and reverse sequence with the mergePairs function, and the merged ASVs were used to build an ASV sequencing table using the makeSequenceTable function. Then, we removed chimaeras from the sequence table with the removeBimeraDenova function. The resulting ASVs were mapped to SILVA 132 database (Quast *et al.*, 2013). We filtered ASVs that were assigned to chloroplast, mitochondria, oomycetes, archaea or did not have a known

kingdom assignment. After filtering, the remaining ASVs with more than 1000 reads per sample were used to create the raw count abundance tables. Those tables were then processed and analysed with functions from the ohchibi package (https://github.com/isaisg/ohchibi).

2.3.4 16S rRNA amplicon sequence analysis

To compare alpha diversity across soils, roots, and individual leaves, we calculated the Shannon diversity index using the diversity function from the vegan package v2.5-5 (Oksanen, 2007). We used Kruskal-Wallis to test for alpha diversity differences between the fractions. Beta diversity analyses (Principal coordinate analysis, PCo, and canonical analysis of principal coordinates, CAP) were based on Bray-Curtis dissimilarity matrices calculated from the rarefied abundance tables. Additionally, we estimated the variance explained by soil origin, fractions and leaf number by performing PERMANOVA via the function adonis from the vegan v2-5.5 R package (Oksanen, 2007). We applied PCoA using the function prcomp in R.

The relative abundance of bacterial phyla was depicted using a stacked bar representation encoded in the function chibi.phylogram from the ohchibi package (https://github.com/isaisg/ohchibi).

We used the R package DESeq2 v.1.24.0 (Love et al., 2014) to compute the fractions-specific enrichment profiles for the enrichment profile in the natural soil characterisation experiment. For each ASV in the abundance tables, we estimated its difference in abundance with the different fractions (leaf, root, and soil) by fitting a Generalized Linear Model (GLM) with the following design:

$$\text{Abundance} \sim \text{Fractions}$$

We extracted the following comparisons from the fitted model: leaf vs. root, leaf vs. soil, and root vs. soil. An ASV was considered statistically significant if it had a p-value < 0.05. We visualized the results for the GLM using a heat map. The rows in the heat map were ordered according to the dendrogram obtained from the ASV analysis. The heat map was coloured based on the $\log_2$-

transformed fold change output by the GLM. The significant comparisons (*p*-value < 0.05) were highlighted with black framing.

For the enrichment at the single-leaves level, we used the same approach described above with a different linear model design:

Abundance ~ Leaf number (NumLeaf) + Soil

3 Results

3.1 Plant shoot growth integrates soil (a)biotic cues at the level of individual leaves

To identify the main drivers of maize leaf growth under controlled conditions, we grew four *Zea mays* (common name maize) in two arable soils

similar in mineral nutrient composition from Santiago island, Cabo Verde (Tarrafal and São Domingos ~ 33.49 Km apart) and in a natural soil from the UK (Sutton-Bonington campus ~ 4.635 Km away) (Figure 1).

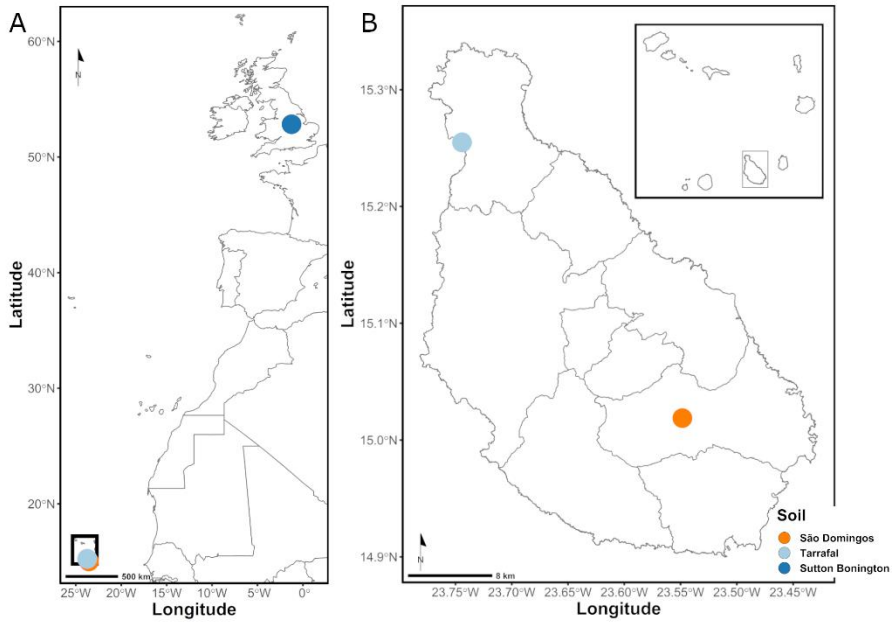


Figure 1: Geographic location of the soils collected for this study. A) Map of the regions in two countries selected for this study. The small black square showed Cabo Verde country with the two field locations. Different colours represent the different locations of the three soils. **B)** Map of the two regions selected in Santiago island. Different colours represent the different locations.

The soils of Santiago Island in Cabo Verde were selected because they have low mineral nutrient content and they are exposed to very short rainy seasons limiting the growth and production of maize on the island. On the contrary, the UK soil is nutrient-rich and receives frequent precipitations. This selection of diverse soils was intended to facilitate the study of parameters affecting plant growth under controlled conditions. Furthermore, the selected genotypes include local varieties well-adapted to the poor soil conditions found in some African areas.

Before starting the experiment, six replicates of each soil were analysed in two independent experiments to determine the nutrient content. We examined differences in soil mineral content in the different soils using Inductively Coupled Plasma Mass Spectrometry (ICP-MS). We found that soil mineral content differed among sites. Principal component analysis (PCA) demonstrated that ~84 % of variability found in soil mineral content was explained by the soil origin (adonis: degrees of freedom (df) = 2; coefficient of determination (R^2) = 0.84; p-value < 0.001) (Figure 2). The primary axis of variation distinguished the Cabo Verde soils (Tarrafal and São Domingos) from the UK soil (Sutton Bonington), while the second axis of variation separated the soils from Tarrafal and São Domingos (Figure 2).

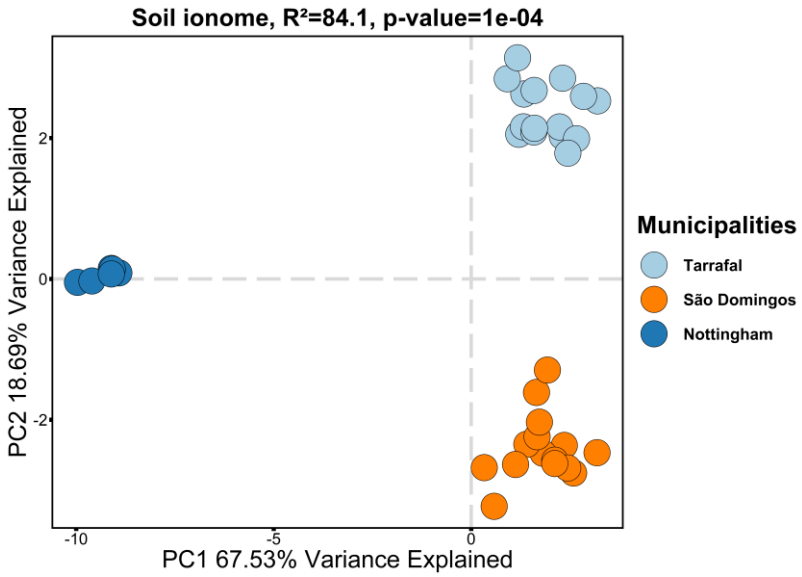


Figure 2: The selected soils differ strongly in mineral content. Principal component analysis (PCA) based on Euclidean distance between mineral content across the three soils tested. The values on the top of the plot represent the percentage of variance explained by statistically significant (p<0.05) variable in a PERMANOVA model.

This pattern suggested that Tarrafal and São Domingos soils have a similar nutritional composition that is entirely different from Sutton Bonington soil.

We next assessed the effect of the different soils on plant growth by growing the four maize genotypes in the different soils in glasshouse conditions. For each plant genotype, we measured total shoot length as a proxy for plant

productivity. To minimize changes in plant growth due to water availability and plant development, a water regime of 50 % of the soil's water content capacity was applied in each case, and all plants were harvested at the V4 developmental stage (stage in which a plant has four fully expanded leaves).

We observed significant growth differences in plants cultivated in São Domingos soil compared to plants grown in a similar nutritional soil from Tarrafal (Figure 3A). We also found soil-specific differences in leaf length but could not identify genotype-specific differences (Kruskal-Wallis, p-value = 0.09) (Figure 3B). Maize plants grown in São Domingos soil showed no significant differences in shoot length compared to those grown in soil from Sutton Bonington, despite the evident contrasting amount of mineral nutrients (Figure 2 and Figure 3A). In contrast, plants grown in Tarrafal poor soil were significantly smaller than those growing on the other two soils, and their development was delayed, consistent with plants grown in poor soil (Figure 3C). We noticed that these differences in total shoot length did not fully translate to the level of individual leaves (Figure 3D). In all cases, leaves 1 and 2 showed no differences in length across the soils used. On the contrary, leaves 3, 4, and 5 showed significant differences that recapitulated those observed in the total length of the plant shoot (Figure 3A and Figure 3D).

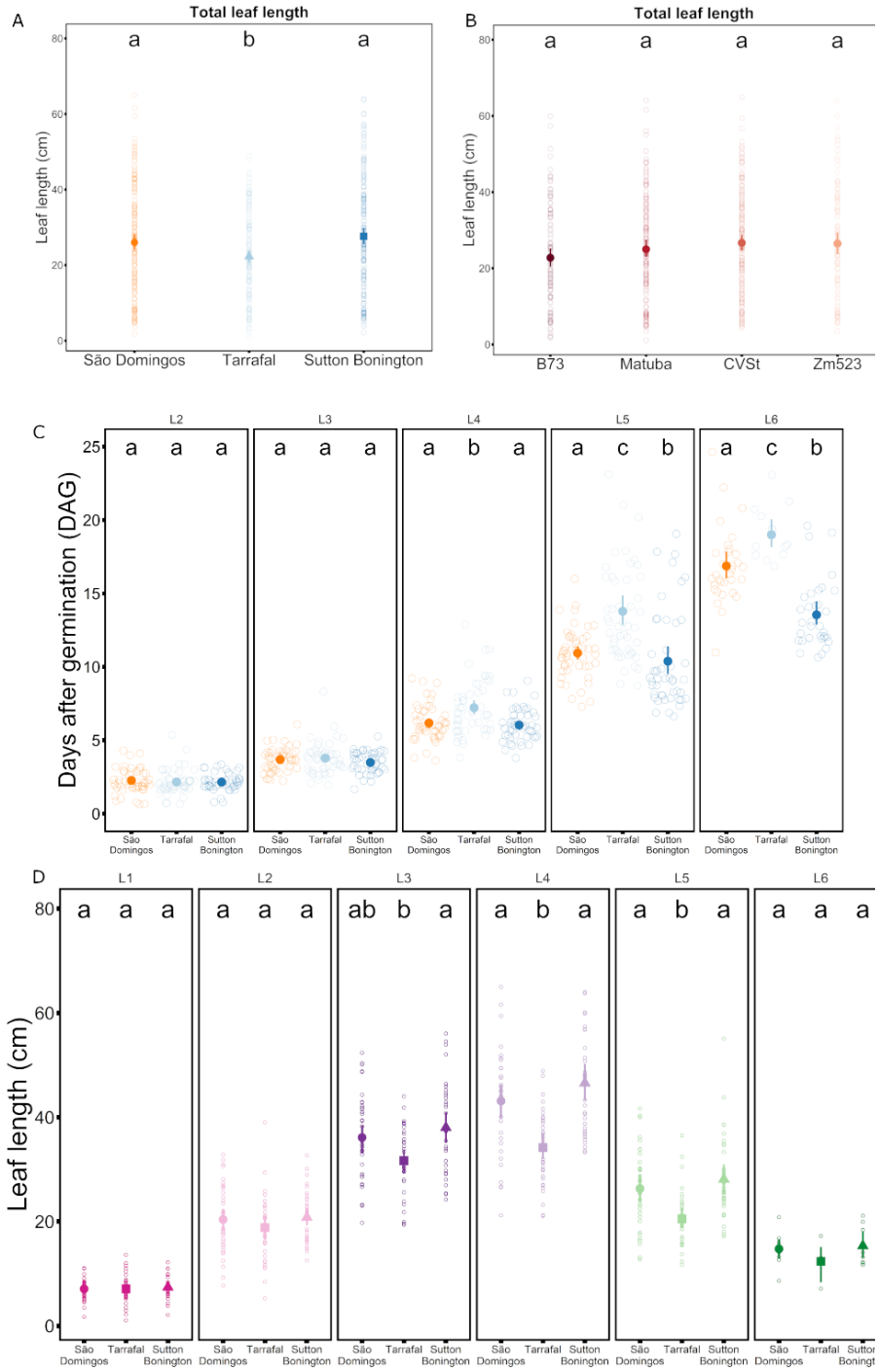

A
Total leaf length
Leaf length (cm)
a
b
a
São Domingos
Tarrafal
Sutton Bonington
B
Total leaf length
Leaf length (cm)
a
a
a
a
B73
Matuba
CVSt
Zm523
C
L2
L3
L4
L5
L6
Days after germination (DAG)
a a a
a a a
a b a
a c b
a c b
São Domingos
Tarrafal
Sutton Bonington
D
L1
L2
L3
L4
L5
L6
Leaf length (cm)
a a a
a a a
ab b a
a b a
a b a
a a a
São Domingos
Tarrafal
Sutton Bonington

Figure 3: Maize leaves differentially integrate soil effect as assessed from growth at the V4 developmental stage in the three different soils. A) Total leaf length reached by the four genotypes. **B)** Total leaf length per maize genotype in the three different soils. **C)** Soil effect on maize development, measured as number of days needed to reach full development of the different leaves. **D)** Individual leaf length (in cm) of four-week-old maize plants (pooled data from the four genotypes) grown in the three different soils. In all plots, letters indicate statistical significance corresponding to the Kruskal-Wallis test with Dunn's *post hoc* test (α=0.05).

In fact, the differences in total shoot growth found across soils were justified by the differences found in leaves 3 and 4. When removing these leaves from the analysis, all plants show similar leaf growth despite the soil used (Figure 4A). In general, no significant differences could be found in root dry weight across the three soils (Figure 4B). Therefore, we ruled out root-associated functions as causal factors for the observed differences in plant shoot growth. All these results indicate that maize shoot growth plasticity responds differentially to soil cues at the level of individual leaves.

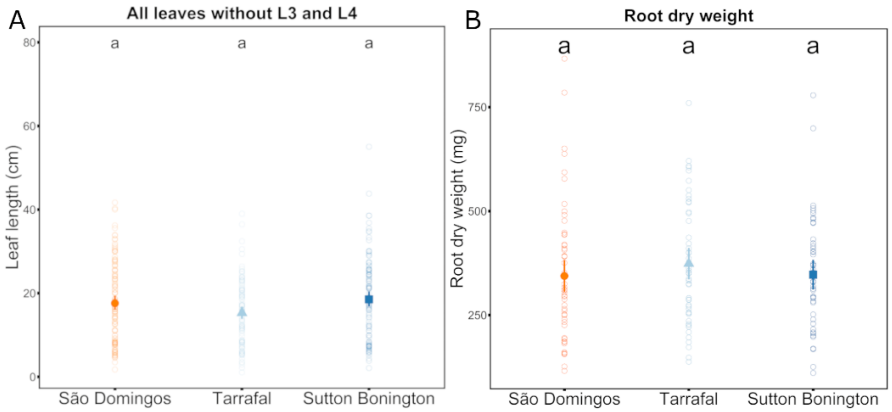


Figure 4: . Leaves 3 and 4 are responsible for the differential integration of soil effect. Leaf length **(A)** and root dry weight **(B)** of four maize genotypes (B73, Matuba, CVSt and Zm523 pooled data) grown for 4 weeks in three different soils. In A, the leaf length is only for leaves 1, 2, 5 and 6, leaving out leaves 3 and 4. Letters indicate statistical significance corresponding to Kruskal-Wallis with Dunn's *post hoc* test (α=0.05). Three independent biological replicates (n=150 plants) in two independent experiments.

3.2 Leaf ionome alone is not sufficient to explain leaf growth differences across soils

To investigate whether the accumulation of mineral nutrients and trace elements (ionome) in the individual leaves is relevant for leaf growth plasticity, we determined the leaf ionome in maize plants grown in different soils (Figure 5). We found that individual leaves showed distinct ionomic profiles, and this pattern was conserved across all soils analysed (Figure 5A and Figure 5B). We observed that independently of the soil used to grow the plants, nutrients such as P, Ni, and K have positive correlations with the leaves age while Ca, Mg, and Mn have negative correlations (Figure 5C - Figure 5E). This suggests that the plant's nutrient allocation and recycling strategies are robust to soil composition. Nevertheless, the soil's nutrient availability influenced the total concentration of the different elements per leaf. Indeed, plants grown in the two soils from Cabo Verde (Tarrafal and São Domingos), with similar mineral nutrient content, exhibited similar single-leaf ionomes compared with plants grown in Sutton-Bonington (Figure 6). These results suggest that the ionomic profiles of individual leaves alone are not sufficient to explain the differences in the growth of maize leaves in plants grown under realistic conditions.

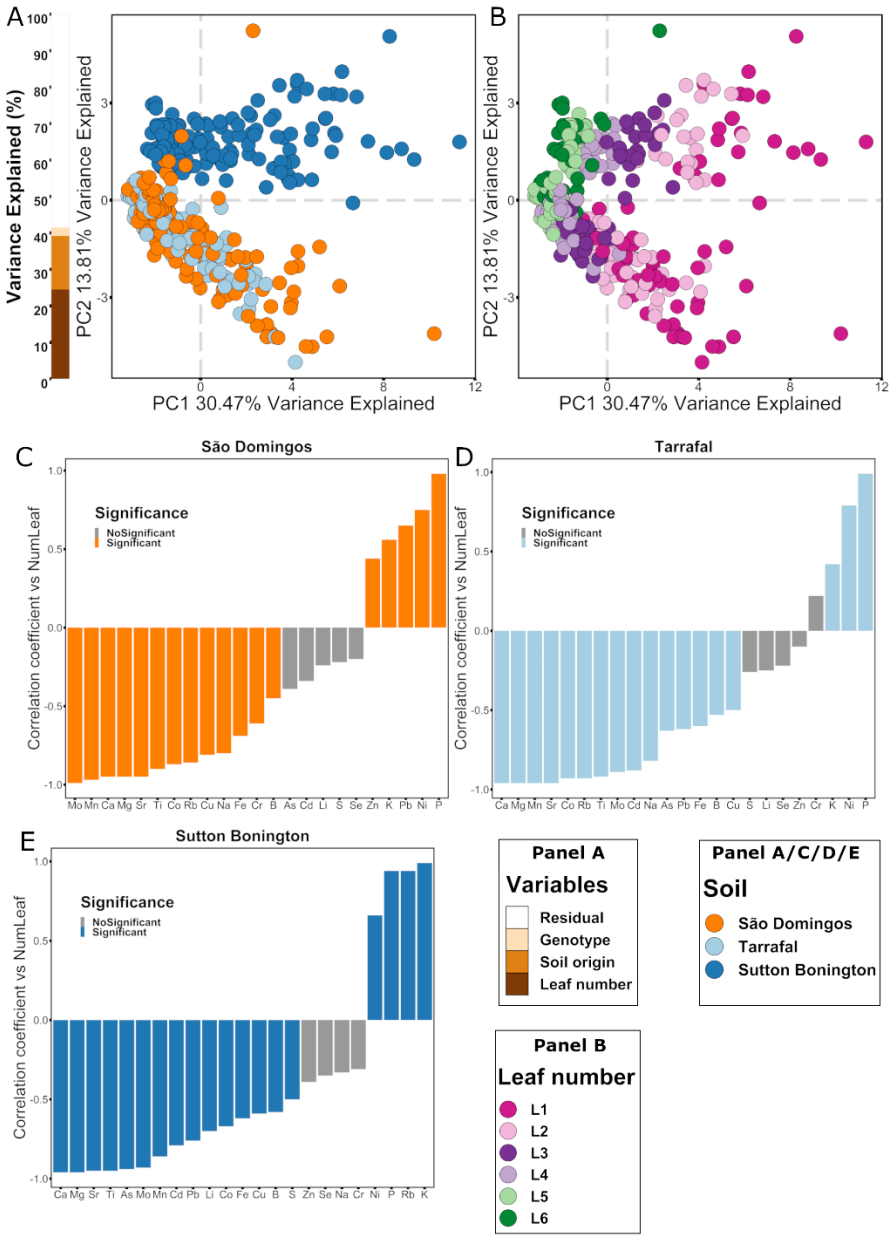


Figure 5: Individual leaves show distinct ionomic profiles. A) PCA analysis based on Euclidean distance of the leaf mineral composition (ionome) showing the ionomic profile of the pooled maize leaves (four genotypes) in the three soils. The bar plot to the left of the PCA represents the percentage of variance explained by statistically significant ($p < 0.05$) variables in a PERMANOVA model. Different colours show the leaf mineral profile in each soil. B) PCA

analysis based on Euclidean distance of leaf ionome showing the mineral profiles of maize single-leaf ionome grown in the three soils. Different colours identify different leaves (L1-L6). C)-E) Correlation coefficient between the concentration in leaf of each mineral nutrient and the leaf age from plants grown in São Domingos soil (C), Tarrafal soil (D) or Sutton Bonington soil (E). Grey bars are not significant ($q > 0.05$).

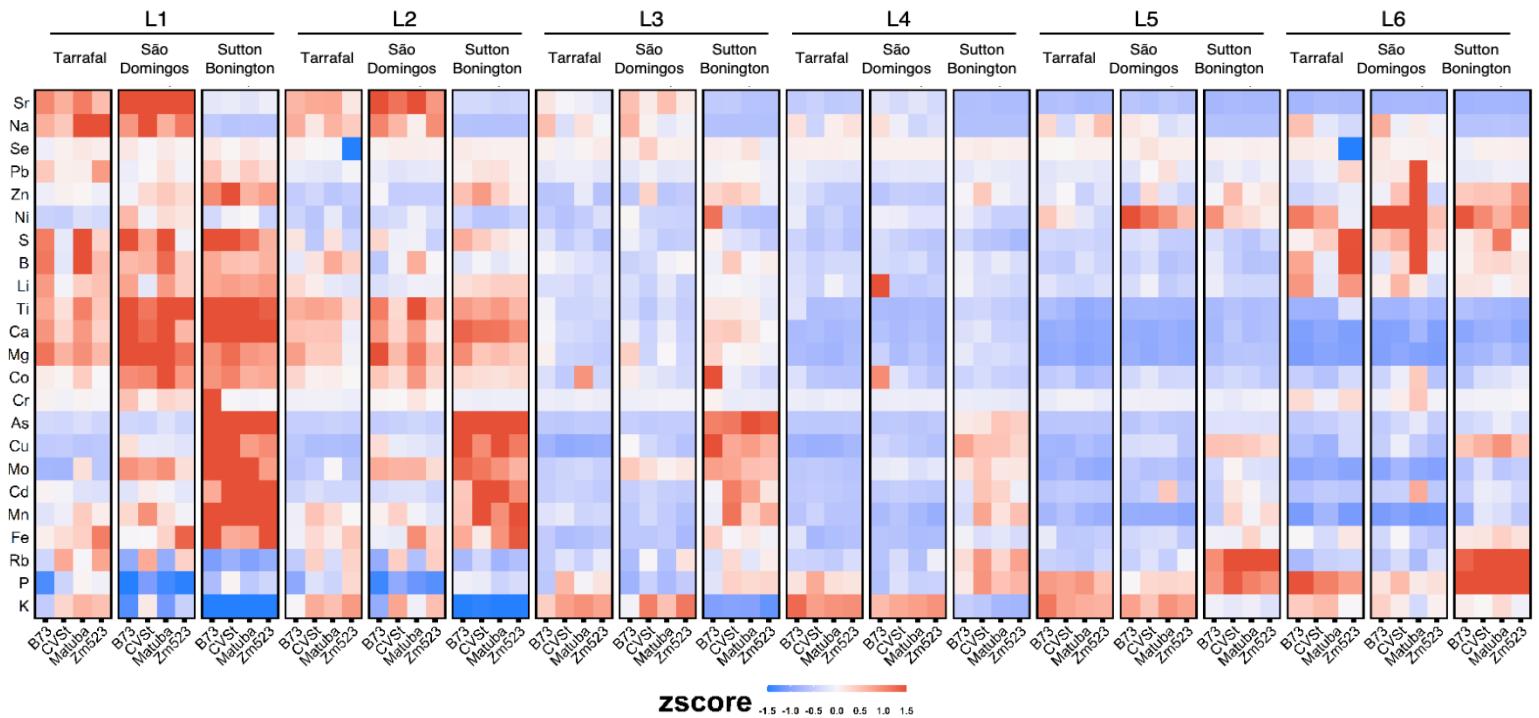


Figure 6: Single-leaf ionome alone is not sufficient to explain plant growth differences. Heatmap showing standardized mineral nutrient concentrations in individual leaves. Each row represents the average zscore for each ion (degree of variation above or below the mean variability for that ion) analysed in this study. Each column in the heatmaps represents the leaf number of the plants grown in each soil. The heat maps are coloured based on z-score.

3.3 Leaves host a distinct microbiota

To examine whether microbial composition contributes to the observed plant leaf phenotypes, we first profiled bacterial communities of maize plants grown in the three soils selected (see materials and methods). We harvested soil, roots, and leaves from plants at V4 developmental stage (with four fully expanded leaves). We also harvested soil from unplanted pots that we used as controls. Next, we identified bacterial community profiles for each sample via PCR amplification of the 16S ribosomal RNA (rRNA) gene followed by sequencing. We obtained 10,357,079 sequences from 514 samples. After removing chimeric and organelle sequences, we obtained 40,761 unique amplicon sequence variants (ASVs).

As expected, we observed that the soil bacterial community was significantly more diverse than the root and leaf bacterial communities across all soils used. The diversity decreased from root to leaf in all soils, as judged by α-diversity calculated using the Shannon diversity index (Figure 7).

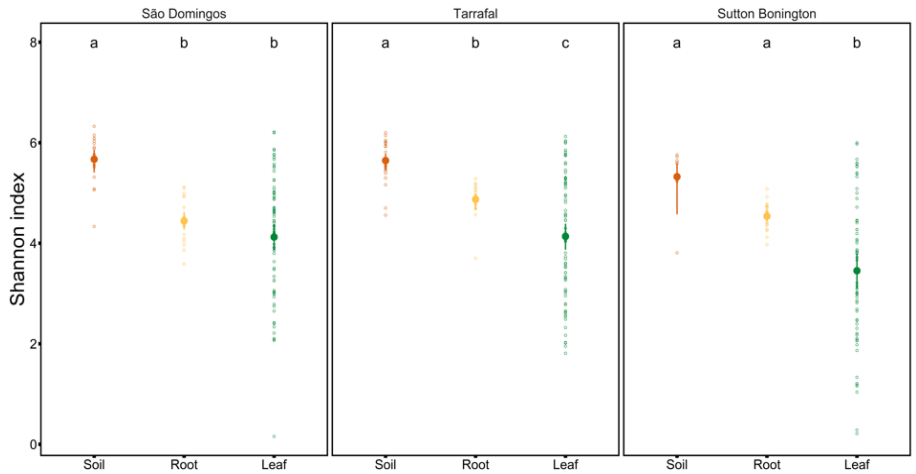


Figure 7: Alpha diversity substantially decreases from soil to leaves in all soils tested. Plots showing the alpha diversity across sample compartments (soil, root, and leaf) in the different soils were estimated using the Shannon diversity index. Letters indicate statistical significance corresponding to Kruskal-Wallis with Dunn's *post hoc* test (α = 0.05).

Next, we explored the community composition for each sample fractions (soil, root, and leaf). Canonical analysis of principal coordinates (CAP) showed

that roots and leaves harbour bacterial communities distinct from the surrounding soil community and from each other (Figure 8A).

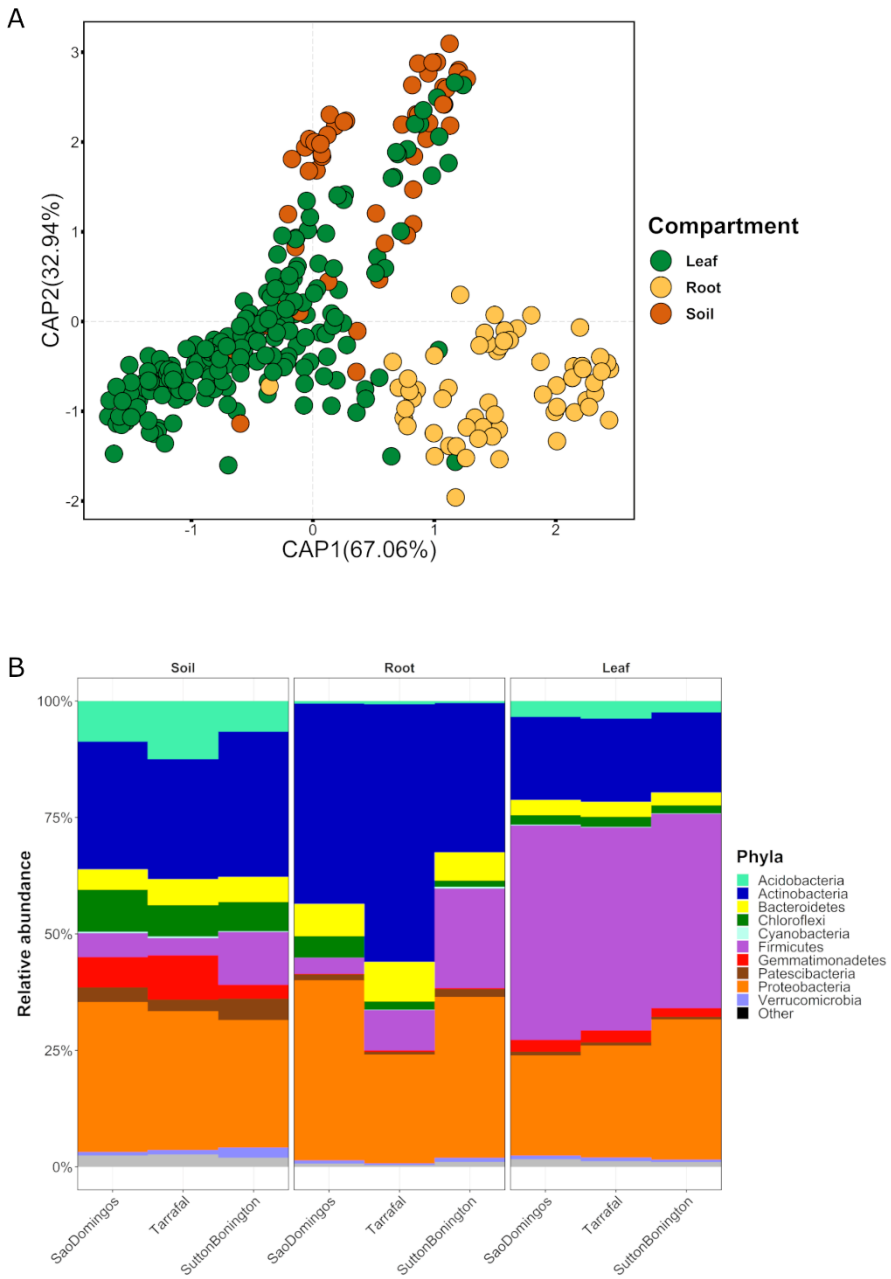


Figure 8: Sample fractions harbour distinct bacterial communities. A) Canonical analysis of principal coordinates (CAP) showing the influence of fraction (soil, root and leaf) on the assembly of the bacterial community. Different colours differentiate the compartments. **B)** Relative abundance profiles of the main bacterial phyla across compartments (soil, root and leaf).

The primary axis (CAP1) distinguished the soil and root compartment from the leaf compartment. In contrast, the secondary axis of variation (CAP2) separated the plant compartment from the soil. Then, we investigated the relative abundance patterns of bacterial phyla in each compartment. The general taxonomic patterns in each compartment were driven mainly by differences in the relative abundances of taxonomic groups. In comparison to soil, root-derived samples were primarily enriched in members of the phyla Actinobacteria, Firmicutes and Proteobacteria, and depleted in Gemmatimonadetes, Patescibacteria, and Chloroflexi members (Figure 8B). As compared to soil, leaf-enriched ASVs belonged mainly to the Firmicutes phylum, while leaf-depleted belonged to Actinobacteria (Figure 8B). This pattern was, in general robust across all soils.

To further explore microbiota composition in the different fractions and study the enrichment profile at ASV, Family, and Class levels, we used a generalized linear model (GLM) (Figure 9 - Figure 11) in the different fractions. We detected a notable significant enrichment pattern at the ASVs level at soil (1353 ASVs) and root compartment (2302 ASVs) levels across the three different soils (Figure 9). In contrast, the enrichment pattern significant in the leaf (16 ASVs) was very low, suggesting that the bacterial communities in the different soils do not differ much. Interestingly, this pattern was maintained also at the Family and Class level (Figure 10-Figure 11).

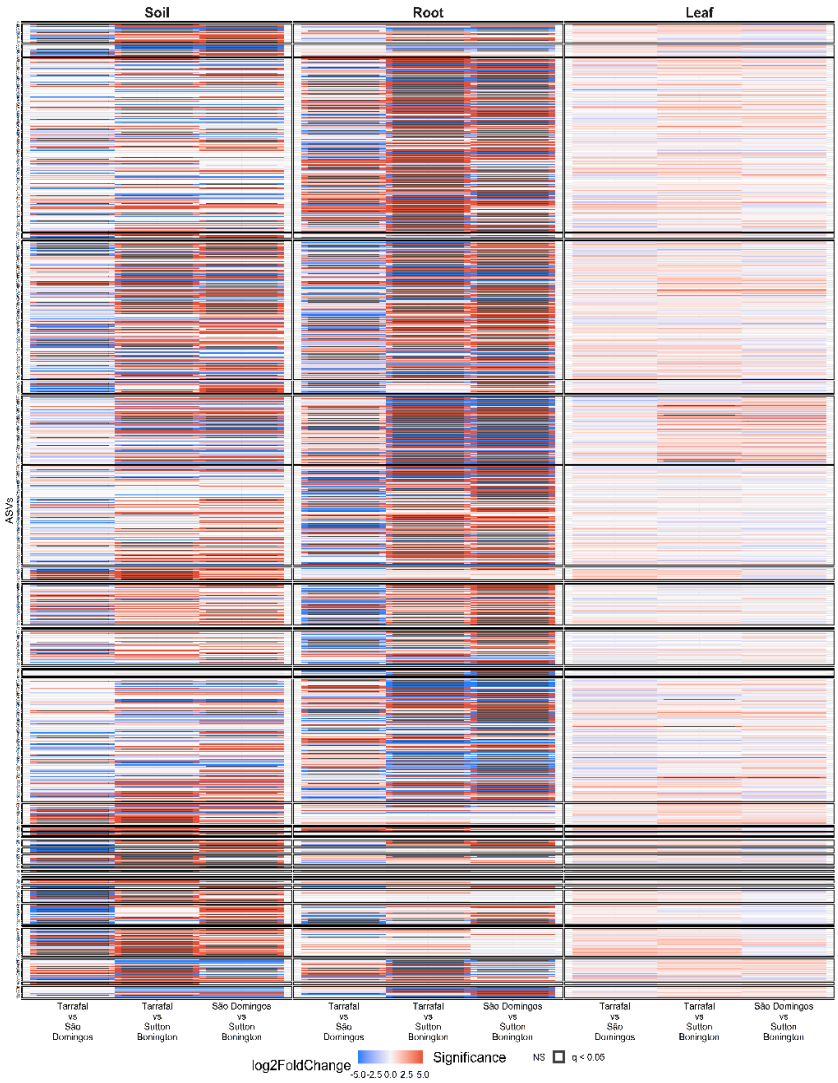


Figure 9: Enrichment pattern of bacterial communities at ASV level in the different sample fractions across the three soils used. Each row in the different panels represents a unique ASV identified in the different fractions (soil, root, and leaf). Each column in the heatmaps represents a specific contrast in the enrichment model. The heat maps are coloured based on $\log_2$ fold changes derived from fitted GLM. Positive fold changes (coloured in red gradient) represent enrichments on the upper side of the name of the contrast (e.g. Tarrafal vs. São Domingos, enriched in Tarrafal as compared to São Domingos), whereas negative fold changes (coloured in blue) represent enrichments on the bottom side of the name of the contrast (e.g. Tarrafal vs. São Domingos, enrichment in São Domingos in comparison to Tarrafal). Significant differences are framed in black ($q < 0.05$)

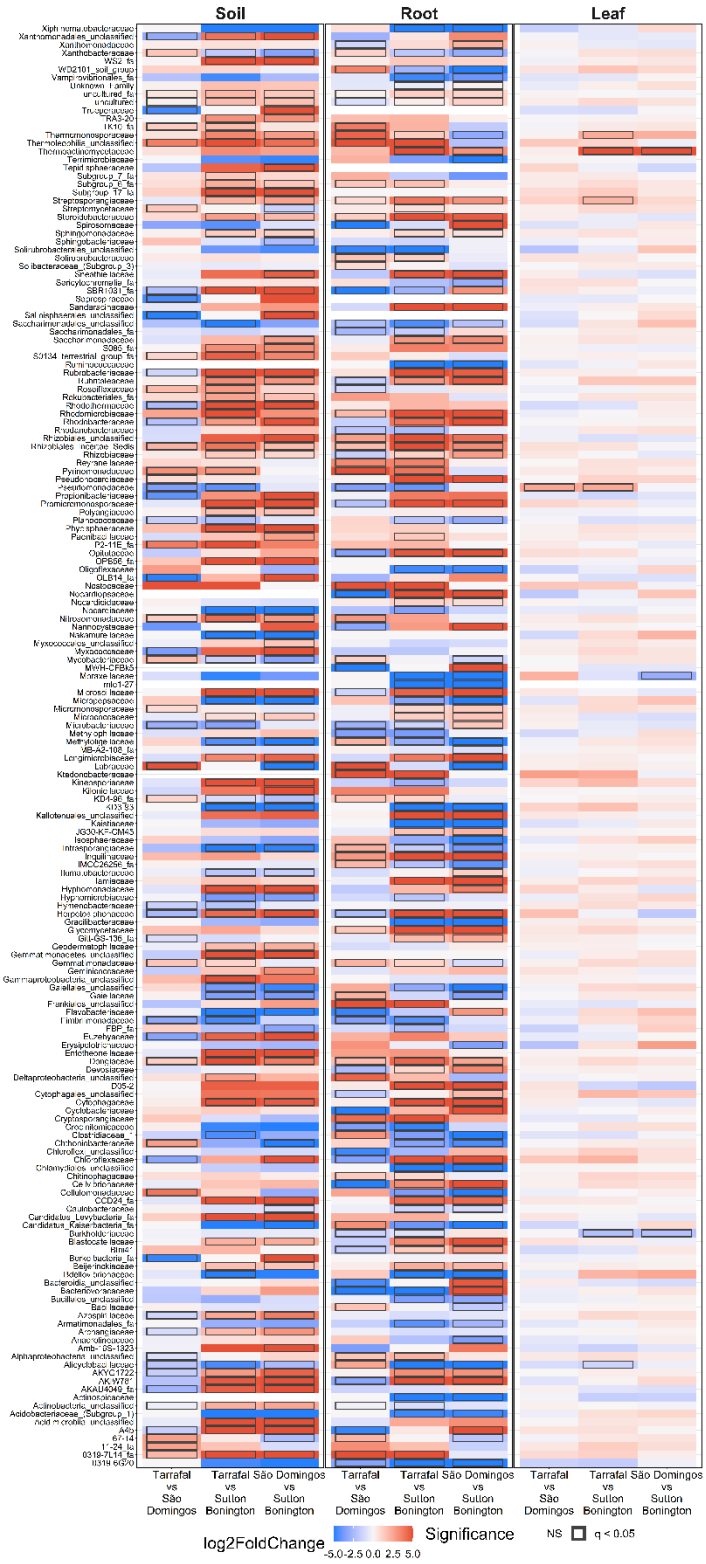


Figure 10: Enrichment pattern of bacterial communities at Family level in the different fractions identified using the three soils. Each row along the different panels represents a unique Family identified in the different compartment (soil, root, and leaf). Each column in the heatmaps represents a specific contrast in the enrichment model. The heat maps are coloured based on $\log_2$ fold changes derived from fitted GLM. Positive fold changes (coloured in red gradient) represent enrichments on the left side of the name of the contrast (e.g. Tarrafal vs. São Domingos, enriched in Tarrafal as compared to São Domingos), whereas negative fold changes (coloured in blue) represent enrichments on the right side of the name of the contrast (e.g. Tarrafal vs. São Domingos, enrichment in São Domingos in comparison to Tarrafal). Significant differences are framed in black ($q < 0.05$).

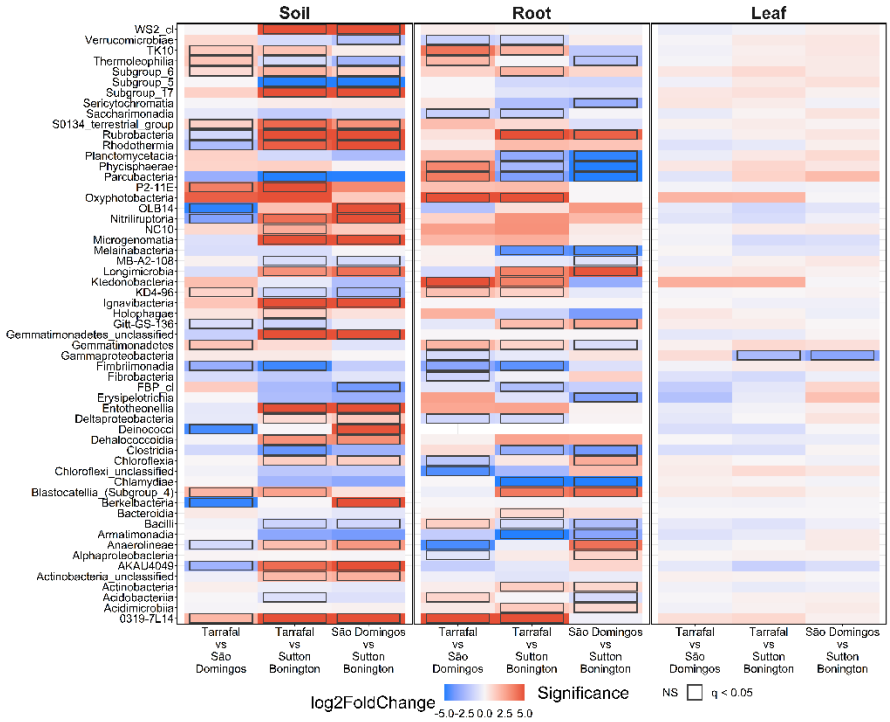


Figure 11: Enrichment pattern of bacterial communities at Class level in the different fractions identified in the three soils. Each row along the different panels represents a unique Class identified in the different compartment (soil, root, and leaf). Each column in the heatmaps represents a specific contrast in the enrichment model. The heat maps are coloured based on $\log_2$ fold changes derived from fitted GLM. Positive fold changes (coloured in red gradient) represent enrichments on the upper side of the name of the contrast (e.g. Tarrafal vs. São Domingos, enriched in Tarrafal as compared to São Domingos), whereas negative fold changes (coloured in blue) represent enrichments on the bottom side of the name of the contrast (e.g. Tarrafal vs. São Domingos, enrichment in São Domingos in comparison to Tarrafal). Significant are framed in black ($q < 0.05$)

To analyse whether the microbial community on the different fractions (soil, root, and leaf) converged according to soil type, we independently calculated the compartment-specific enrichment profile across the three soils. We estimated, at the ASV level, their relative abundance differences in each plant compartment (leaf vs. root) and against soil (leaf vs. soil and root vs. soil) for each soil. In this way, we could identify five clusters (Figure 12). Cluster 1 (Cl1) contained mainly ASVs belonging to Actinobacteria and Proteobacteria, which

were phyla predominantly depleted in leaves of plants grown on Santiago island soils (Figure 12). We also observed that cluster Cl2 was composed mainly of ASVs belonging to the Firmicutes phylum (Figure 12), which was consistently enriched in leaves from plants grown in Santiago soils and to a lesser extent in Sutton Bonington soil. Cl3 and Cl4 contained a high abundance of Actinobacteria and Firmicutes with an enrichment pattern again more similar between the two Santiago soils and different between these two and Sutton-Bonington soil (Figure 12). Finally, Cl5 was composed of Actinobacteria, Proteobacteria and Firmicutes, mainly found enriched in the roots of plants grown in the Santiago soils and depleted in leaves (Figure 12). Altogether, these data indicate that the leaves host more specific bacterial populations than roots and these populations are, in general, more similar between the two African soils studied. This suggests that the differences found in leaf growth across soils (Figure 12) cannot be explained by the differences observed in the total bacteria populations inhabiting plant leaves. Thus, we speculated that a reduced population of highly abundant bacteria, highly adapted to specifically colonise the unique environment of single leaves, may explain differences observed in leaf growth across soils. To test this hypothesis, we explored bacterial populations at a single leaf resolution across the different soils.

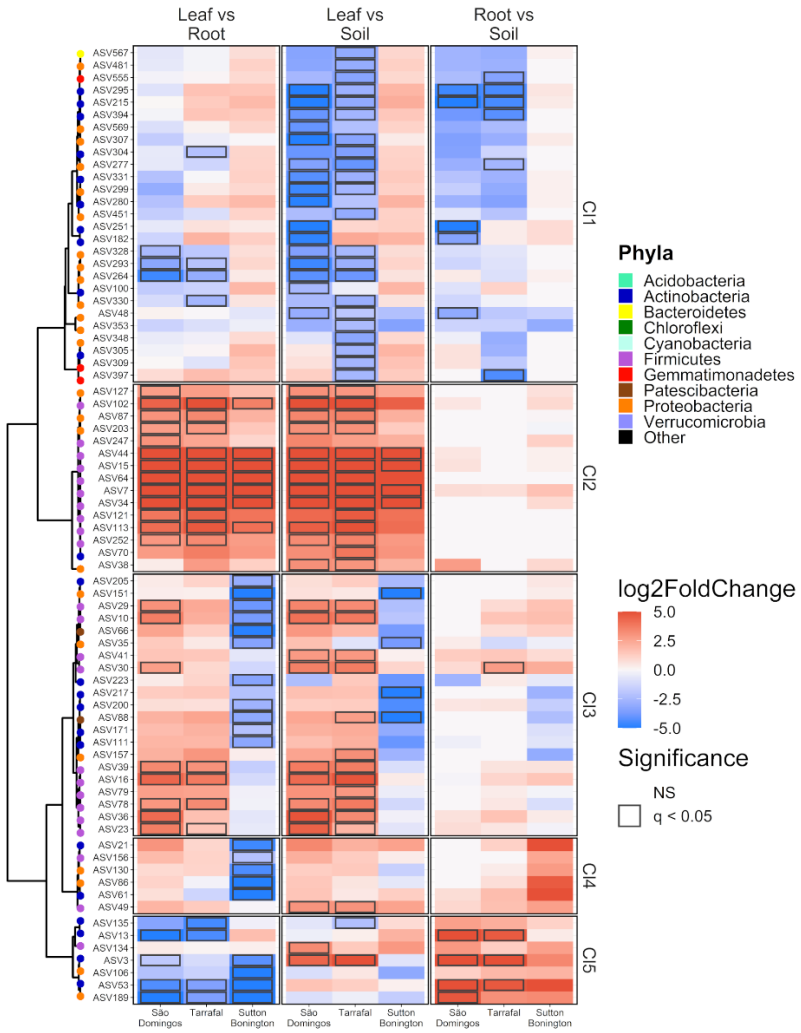


Figure 12: Maize leaves of plants grown on three different soils show a high convergence of bacterial communities across those soils. Heatmaps depicting the enrichment pattern of bacterial communities in the three different soils tested. Dendrogram rounded ends are coloured according to phylum taxonomy. Each row represents a unique significant ASV (q < 0.05) identified in the soils. Each column in the heatmap represents a specific contrast in the enrichment model. The heatmaps are coloured based on $\log_2$ fold changes derived from a fitted Generalized Linear Model (GLM). Positive fold changes (coloured in red gradient) represent enrichments on the upper side of the name of the contrast (e.g. Leaf vs. Root, enriched in leaf as compared to root). In contrast, negative fold changes (coloured in blue gradient) represent

enrichments on the bottom side of the name of the contrast (e.g. Leaf vs. Root, enrichment in root in comparison to leaf). Significant comparisons are outlined in a black frame ($q < 0.05$).

To explore single-leaf bacterial community composition, we compared the microbiota structure between the different leaves. We observed that the bacterial community diversity decreased from old leaves (L1 and L2) to young leaves (L5 and L6) (Figure 13), suggesting that different leaves harbour different bacterial diversity and that unique ecological mechanisms of selection operate in individual leaves configurate the bacterial niches.

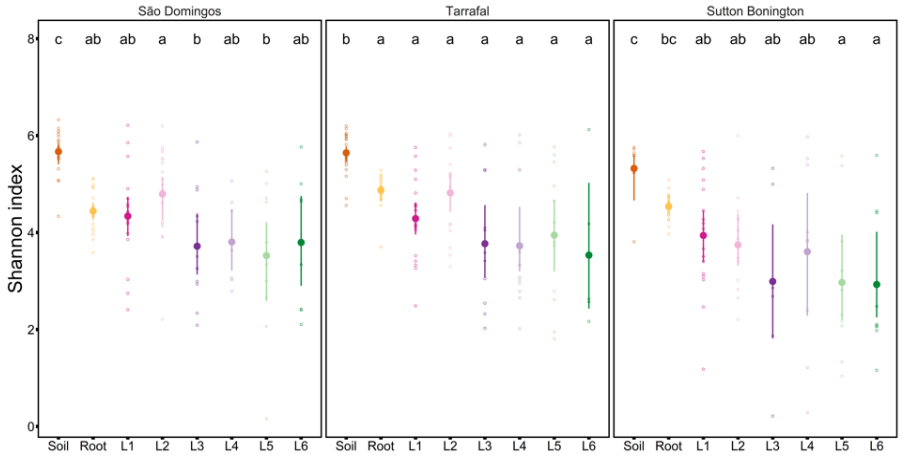


Figure 13: Single-leaves display reproducible effects on alpha-diversity. Representation of the bacterial α-diversity (estimated using Shannon diversity index) calculated for the different compartments analysed for plants growing in São Domingos, Tarrafal or Sutton Bonington soils. Letters represent the results of the *post hoc* Dunn test of the Kruskal-Wallis model.

Canonical analysis of principal coordinates (CAP) based on Bray-Curtis distances revealed significant differences in bacteria composition according to the leaf age (São Domingos: $R^2 = 0.08$, p-value < 0.01; Tarrafal: $R^2 = 0.08$, p-value < 0.01; Sutton Bonington: $R^2 = 0.1$, p-value < 0.01) (Figure 14). These results indicate that single-leaves accommodate distinct bacterial communities.

Next, we independently calculated the single-leaf-specific enrichment profile in the three soils. We identified five distinct clusters by estimating the ASVs abundance differences against the natural inoculum (single-leaf vs. soil) (Figure 15). We found that cluster 1 (Cl1) grouped ASVs belonging to Actinobacteria, Proteobacteria, Firmicutes, Gemmatimonadetes, and Bacteroidetes and were

predominantly more abundant in leaves from plants grown in Sutton Bonington soil as compared to those grown in Santiago island soils (Figure 15).

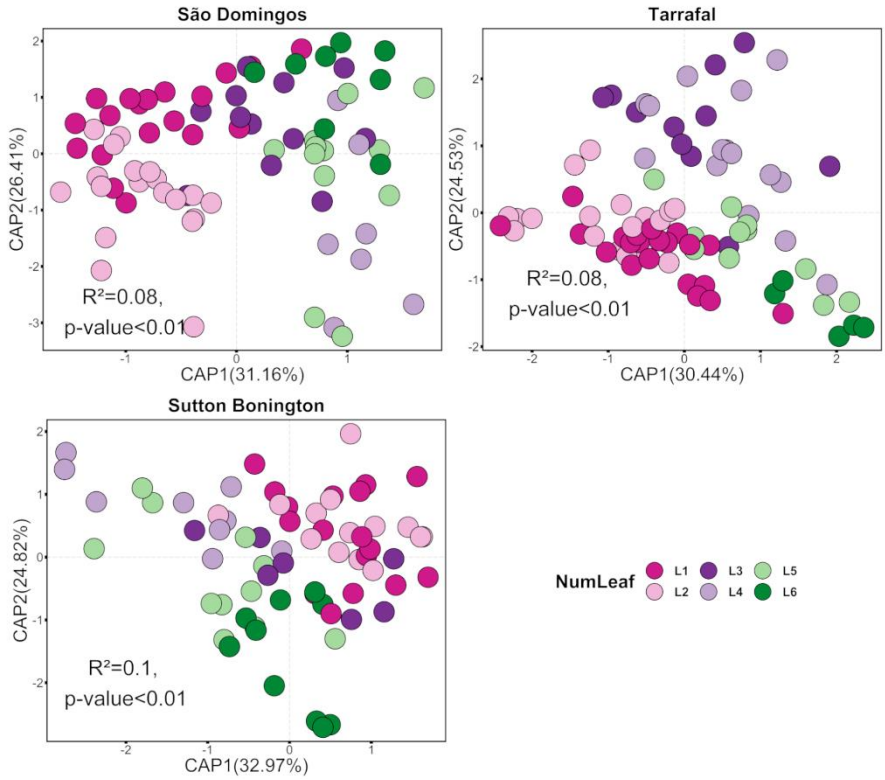


Figure 14: Single-leaves harbour distinct microbiota. Canonical analysis of principal coordinates showing the single-leaf effect on the composition of the bacterial community in each soil. PERMANOVA R^2 and p-value are shown within each plot.

Cl2 was composed by 6 ASVs belonging to Actinobacteria, Proteobacteria, and Gemmatimonadetes, and interestingly these ASVs were more abundant in leaves of plants growing in Sutton Bonington and São Domingos as compared to leaves of plants grown in Tarrafal soil (Figure 15). Cl3 was composed by ASVs belonging to Actinobacteria, Proteobacteria and Firmicutes phyla and they were predominantly more abundant in leaves from plants growing in Santiago soils as compared to those grown in Sutton Bonington soil (Figure 15). The Cl4 and Cl5 were composed mainly of Firmicutes and showed leaf enrichment across all soils (Figure 15). Interestingly, in Cl5, the bacteria

community convergence in L4 to L6 shifts from more similar communities between plants grown in São Domingos and Sutton Bonington soils (Figure 15).

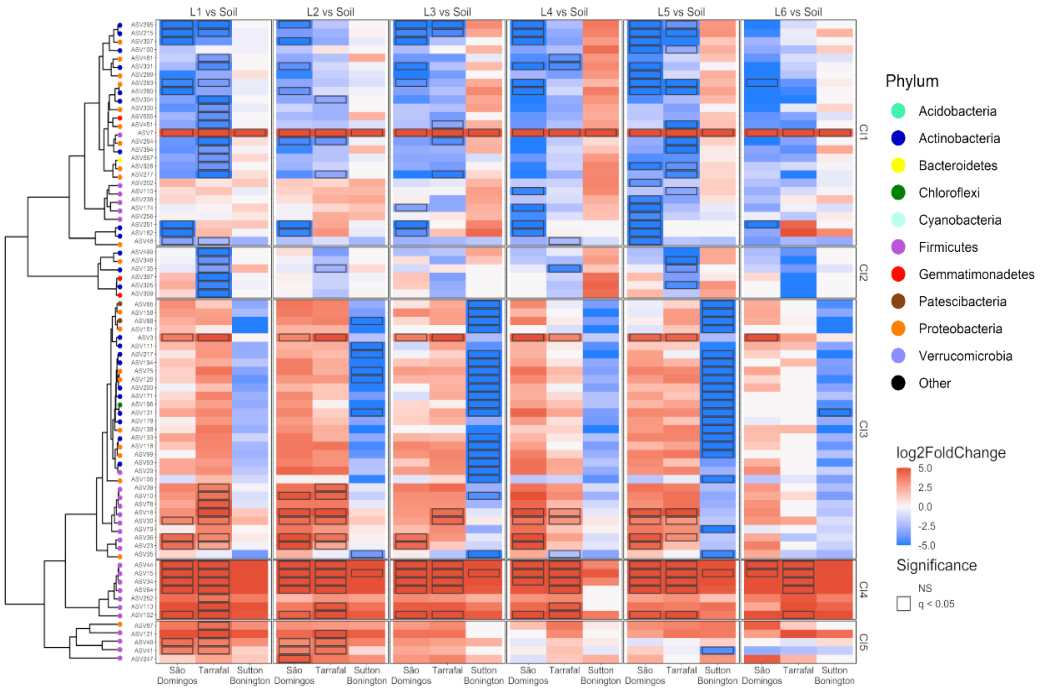


Figure 15: Bacterial community convergence shift from soil to leaves. Leaf enrichment patterns of the bacterial community across soils. Each row represents a unique ASV. Dendrogram rounded ends are coloured according to phyla. Heatmaps are coloured by log_2-transformed fold changes derived from a fitted generalized linear model and represent enrichments in maize leaves compared to soil. Comparisons with false-discovery-rate (FDR)-corrected *q*-value < 0.05 are framed in black. n=144 biological replicates from 2 independent experiments.

3.4 Different leaves assemble microbiota with distinct taxa abundance

Since single leaves showed different bacterial communities, we hypothesized that each leaf may have a bacterial community signature. To identify bacterial ASVs that are specific to individual leaves, we calculate their prevalence across individual leaves in each soil, independently. The ASVs detected in > 30% of samples and with high relative abundance were defined as specific to the correspondent leaf or leaf signature ASVs (Figure 16). The prevalent ASV revealed a total of 119 specific ASVs in leaf samples from plants

grown in São Domingos soils, and 132 and 133 specific ASVs in plants grown in Tarrafal and Sutton Bonington soils, respectively (Figure 16).

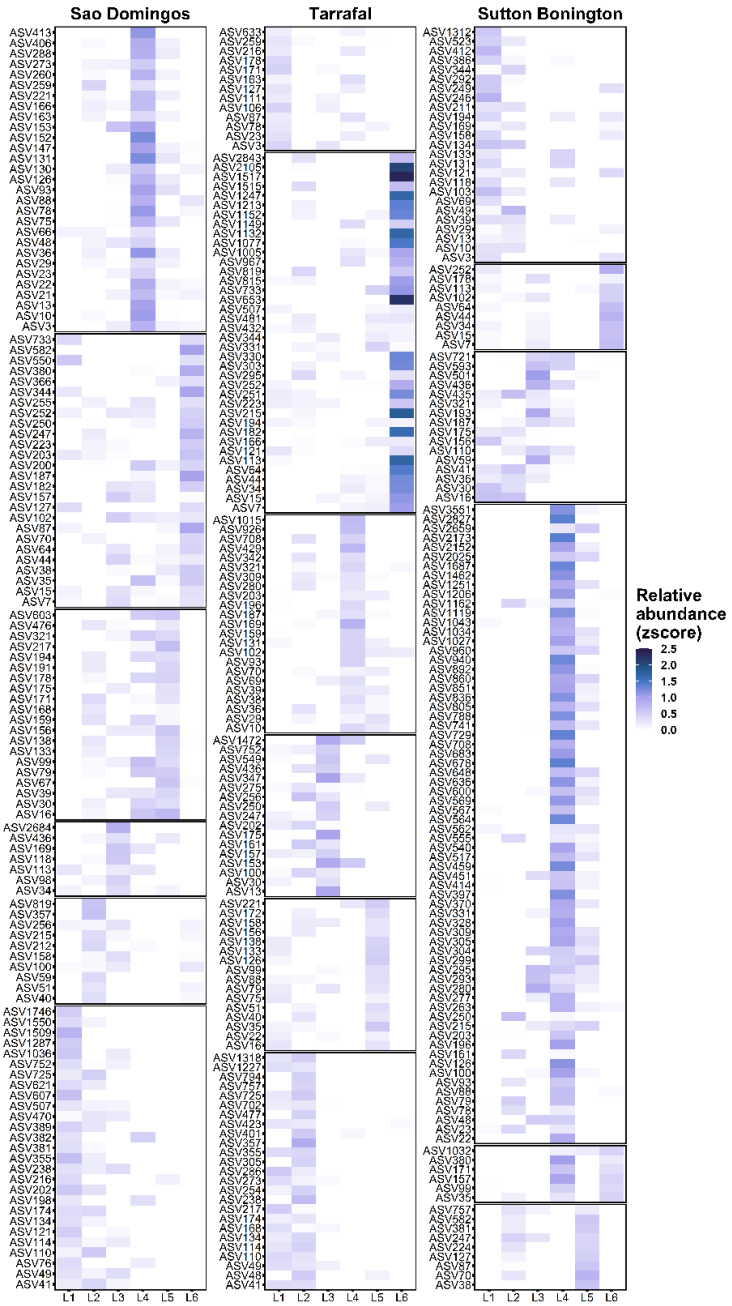


Figure 16: Bacterial community signature in single-leaves. Heatmap showing the positive z-score for the ASVs relative abundance in individual leaves. The different horizontal panels were created according to the cluster method ward.D2.

Analysis of inter-soil genera shared in leaves across soils revealed a substantial number of genera shared across plant leaves grown in São Domingos and Tarrafal soil. These genera were mainly shared in L3 and L5 (Figure 17). Conversely, plants grown in Tarrafal soil shared a smaller number of genera with plants grown in Sutton Bonington Soil (Figure 17). In general, plants grown in São Domingos and Sutton Bonington soil shared a larger amount of genera as compared to genera shared between Tarrafal and Sutton Bonington (Figure 17).

These data suggested that bacterial populations may adapt to individual maize leaves according to developmental stages, thus affecting their growth.

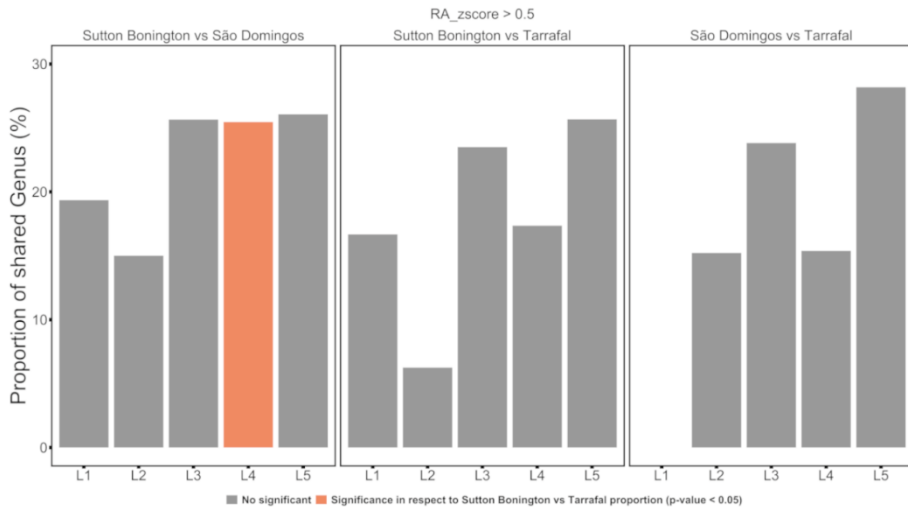


Figure 17: ASVs shared between the different soils. Bar graph showing the shared ASVs between leaves of plants grown in the different soils.

4 Discussion

Plant growth is complex, dynamic and multifactorial. Our study demonstrates that external cues, such as soil mineral content and soil microbiota modulate leaf growth and individual leaves of the plant differentially integrate their effect. We provided evidence that under realistic conditions, changes in leaf ionome profiles are not sufficient to explain differences found in maize plants grown in contrasting soils. Furthermore, the expected effects of soil mineral nutrients on plant growth might be overridden by the leaf microbial community composition, thus suggesting that regarding plant growth regulation, leaf microbiota can have an epistatic effect on mineral nutrients.

The capacity of plant leaves to individually integrate abiotic environmental perturbations has been suggested in another publication (Berens *et al.*, 2019). In that report, it is established that leaf growth in response to abiotic stresses is controlled by the differential activation of the immune system in old and young leaves. Therefore, we can speculate that this trade-off between plant growth and plant defence activation could be modulated by the presence of the plant microbiota. In that case, an unexpected growth promotion effect on leaves could be explained by changes in the activation of the plant immune system upon microbiota colonization. It is reasonable to believe that groups of bacteria colonising individual leaves have evolved the capacity to modulate the plant immune system to enhance or not leaf growth, bypassing the trade-off defence/growth in plants. This hypothesis needs to be tested in an ecological-relevant system.

Factors and mechanisms determining microbiota control of plant growth are beginning to be defined (Salas-González *et al.*, 2020; Hou *et al.*, 2020). In the context of the observed differential leaf microbiota, we can ask how these bacterial populations acquired the capacity to adapt to the specific conditions that individual leaves imposed. This adaptation could be through several mechanisms, such as the ability to extract nutrients, production of hormones and surfactants, as well as motility and biofilm formation (Nadakuduti *et al.*, 2012; Ueda *et al.*, 2018; Leveau, 2019; Oso *et al.*, 2019).

In conclusion, our work here describes the importance of the leaf environment in selecting different microbial communities in plants, likely modulating plant growth in a natural environment. This opens the possibility of developing microbial-based strategies to control plant growth at the organ level, raising opportunities to improve plant edible parts without affecting their nutritional content. Therefore, to reach such a final goal, a deeper understanding of the mechanisms of the different microbial communities in plant development is needed, thus requiring further studies using more appropriate methods to determine the role of the single-leaf community in the plant phenotype.

5 References

Bai, Y., Müller, D.B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., Dombrowski, N., Munch, P.C., Spaepen, S., Remus-Emsermann, M., Huttel, B., McHardy, A.C., Vorholt, J.A., Schulze-Lefert, P. (2015) Functional overlap of the Arabidopsis leaf and root microbiota. *Nature*, **528**, 364–369.

Berens, M.L., Wolinska, K.W., Spaepen, S., Ziegler J, Nobori T, Nair A, Krüler V, Winkelmüller TM, Wang Y, Mine A, Becker D, Garrido-Oter R, Schulze-Lefert P, Tsuda K. (2019) Balancing trade-offs between biotic and abiotic stress responses through leaf age-dependent variation in stress hormone cross-talk. *Proc Natl Acad Sci U S A*, **116**, 2364–2373.

Bilskie, J. and Campbell Scientific (2001) Soil water status: content and potential.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A. and Holmes, S.P. (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*.

Chaín, J.M., Tubert, E., Graciano, C., Castagno LN, Recchi M, Pieckenstain FL, Estrella MJ, Gudesblat G, Amodeo G, Baroli I. (2020) Growth promotion and protection from drought in Eucalyptus grandis seedlings inoculated with beneficial bacteria embedded in a superabsorbent polymer. *Scientific Reports 2020 10:1*, **10**, 1–17.

Chen, J., Xu, W., Velten, J., Xin, Z. and Stout, J. (2012) Characterization of maize inbred lines for drought and heat tolerance. *J Soil Water Conserv*, **67**, 354–364.

Durán, P., Thiergart, T., Garrido-Oter, R., Agler, M., Kemen, E., Schulze-Lefert, P. and Hacquard, S. (2018) Microbial Interkingdom Interactions in Roots Promote Arabidopsis Survival. *Cell*, **175**, 973-983.e14.

Hou, S., Thiergart, T., Vannier, N., Mesny, F., Ziegler, J., Pickel, B. and Hacquard, S. (2020) Microbiota-root-shoot axis modulation by MYC2 favours arabidopsis growth over defence under suboptimal light. *bioRxiv*, 1–29.

Kozak, M. and Piepho, H.P. (2018) What's normal anyway? Residual plots are more telling than significance tests when checking ANOVA assumptions. *J Agron Crop Sci*, **204**.

Leveau, J.H. (2019) A brief from the leaf: latest research to inform our understanding of the phyllosphere microbiome. *Curr Opin Microbiol*, **49**, 41–49.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*.

Nadakuduti, S.S., Pollard, M., Kosma, D.K., Allen, C., Ohlrogge, J.B. and Barry, C.S. (2012) Pleiotropic phenotypes of the sticky peel mutant provide new insight into the role of CUTIN DEFICIENT2 in epidermal cell function in tomato. *Plant Physiol*, **159**, 945–960.

Ogle, D., Wheeler, P. and Dinno, A. (2020) Package "FSA" Title Simple Fisheries Stock Assessment Methods.

Oksanen, J. (2007) vegan : Community Ecology Package. R package version 1.8-5.

Oso, S., Walters, M., Schlechter, R.O. and Remus-Emsermann, M.N.P. (2019) Utilisation of hydrocarbons and production of surfactants by bacteria isolated from plant leaf surfaces. *FEMS Microbiol Lett*, **366**, 61.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res*, **41**.

Salas-González, I., Reyt, G., Flis, P., Custódio V., Gopaulchan D., Bakhoum N., Dew T.P., Suresh K., Franke R.B., Dangl J.L., Salt D.E., Castrillo G. (2021) Coordination between microbiota and root endodermis supports plant mineral nutrient homeostasis. *Science (1979)*, **371**.

Sapeta, H., Costa, J.M., Lourenço, T., Maroco, J., Linde, P. van der and Oliveira, M.M. (2013) Drought stress response in Jatropha curcas: Growth and physiology. *Environ Exp Bot*, **85**, 76–84.

Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*, **9**, 671–675.

Soltanpour, P.N. and Schwab, A.P. (1977) A new soil test for simultaneous extraction of macroand micro-nutrients in alkaline soils. *Commun Soil Sci Plant Anal*.

Thierfelder, C., Rusinamhodzi, L., Setimela, P., Walker, F. and Eash, N.S. (2016) Conservation agriculture and drought-tolerant germplasm: Reaping the benefits of climate-smart agriculture technologies in central Mozambique. *Renewable Agriculture and Food Systems*, **31**, 414–428.

Ueda, H., Kurose, D., Kugimiya, S., Mitsuhara, I., Yoshida, S., Tabata, J., Suzuki, K. and Kitamoto, H. (2018) Disease severity enhancement by an esterase from non-phytopathogenic yeast Pseudozyma antarctica and its potential as adjuvant for biocontrol agents. *Sci Rep*, **8**, 16455.

Vega, A., O'Brien, J.A. and Gutiérrez, R.A. (2019) Nitrate and hormonal signaling crosstalk for plant growth and development. *Curr Opin Plant Biol*, **52**, 155–163.

Wit, M. de, Galvão, V.C. and Fankhauser, C. (2016) Light-Mediated Hormonal Regulation of Plant Growth and Development.

Chapter IV

Microbiota of individual maize leaves modulate the growth-defence trade-off to promote leaf growth

This chapter was adapted from:

Custódio, V., Salas-González, I., Gopaulchan, D., Flis, P., Gao, Y., Amórós, R., Moreno, A., Oliveira, M.M., Salt, D. and Castrillo, G. (2023), Individual leaf microbiota tunes a genetic regulatory network to promote leaf growth in maize. (*In preparation)*.

Author's contribution

Custódio, V., Salas-González, I., and Castrillo, G. conceptualize the project and perform data curation. Custódio, V. and Salas-González, I., formally analyse the data. Custódio, V., Salas-González, I., Gopaulchan, D., Flis, P., Gao, Y., Amórós, R., Moreno, A., and Castrillo, G. perform the investigation. Custódio, V., and Castrillo, G wrote the original draft. Custódio, V., Oliveira, M.M., Salt, D. and Castrillo, G. Contributed to the text and corrected the different versions of the article.

Abstract

In natural ecosystems, microbes have evolved the ability to modulate the development of healthy plant organs. However, it is still poorly understood how the microbiota inhabiting the phyllosphere influence individual leaf growth. Here, we built a bacterial strain collection that we used in recolonization experiments to study the microbial mechanisms involved in the leaf growth-promoting effect. We established that highly abundant bacteria inhabiting individual young leaves promote the growth of individual leaves. Using transcriptomics analyses, we revealed a genetic network that integrates the beneficial effect of the phyllosphere microbiota into the leaf developmental program. We demonstrated that microbiota-driven repression of a cluster of defence-related genes allows the plant to modulate the growth-defence trade-off at single-leaf resolution to modulate plant leaf growth. Our results explain the growth promotion effect caused by the microbiota in natural soils. This knowledge can be used in agricultural settings to increase crop yields.

1 Introduction

Similarly to other biological systems, including animals, plants harbour populations of microbes functionally adapted to colonize their different organs, enhancing their function and, as a consequence, the health of the entire organism (Custódio *et al.*, 2022; Hacquard *et al.*, 2015). This organ compartmentalization is driven by the microbial ability to thrive under the ecological constraints imposed by different plant organs (Bai *et al*., 2015; Massoni *et al.,* 2021). Plant leaves constitute a very unstable microbial niche, relatively poor in nutrients and permanently affected by multiple environmental stresses (Murillo-Roos *et al.*, 2022). The microbes inhabiting this stressful and selective environment that change along development and ageing (Wagner *et al.*, 2016) are known to confer some protection against plant pathogens (Vogel *et al.*, 2021).

To successfully colonise the leaves, microbes developed the capacity to efficiently extract nutrients (Murillo-Roos *et al.,* 2022; Bai *et al.,* 2015), produce hormones (Eichmann *et al.*, 2021), and modulate the plant immune system (Vogel *et al.*, 2021; Ma *et al.*, 2021; Teixeira *et al.*, 2019). These microbial capacities are likely to affect the endogenous program of leaf development that under axenic conditions is controlled by the phytohormones auxins, cytokinins, and gibberellins (Nakayama *et al.*, 2022). Other factors such as nutrient availability (Walter and Schurr, 2005), pathogen infection (Singkaravanit-Ogawa *et al.*, 2021), and individual members of the root microbiota have been described as influencing leaf growth (Berens et al., 2019; Janie Zhang et al., 2021). However, how the leaf integrates its resident functional microbiota into its developmental program is still not fully understood.

Aiming to contribute to better understand the role of leaf microbiota, we have used a bacterial culture collection, obtained from leaves, roots, and rhizosphere in synthetic community experiments. We could establish that bacterial populations adapted to colonise young leaves promote individual leaf growth. Furthermore, using a transcriptomics approach we identified a new genetic network controlling the integration of the beneficial effect of the leaf microbiota into the plant leaf developmental program. This regulatory mechanism involves the repression of a defence-related gene cluster thus

allowing the plant to integrate the beneficial effects of the microbiota in leaf growth promoting programs. This defence-related mechanism can explain a leaf growth promotion effect, taking in place in natural soils with complex microbiota.

2 Materials and methods

2.1 Sampling of maize plants and isolation of root-, leaf-, and soil-derived bacteria

To obtain a taxonomically diverse representation of maize root and leaf microbiota, we used tissues from different maize plant genotypes (B73, Matuba, CVSt, and Zm523) grown in the three soils described in Chapter III. Briefly, fragments of maize roots and leaves were collected from plants at V4 developmental stage (leaf 4 fully expanded) and washed twice with PBS 1x, in the case of roots, and with 10 mM $MgCl_2$ for the leaves (Bai et al., 2015; Jingying Zhang et al., 2021). As a control we collected rhizosphere soil samples that were washed with 1X PBS buffer, supplemented with 0.02% Silwet L-77. The supernatant buffer with the microbiota was directly inoculated in liquid culture media.

Roots and leaves were crushed using sterile mortars and pestles with 5 ml of 10 mM $MgCl_2$ under sterile conditions. The resulting homogenates were diluted to 1:200.000, and 100 µL of each dilution was plated on ten different growth media (Table 1). Plates were grown in an incubator at 28 ºC, for two-three weeks.

Bacterial colonies with different morphologies and colours were individually transferred from each plate to freshly prepared LB plates. Individual bacterial strains were isolated by successive sub-cultures to new LB plates until no signs of contamination could be detected. For every individual bacterial isolate, we prepared a LB liquid culture that was mixed (1:1, v/v) with 80% glycerol (v/v) and stored at −80°C for long-term conservation.

Table 1: Culture media used for bacteria isolation. All media (Bai *et al.,* 2015) are appropriate to growth Actinobacteria, Bacteroidetes, Firmicutes, Proteobacteria.

Media	Compartment target
Trypic Soy Broth (TSB)	Root, Soil
Tryptone Yeast Extract Glucose (TYG) Medium	Root, Soil
Yeast Extract Mannitol (YEM) Medium	Root, Soil
M408 Medium	Root, Soil
M715 Medium	Root, Soil
Flour medium	Root, Soil
Tap Water Yeast Extract (TWYE) Medium	Root, Soil
MYX Medium	Leaf
Minimal Media plus Methanol (MM + MeOH) Medium	Leaf
R2A Medium	Leaf

2.2 Characterization of the bacterial strain collection. Sanger sequencing

To confirm the purity of the bacterial cultures and remove possible duplicates in the collection, we amplified, and Sanger sequenced bacterial *16S rRNA* of all individual cultures. Prior to *16S rRNA* amplification, total DNA was extracted from the individual glycerol stocks following the protocol described by Bai et al. (2015) Briefly, 6 µL of each individual stock culture was diluted in 10 µL of buffer I (25 mM NaOH, 0.2 mM EDTA, pH 12) applied in each slot a 96-well PCR plate). The bacterial culture was lysed at 95 ºC for 30min, and the lysis was stopped by neutralization with buffer II (40 mM Tris-HCl, pH 7.5, 10 µL applied per slot). Using the resulting solutions as templates, the V1-V9 region of the bacterial *16S rRNA* gene was amplified with the primers 27F (5′-AGRGTTTGATCMTGGCTCAG-3′) and 1492R (5′- ACCTTGTTACGACTT-3′). PCR conditions were as follows: 12.5 µL of Readymix, 0.5 µL of primer 27F, and 0.5 µL of primer 1492R in a final volume of 50 µL. The thermocycling

conditions were: denaturation at 95 ºC for 5 min, followed by 34 cycles of 95 ºC for 30 s, 55 ºC for 30 s for primer annealing, and 72 ºC extension for 1 min, with a final elongation step at 72 ºC for 5 min. PCR products were purified using QIAquick Gel Extraction Kit (product) following the manufacturer's specifications. The PCR products were sequenced using Sanger technology with the 27F primer.

2.3 Curation of the bacterial strain collection

To select a representative group of strains encompassing all the diversity of the maize bacterial collection, we defined three criteria for selection. The first criterion was to select the abundant genera identified in the natural community experiment (Chapter III – section 3.3 and section 3.4). The second criterion was to select the low abundant genera, and the third was to maximize phylogenetic diversity by exploring the tree topology, while selecting the strains. Additionally, we wanted to add to the collection other leaf derived-strains described as maize phyllosphere bacteria (Wallace *et al.*, 2018). For this, we isolated 15 more strains from leaves using the Minimal medium and R2A (Table 1).

2.4 Bacterial strains whole genome sequencing. DNA extraction

To fully characterise the curated collection of bacterial strains, we sequenced the genomes of each selected isolate. For bacterial DNA purification, we used a modified version of the DNA isolation protocol of Bai et al. (2015) modified by Finkel et al. (2020). Each bacterial strain was grown individually in 4 mL LB for 24h to 48h, until saturation, and bacterial cells were collected by centrifugation at 3,220 g for 15 min. In each case, pellets were resuspended in 5 mL SET buffer (75 mM NaCl, 25 mM EDTA, 20 mM Tris-HCl, pH 7.5). Cellular suspensions were then digested with 20 µL of 50 mg/mL Lysozyme (Sigma) at 37°C, for 30 min. After the cell lysis, 2 mL of 5 M NaCl and 5 mL of chloroform were added to each sample, followed by agitation in an orbital shaker for 30 min. Samples were then centrifuged at 3,220 g for 15min, and 6 mL of the aqueous phase was transferred to a fresh Falcon tube. DNA was precipitated with 3.6 mL isopropanol at -20°C, overnight. Samples were centrifugated at 3,220 g for 5min, and the pellet washed with 1 mL 70% ethanol. Ethanol was eliminated by centrifugation at 3,220 g for 5min and the

pellet dried at room temperature for 15min. The final DNA was dissolved in 50 µL of water. To remove possible contamination with RNA, all samples were treated with 2µL of 4 mg/mL RNase A (Sigma-Aldrich Chemie GmbH) and incubated overnight at 4 °C. The DNA was quantified using Qubit dsDNA HS (Invitrogen).

2.5 Bacterial strains whole genome sequencing. DNA library preparation

For the libraries preparation, isolated bacterial genomic DNA was digested with NEBNext dsDNA Fragmentase (New England Biolabs), according to the manufacturer's instructions. Briefly, 200 ng of DNA of each strain were digested with a 1:1 (v/v) mixture of 1x Fragmentase Reaction Buffer V2 and 1X dsDNA Fragmentase in a final volume of 20 µL at 37°C for 15 min. Digestion was stopped by adding 5 µL of 0.5 M EDTA and the samples diluted with an equal volume of nuclease-free water. DNA fragments with sizes between 300-500 bp were selected using AMPure XP magnetic beads (Beckman Coulter).

DNA fragments were end-repaired using 2.5 µL of 3U/µL T4 DNA polymerase, 0.5 µL of 5U/µL Klenow DNA polymerase, 2.5 µL of 10U/µL T4 PNK, 5 µL of 10X T4 DNA ligase buffer with 10 mM ATP, 0.8 µL of 25 mM dNTP mix, and 8.7 µL of nuclease-free water. The reactions were incubated at 20°C for 30 min and the resulting fragments were purified using AMPure XP beads. Next, DNA fragments adenylation was performed using 18 µL of the end-repaired DNA, 3 µL of 5U/µL Klenow exo$^-$, 5 µL of 10x Enzymatics Blue Buffer, 1 µL of 10mM dATP, and 9 µL of nuclease-free water. The mixture was incubated at 37^{o}C for 30min and then at 70^{o}C for 5min, to stop the reaction. The final product was purified using AMPure XP magnetic beads. For the adapter ligation, DNA fragments were treated with 1 µL of 600U/µL T4 DNA ligase, 12.5 µL of 2X Rapid Ligation Buffer, and 2.5 µL of the adapter index (2.08 µL of KAPA Adapter Dilution Buffer and 0.42 µL of 15 µM KAPA Dual-Indexed Adapters). Samples were incubated at room temperature for 15min and then the reaction was stopped by adding 5 µL of 0.5 M EDTA. The final products were purified twice using AMPure XP beads.

The DNA fragments containing the adapters were enriched by PCR using the following conditions: 25 µL of 2X KAPA HiFi HS Mix, 2.5 µL of 5 µM I7 primer, 2.5 µL of 5 µM I5 primer, and the following temperature cycling: 98 °C for 45 s; 14 cycles of 98 °C for 15 s; 60 °C for 30 s; 72 °C for 30 s; 72 °C for 1min; 12°C until use.

Finally, PCR products were purified using AMPure XP magnetic beads, and library yields were measured using Qubit dsDNA HS (Invitrogen). Library size was assessed using the High Sensitivity D1000 ScreenTape (Agilent) on the Agilent 4200 TapeStation System, and equimolar quantities of individual barcoded DNA libraries were pooled (3 pools). The library pools were sequenced in the Beijing Genomics Institute (BGI) using MGI Tech MGISEQ-2000 sequencing platform to generate a minimum of 10 million paired-end, 150 bp reads /sample

2.6 Genome assembly

To generate the draft genome sequence of the different bacterial strains, paired-end Illumina reads were trimmed using Sickle (v0.7.0) (https://github.com/najoshi/sickle) and *de novo* assembled using SPAdes (v3.15.5) (Bankevich *et al.*, 2012). Then, the draft genome for each sample was ordered and oriented into longer sequences using 'scaffold' function and default parameters from RagTag (v2.1.0) (Alonge *et al.*, 2022).

2.7 Phylogenetics analysis based on whole genome and proteins comparison

The pairwise average nucleotide identity (ANI) alignment was performed for the bacterial genomes using fastANI (v1.32) (Jain *et al.*, 2018), and hclust from the R package was used to cluster at the cutoff species level (ANI ≥ 97%). GTDB-Tk (v2.1.0, database r86, 'classify_wf' function and default parameters) was used to perform taxonomic annotation of each genome and reconstruct the maximum-likelihood phylogenetic tree based on 120 conserved single-copy genes (Chaumeil *et al.*, 2020). The phylogenetic tree was viewed using the display and annotation tool iTOL (v6.1.1) (Letunic and Bork, 2021).

2.8 Experiments with bacterial synthetic communities

2.8.1 Bacterial growth conditions

One week before each experiment, the 211 bacterial strains used in the synthetic communities were individually recovered from the glycerol stocks and inoculated in 96 DeepWell plates containing 1 mL of LB medium per wells. Bacterial cultures were grown in an incubator at 28 °C, with agitation at 120 r.p.m for 5 days. One hundred microliters of all cultures were then inoculated in new 96-deep-well plates containing 1 mL of LB medium and were grown in the same conditions for another 48 h. Both sets of cultures were combined and centrifuged (Eppendorf 5810R) at 3,220 g for 8 min. To remove used medium and cell debris, all bacterial pellets were washed three times with 10 mM $MgCl_2$. Bacterial cells were then harvested by centrifugation, resuspended in 1 mL of 10 mM $MgCl_2$, and the OD_{600nm} was measured. To equalise bacterial cultures, we assumed that $1OD_{600nm}$ unit was equal to 10^9 c.f.u/mL, and a volume equivalent to 10^5 c.f.u/ml from each individual culture was mixed to build the full synthetic community and the drop-out communities. For the leaf markers, four drop-out synthetic communities were built: i) leaf 1 and leaf 2 markers were excluded from the full synthetic community; ii) leaf 3; iii) leaf 4 and iv) leaf 5 markers were excluded from the full synthetic community. The resulting synthetic communities' mixtures were adjusted to an $OD_{600nm} = 0.2$ and 2 mL of each final synthetic community was mixed with 400 mL of Hoagland solution. The individual pots containing maize seeds were then supplemented with 64 mL of Hoagland solution inoculated with the bacterial synthetic communities.

2.8.2 Plant Growth conditions

We used a sterilised mixture of sand and clay 5:1(v/v) for the synthetic community experiments to fill 9-cm-diameter pots. Maize seeds were pre-germinated in water plates supplemented with 1% bacto-agar for 48 h at 30°C before being transferred to the pots. Half of the pots were then supplemented with 64 ml of ¼ Hoagland solution (Low nutrient condition) and the other half of the pots was supplemented with full Hoagland solution (high nutrient condition). All pots were inoculated with the different bacterial combinations as described above. As a control, we grew maize plants in pots filled with the same sterilized

mixture of sand and clay, supplemented with the two solutions, and without bacteria (Control).

All pots were randomly placed in ethanol-sterilised 63 L plastic boxes. To allow the gas exchange while maintaining sterility, we opened 9 windows on the lid of the boxes that were covered with Breath-easy gas permeable sealing membrane (Research Products International). The boxes were transferred to a growth chamber, and plants were grown under a 16-h light/8-h dark regime at 26°C day/21°C night until the plants reaching the V4 developmental stage. All the experiments were performed in two independent replicates, and each replicate included six pots per condition.

2.8.3 Leaf ionome

Samples for ionomic analysis were collected from plants at V4 stage. Samples were analysed as described in Chapter III (section 2.2.3).

2.8.4 Leaf size quantification

Maize plants were manually imaged with an EOS 4000D DSLR camera (Canon) using an 18-55 mm lens in the automatic mode. These images were used to determine individual leaf length using the Measure function in ImageJ software (Schneider *et al.*, 2012).

2.8.5 16S rRNA Sequences analysis. Synthetic community

Synthetic community sequencing data were demultiplexed and trimmed with cutadapt (Martin, 2011). The resulting sequences were then merged and aligned to a reference set of 16S rDNA sequences extracted from genome assemblies of members of the synthetic community. The reference database also included sequences from known bacterial contaminants and maize organellar sequences. Sequence alignment was performed with USEARCH v.7.1090 (Edgar, 2010) with the option usearch_global at a 98% identity threshold. On average, 85% of sequences matched an expected isolate. Our 211 isolates could not all be distinguished from each other based on the V3–V4 sequence and were thus classified into 106 unique sequences. A unique sequence encompasses a set of identical (clustered at 100%) V3–V4 sequences coming from single or multiple isolates. Sequence mapping results

were used to produce an isolate abundance table. The remaining unmapped sequences denoised and collapsed into amplicon sequence variants (ASVs), using DADA2 v.1.10.1 (Callahan et al., 2016). The resulting ASVs were mapped to SILVA 132 database (Quast et al., 2013). We filtered ASVs that were assigned to the chloroplast, mitochondria, oomycete, archaea or did not have a known kingdom assignment. The vast majority of the remaining unassigned ASVs belonged to the same genus as isolates in the synthetic community. We combined the assigned unique sequence and unassigned ASVs count tables into a single count table. In addition to the raw count table, we created rarefied (1,000 reads per sample) and relative abundance versions of the abundance matrix for further analyses.

2.8.6 Statistical analysis

The resulting abundance tables were processed and analysed with functions from the ohchibi package (https://github.com/isaisg/ohchibi). An α-diversity metric (Shannon diversity) was calculated using the diversity function from the vegan package v2.5-5 (Oksanen, 2007). We used Kruskal-Wallis to test for alpha diversity differences between the fractions and single-leaf. β-Diversity analyses (principal coordinate analysis and canonical analysis of principal coordinates (CAP)) were based on Bray–Curtis dissimilarity calculated from the relative abundance matrices. We used the capscale function from the vegan R package v2.5-5 to compute the CAP. To analyse the full dataset (all fractions, and single leaves), we constrained by fractions and single leaves while conditioning for the replica and experiment effect. In addition to CAP, we performed PERMANOVA using the adonis function from the vegan package v2.5-5. We used the package DESeq2 v.1.24.0 (Love et al., 2014) to compute the enrichment profiles for unique sequences present in the count table.

2.9 Experiments to define the molecular mechanisms regulating the microbiota effect on individual leaf growth

2.9.1 RNA extraction

RNA was extracted from maize leaf samples following Logemann et al. (Logemann *et al.*, 1987). Briefly, approximately 5 cm of leaf from the tip were harvested after discarding the first 1 to 2 cm of each leaf of wild-type maize and

flash frozen in liquid nitrogen. Frozen tissue was pulverized in a TissueLyzer II (Qiagen), using 2 cycles of 1 min, frequency 30 s^{-1}. We added to each sample 400 µL of Z6-buffer (8M guanidine HCl, 20 mM MES, 20 mM EDTA at pH 7.0). Then, samples were treated with 400 µL phenol:chloroform:isoamylalcohol; 25:24:1 (Sigma), vortexed, and centrifuged at 20,000 g for 10 min to separate the phases. The aqueous phase, containing the RNA, was transferred to a clean 1.5 ml tube and the RNA was precipitated by adding 0.05 volumes of 1N acetic acid and 0.7 volumes of 96% ethanol and incubated at -20°C overnight. Samples were centrifuged at 20,000 g for 10 min, at 4°C, and the pellet was washed sequentially with 200 µL sodium-acetate (pH 5.2) and 70% ethanol. The RNA was dried for ten minutes, resuspended in 30 µL of ultrapure water, and stored at −80°C until use.

2.9.2 RNAseq

Isolated RNA was treated with Turbo DNA-free (Applied Biosystems), according to the manufacturer's instructions to remove contaminating DNA. RNA libraries were then prepared using 1 µg RNA following Finkel et al (Finkel *et al.*, 2020) with modifications. Briefly, mRNA was purified from total RNA with Sera-mag oligo(dT) magnetic beads (GE Healthcare Life Sciences) and fragmented at 94 °C for 6 min. Fragmented mRNA was used for first-strand cDNA synthesis using reverse transcriptase and random hexamers, then the second-strand cDNA was synthesized using DNA polymerase I and the contaminating RNA removed using RNaseH. Double-stranded cDNA was end-repaired using T4 DNA polymerase, T4 polynucleotide kinase, and Klenow polymerase. The adenylation of the DNA fragments were performed using Klenow exo-polymerase to allow the ligation of KAPA Dual-Indexed adapters (KK8722).

Library yield was measured by Qubit dsDNA HS (Invitrogen; Q32851) and library size was assessed using the High Sensitivity D1000 ScreenTape (Agilent; 5067- 5584) on the Agilent 4200 TapeStation System (Agilent; G2991A). Equimolar quantities of individual barcoded RNA libraries were pooled in a randomized manner, and shipped in dried ice to Beijing Genomics Institute (BGI), Shenzhen, China. Each library pool was sequenced on a MGI

Tech MGISEQ-2000 sequencing platform to generate a minimum of 10 million paired-end, 100 bp reads per sample.

2.9.3 Phyto-hormones quantification

The hormone profile was determined in maize leaf 3 (L3) from plants growth with SynCom. For the L3, 50 mg of the leaf fragments were placed in a weighted 1.5-mL Eppendorf tube and flash-frozen in liquid nitrogen. Subsequently, the frozen tissue was shipped in dried ice to Servicio de Cuantificacion de Hormonas Vegetales, IBMCP (CSIC/UPV), Spain, for hormone profile analysis.

2.9.4 Quantification of the size and number of leaf cells

The cell size and number from the individual leaves were quantified from pictures obtained on a Leica TCS SP5 confocal microscopy on the bright field mode. Briefly, we fixed the single-leaf fragments in 100% ethanol for one week. To undergo bleaching and chlorophyll removal, ethanol was replaced with 80% acetone and incubated at 4°C for 2-3 days. The cellular content of the leaf was removed by incubating the leaf fragments with 10% KOH for 4h at 85°C. The transparent tissue originated was analysed at confocal microscopy, and cell size and number were quantified using ImageJ software.

2.9.5 Computational and statistical analysis

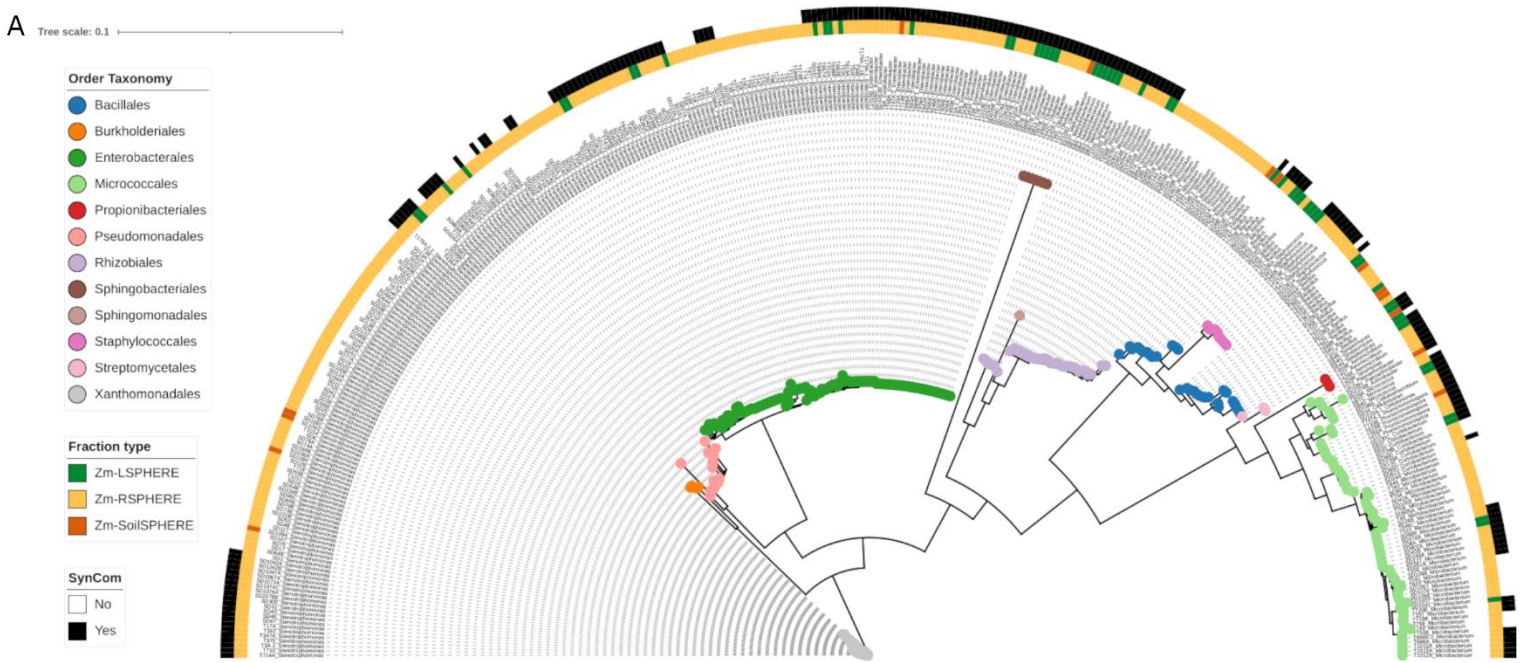
For the analysis of the RNASeq data, reads were first subjected to quality control and filtering via fastp v0.20.1 (Chen *et al.*, 2018). The resulting high-quality reads were pseudo-aligned to the maize genome deposited in the Ensembl Plants version 52, using salmon v1.6.0 (Patro *et al.*, 2017). Gene level summarization from the salmon quantification was performed with the tximport function from the tximport v.1.22.0 R package (Soneson *et al.*, 2015).

Gene ontology analysis was performed over each of the delimited clusters employing the function compareCluster function from the clusterProfiler v.4.2.2, R package (Yu *et al.*, 2012).

3 Results

3.1 Maize-derived bacterial culture collection

To demonstrate how single-leaf bacterial populations control leaves growth in maize, we first generated a bacterial culture collection. We isolated individual bacterial strains from the roots and shoots of 12 *Zea mays* plants at the V4 stage grown in three natural soils from Cabo Verde and the UK. Hundreds of shoot and root fragments, three times washed with $MgCl_2$, were homogenized and the resulting solution was inoculated on ten agar-based growth media with different bacterial selection capacities. Using this method, we were able to isolate 2169 bacterial strains. Initially, the taxonomy assignments of the isolates was made using Sanger sequencing data of the bacterial 16S rRNA gene. With this taxonomical information we could remove strain duplications that in our case were defined as bacterial strains with the same morphology and 100% identical 16S rRNA sequences, resulting in 395 pure bacterial isolates belonging to the classes gammaproteobacteria (59.8%%), actinobacteria (18.3%), alphaproteobacteria (10.3%), bacilli (9.8%), and bacteroidetes (1.8%) (Figure 1A). We checked that the resulting bacterial collection approximated, at the order level, the microbiome taxonomic composition of maize plants grown in natural soils, as defined using culture-independent 16S rRNA sequencing (Figure 1B). Our results showed that it is possible to isolate the most abundant bacterial classes colonizing maize tissues. Moreover, this collection provides us with a tool to define the molecular mechanisms underpinning leaf growth in recolonization experiments under controlled conditions.


A
Tree scale: 0.1
Order Taxonomy
Bacillales
Burkholderiales
Enterobacterales
Micrococcales
Propionibacteriales
Pseudomonadales
Rhizobiales
Sphingobacteriales
Sphingomonadales
Staphylococcales
Streptomycetales
Xanthomonadales
Fraction type
Zm-LSPHERE
Zm-RSPHERE
Zm-SoilSPHERE
SynCom
No
Yes

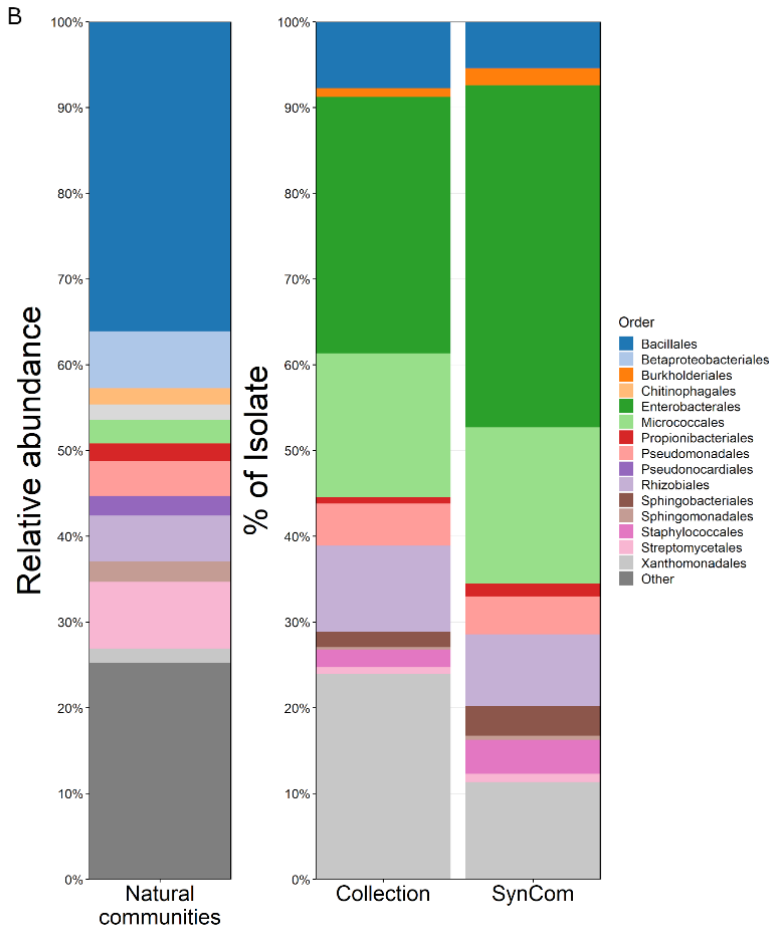


Figure 1: Taxonomic representation of the bacteria isolated from maize plants and soil, and their representation in the profiling study of natural communities. A) Phylogenetic tree built using 16S-sequences of 395 bacterial strains obtained from surface-sterilized maize root, leaves and soil. The tree has tips coloured according to the bacterial taxonomic order. The coloured outer ring represents the leaf (*Zm*-LSPHERE), roots (*Zm*-RSPHERE) and soil (*Zm*-SoilSPHERE) fractions where the strain was isolated. The black ring shows the strain selected to build the synthetic community (SynCom). **B)** Comparison of bacterial taxonomic order proportions in the natural communities' study, bacterial collection and SynCom. The "Natural communities" represents the relative abundance profiles of the most abundant bacterial order identified in the roots and leaves of maize plants grown in soils from Cabo Verde and UK. The proportion of isolates at the order levels in the collection and SynCom collection is also shown.

3.2 Individual leaves bacterial communities influence leaf growth

To understand how microbial function influences the endogenous leaf development program, we designed a clay-sand-based system that reduced the complexity of soil structure. This system reproduced the water retention capacity found in the Sutton-Bonington soil, the highest observed in the three soils used (Figure 2A)**.** Using this system under axenic conditions (no microbiota present), we first determined the individual leaves length in maize plants grown in four different concentrations of nutrients (no nutrient supplementation, ¼ Hoagland, ½ Hoagland, and full Hoagland). We observed that shoots of B73 maize plants grew following a linear trajectory driven by the concentration of mineral nutrients in the mixture of clay-sand used (Figure 2B). This effect was also observed at single-leaf resolution; the leaves of plants exposed to an increasing gradient of nutrient concentrations followed a linear pattern of increased growth (Figure 2C). In agreement with the data retrieved from our experiments in natural soils (Chapter III), the mineral nutrient concentration in the clay-sand matrix influenced the individual leaf ionome (Figure 2D) This indicates that, under sterile conditions, non-toxic concentrations of mineral nutrients in the soil positively correlate with the plant leaf growth. Therefore, nutrient-rich and nutrient-poor soil will always produce plants with larger and smaller leaves, respectively. Based on these observations, we hypothesized that the unexpected leaf growth promotion observed in plants grown in the São Domingos natural soil might be due to the population of microbes inhabiting the leaves.

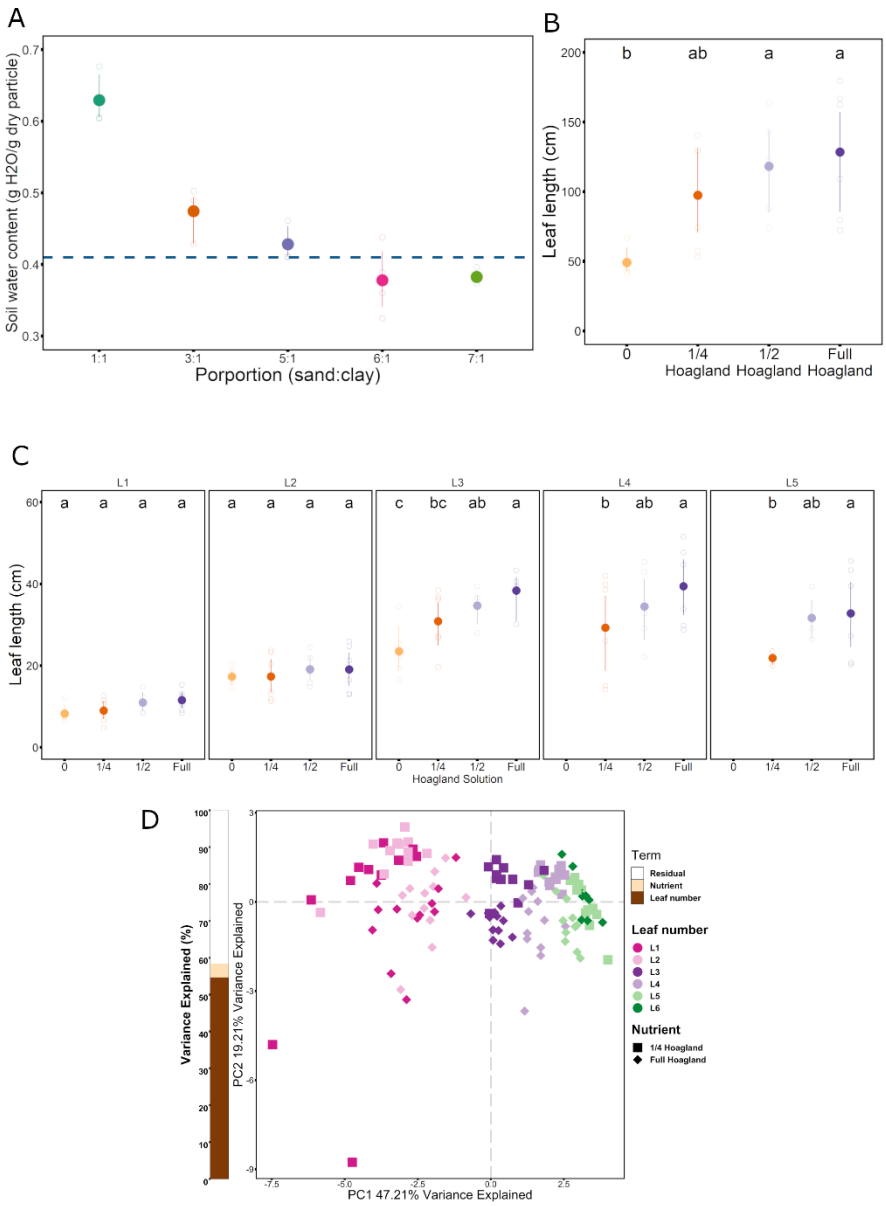


Figure 2: Assessment of maize leaf growth in the absence of bacteria, showing that soil nutrient content is the main driver of leaf growth. A) Soil water content in the different clay-sand matrix. The blue line shows the soil water content in Sutton Bonington soil, the soil with the highest water retention compared to the two soils in Cabo Verde. **B)** Pointrange plots showing the total length of maize leaf growth with different nutrient concentrations. **C)** Individual leaf length (in cm) of four-week-old maize plants grown in different nutrient concentrations. In

plot **B)** and **C)**, letters indicate statistical significance corresponding to the ANOVA test with Tukey's *post hoc* test (α=0.05). **D)** PCA analysis based on Euclidean distance of the leaf mineral composition (ionome) showing the ionomic profile of maize leaves growth under axenic conditions and different nutrient concentrations. The bar plot to the left of the PCA represents the percentage of variance explained by statistically significant ($p < 0.05$) variables in a PERMANOVA model. Different colours show the individual leaves and different shapes show the different nutrients used.

To demonstrate this hypothesis, we designed a 211-bacterial member's synthetic community using our collection of bacterial isolates (Figure 3). The synthetic community bacterial members have unique 16S rRNA sequences and this selection covered 72% of the total relative abundance of all bacterial order found in leaves of maize plants grown in natural soils (Figure 4). To fully assign taxonomical attributes to the synthetic community, we have sequenced the genomes of all individual bacterial strains. The bacterial synthetic community was inoculated in pots containing pre-germinated B73 maize seeds sown in a sterile clay-sand mixture supplemented with two concentrations of Hoogland medium, low (1/4 Hoagland) and full, as a control, we also grew B73 seeds under axenic conditions (without the bacterial synthetic community) in both nutrient concentrations. We observed a growth promotion effect in plant leaves of inoculated maize plants, compared to the uninoculated control plants (Figure 5A). Although the positive microbiota effect on leaf growth was observed in both nutrient concentrations tested, this effect was more obvious in the case of plants grown under low nutrient condition (Figure 5A). Under this condition, the bacterial synthetic community enhanced the plant growth to the level found in plants grown with sufficient nutrition (Figure 5A). Consistent with our previous experiments using natural soils, described in the previous chapter, only young leaves respond to the positive microbiome effect on growth (Figure 5A). Moreover, individual plant leaves showed unique mineral nutrient configurations (Figure 5B) in both low and high nutrient concentrations. This observation reinforces our hypothesis that, when grown in natural soil (as observed in Chapter III) leaf growth is determined by the bacterial population inhabiting the leaves.

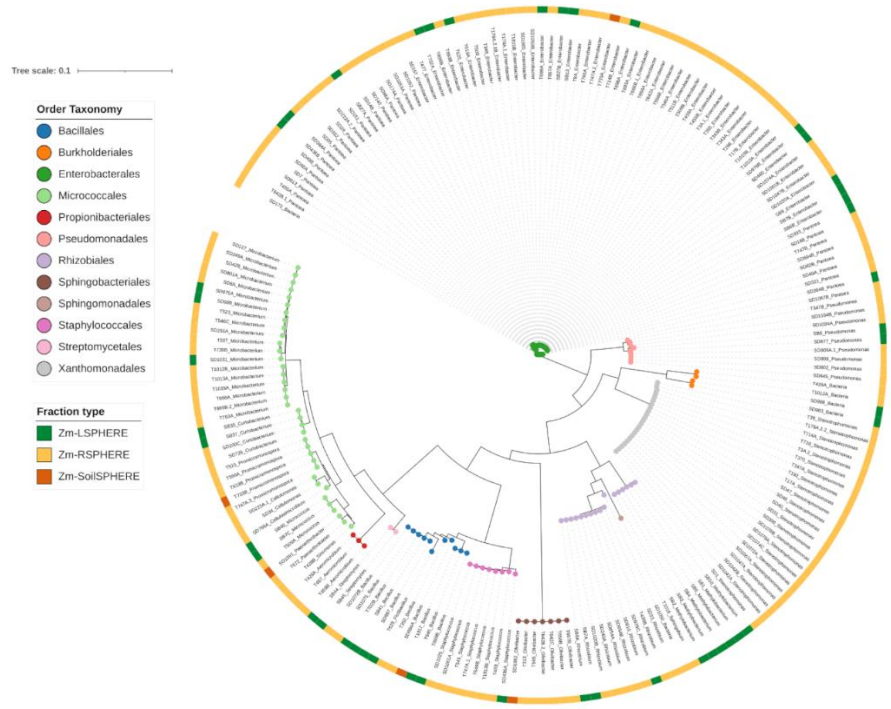


Figure 3: Phylogenetic tree of 211 genome-sequenced isolates used to build a synthetic community (SynCom) resembling the taxonomic profile of natural communities. These bacterial isolates were obtained from surface-sterilized maize roots, leaves and soil. The composition of this SynCom captures the diversity of the natural community bacteria. The tree has tips coloured according to the bacterial taxonomic order. The coloured outer ring represents the leaf (*Zm*-LSPHERE), roots (*Zm*-RSPHERE) and soil (*Zm*-SoilSPHERE) fractions from where the strain was isolated.

To understand the role of the microbiota in the control of leaf growth in maize, we applied a linear model to identify groups of bacterial isolates with high relative abundance and prevalence in the individual leaves. We found specific groups of bacterial strains colonising the different leaves of maize that were classified as leaf markers (Figure 6A). To establish a causal relationship between the bacterial leaf markers and the leaf growth, we compared the lengths of maize leaves exposed to the full bacterial synthetic community and four additional synthetic communities designed to sequentially drop out the individual leaf markers (Figure 6B). We found that the positive effect of the full synthetic community on leaf 3, 4, and 5 growth was abolished when we dropped out individual leaf marker bacteria of leaf 3, 4, 5 from the synthetic community

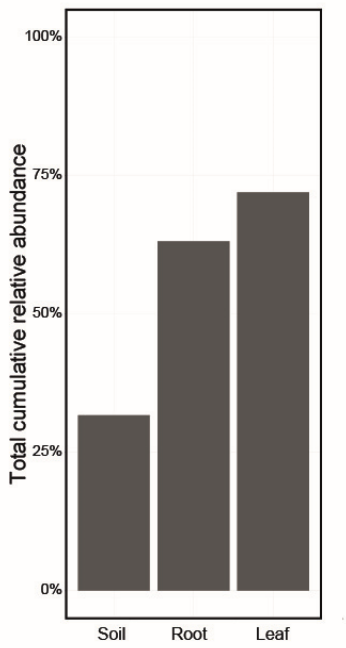


Figure 4 : SynCom captures the majority of order identified in the natural soil experiment. Percentage of bacterial orders found in the bacterial strains collection built in this work and in the full bacterial synthetic community (SynCom) used in our recolonization experiments.

(Figure 6C). Reinforcing these observations, plants exposed to the full synthetic communities had a higher number of bundle sheath cells in leaf 3, 4, and 5 as compared to the same leaf in plants grown under axenic conditions or to plants inoculated with the drop-out synthetic communities for leaf 3, 4, and 5 and (Figure 6D). We did not observe this negative effect on leaf growth when we removed leaf markers for leaf 1 and 2 from the full synthetic community. In this case, the growth of maize plants was similar to that of plants exposed to the full synthetic community (Figure 6C). We verified that the inoculation of the bacterial markers of leaves 3, 4, and 5 alone were not sufficient to induce growth promotion in maize, indicating that leaf bacterial markers are necessary but not sufficient to induce a positive effect on leaf growth, and that an adequate microbial context is needed (Figure 6E). These results indicate that growth promotion in maize leaves is controlled by the population of microbes that efficiently colonize individual young leaves, under the specific nutritional and ecological constraints that each individual leaf represents.

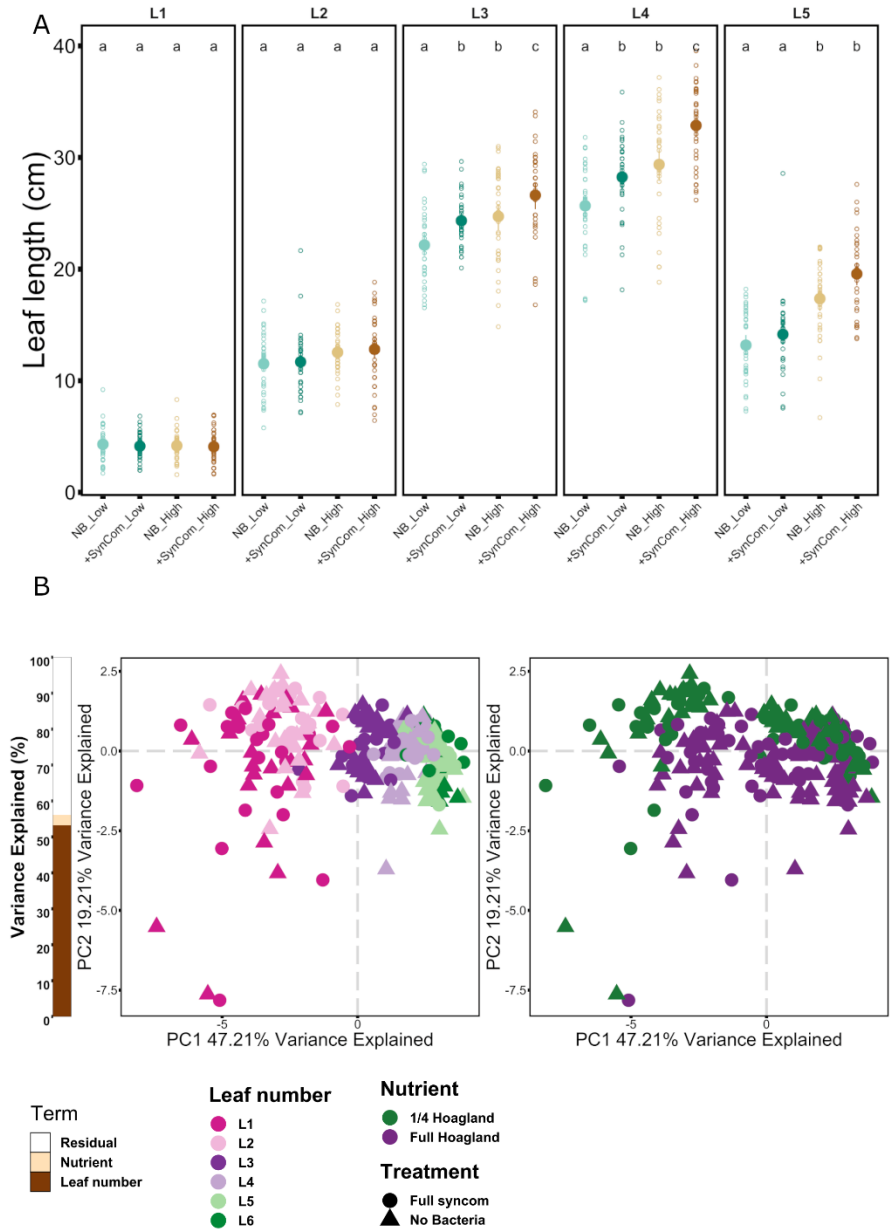


Figure 5: Microbiota is the main drive of leaf growth. A) Individual leaf length (in cm) of four-week-old maize plants grown in different nutrient concentrations. Letters indicate statistical significance corresponding to the Kruskal-Wallis test with Dunn's *post hoc* test (α=0.05). **B)** PCA analysis based on Euclidean distance of the leaf mineral composition (ionome) showing the ionomic profile of maize leaves growth with SynCom (left) with different nutrient concentrations

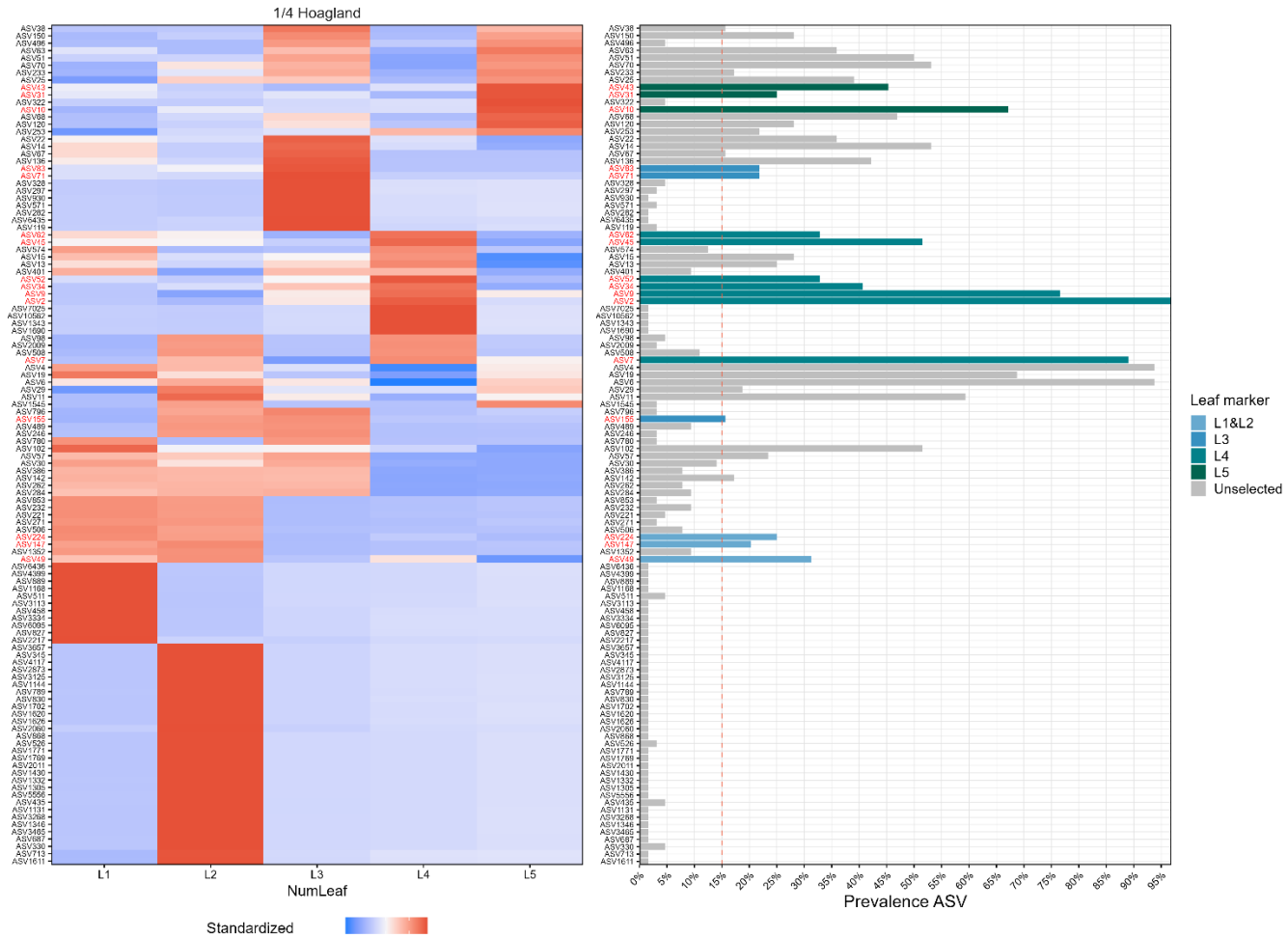
(right). The bar plot to the left of the PCA represents the percentage of variance explained by statistically significant ($p < 0.05$) variables in a PERMANOVA model.

3.3 Repression of defence-related gene cluster coordinates the microbiota effect on leaves

To define the mechanisms underlying the beneficial bacterial effect on leaf growth, we analysed the transcriptional response of leaf 3. To refine our transcriptional analysis, we filtered our gene selections using the transcriptional profiles obtained in wild-type B73 plants inoculated with the three synthetic bacterial communities that sequentially dropped out bacterial leaf markers for leaf 3, leaf 4, and leaf 5 from the full bacterial synthetic community. Plants inoculated with these synthetic communities abolished the microbiota-driven growth-promoting effect on maize leaves. Therefore, we reasoned that the genes showing an opposite transcriptional response in plants inoculated with the full synthetic community compared to those inoculated with the three dropout ones could be related to the microbiota-driven mechanism of leaf growth promotion. With this filtering approach, we identified three clusters of genes (C4, C6, and C8) whose expression in general is repressed in leaf 3 in response to the full synthetic community and induced in response to the three drop-out synthetic communities and in axenic plants (Figure 7A).

The gene ontology analysis of the cluster of these genes showed categories highly enrich in defence-response genes (Figure 7B). Thus, we hypothesized that the microbiota of individual leaves influence leaf growth through: a) inhibition of plant immunity regulators, b) repression of a leaf development-related phytohormone regulatory network, or c) a mechanism involving both defence and development.

A


1/4 Hoagland
L1
L2
L3
L4
L5
NumLeaf
Standardized
Relative Abundance
-1
0
1
Prevalence ASV
0%
5%
10%
15%
20%
25%
30%
35%
40%
45%
50%
55%
60%
65%
70%
75%
80%
85%
90%
95%
Leaf marker
L1&L2
L3
L4
L5
Unselected

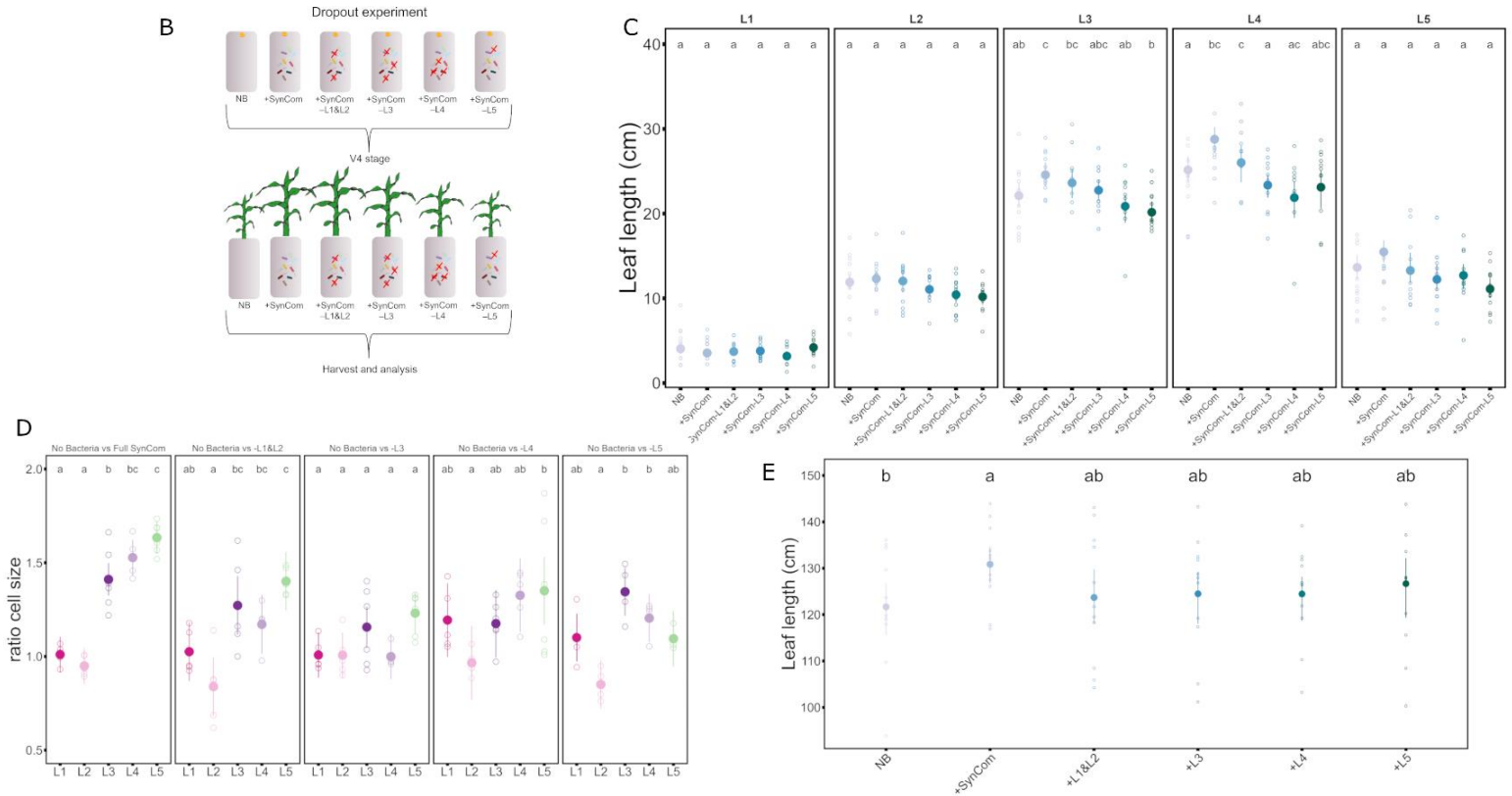
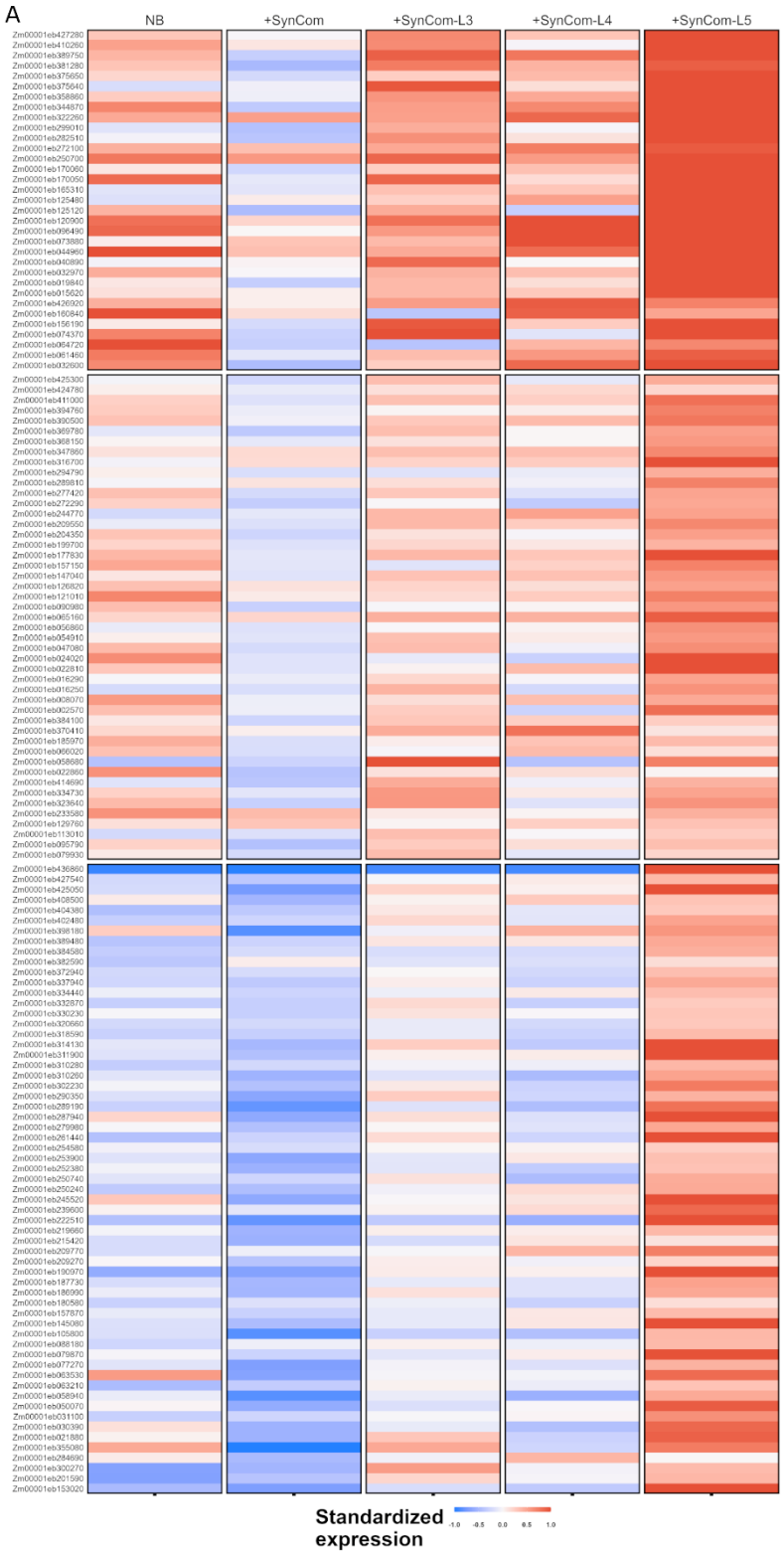

B
Dropout experiment
NB
+SynCom
+SynCom −L1&L2
+SynCom −L3
+SynCom −L4
+SynCom −L5
V4 stage
Harvest and analysis
C
L1
L2
L3
L4
L5
Leaf length (cm)
D
No Bacteria vs Full SynCom
No Bacteria vs -L1&L2
No Bacteria vs -L3
No Bacteria vs -L4
No Bacteria vs -L5
ratio cell size
E
Leaf length (cm)
+L1&L2
+L3
+L4
+L5

Figure 6: Growth promotion in maize leaves is controlled by microbes that efficiently colonize individual leaves. A) Single-leaf enrichment patterns of the synthetic community across different nutrient composition and maize leaves. Each rows represents a unique ASV. B) Schematic description of the dropout experimental approach. C) Individual leaf length (in cm) of four-week-old maize plants growth with full SynCom and drop-out SynCom. Letters indicate statistical significance corresponding to the Kruskal-Wallis with Dunn's *post hoc* test (α=0.05). D) Cell size ratio in maize individual leaves. E) Pointrange plots showing the total length of maize leaf growth with bacterial leave signature. Letters in D) and E) indicate statistical significance corresponding to the ANOVA test with Tukey's *post hoc* test (α=0.05).

To disentangle the defence hypothesis from the developmental one, we first quantified the hormonal profiles of leaf 3 in plants grown under axenic conditions in response to the full synthetic community and the same synthetic community after dropping out leaf 3, 4, and 5 marker bacteria. We did not detect significant differences in the phytohormones auxin, gibberellin, ABA, and cytokinin accumulation among bacterial treatments (Figure 7C). This indicates that the repression of the inhibitory mechanisms that control the induction of the main phytohormones related to leaf development do not explain the differences observed in leaf growth in response to the different bacterial communities used. Therefore, we speculate that the leaf marker bacteria can promote the inhibition of defence-related genes to enhance individual leaf growth.


A
NB
+SynCom
+SynCom-L3
+SynCom-L4
+SynCom-L5
CR4
CR6
CR8
Standardized expression
-1.0
-0.5
0.0
0.5
1.0

B


L3
photosynthesis, light harvesting in photosystem I
photosynthesis, light harvesting
photosynthesis, light reaction
photosynthesis
phylloquinone biosynthetic process
photosynthetic electron transport in photosystem II
fructose 6-phosphate metabolic process
gluconeogenesis
response to light intensity
response to far red light
Golgi to vacuole transport
systemic acquired resistance
regulation of systemic acquired resistance
defense response to bacterium
sphingosine metabolic process
sphingosine biosynthetic process
sphingoid biosynthetic process
negative regulation of ethylene-activated signaling pathway
negative regulation of phosphorelay signal transduction system
actin polymerization or depolymerization
somatic embryogenesis
regeneration
regulation of transcription by RNA polymerase II
glycolytic process
ATP generation from ADP
actin cytoskeleton organization
actin filament-based process
suberin biosynthetic process
calcium ion transmembrane transport
calcium ion transport
peptide catabolic process
CR9 (17)
CR10 (10)
CR7 (25)
CR4 (20)
CR6 (38)
CR8 (37)
CR11 (12)
CR3 (49)
CR2 (52)
CR1 (43)
CR5 (28)
Cluster
GeneRatio
0.1
0.2
0.3
0.4
p.adjust
0.04
0.03
0.02
0.01

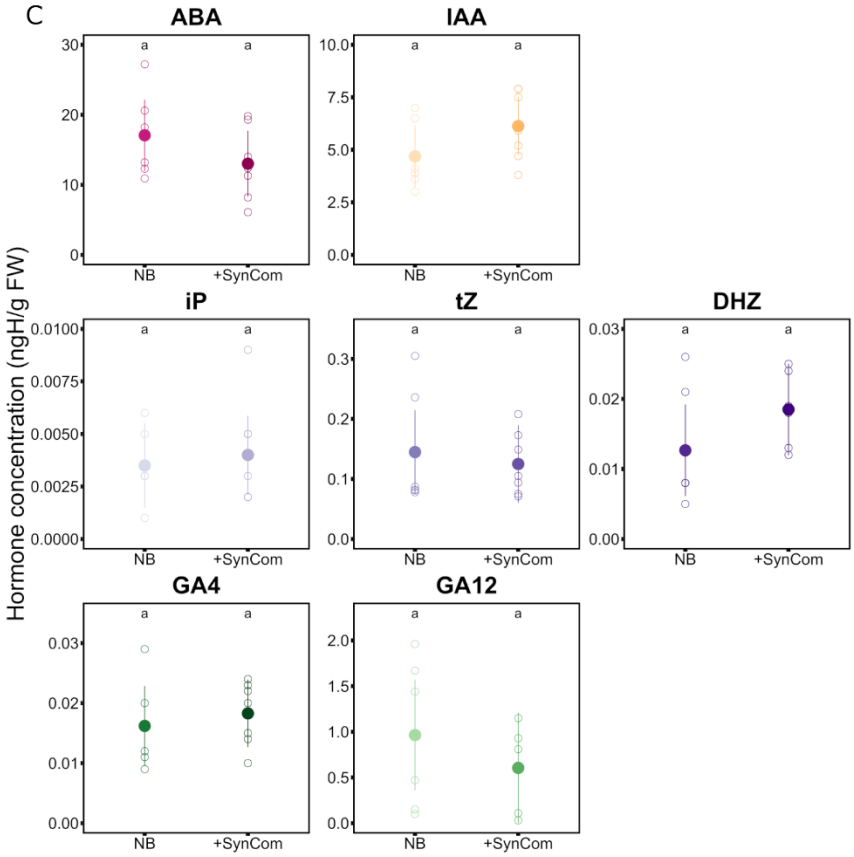


Figure 7: Microbiota-drive plant gene repression of a set of defence-related genes modulating leaf growth in a hormone-independent manner. A) Heatmap showing the average standardized expression of 141 DEGs across the SynCom treatment in leaf 3 (L3) for maize B73 genotype. **B)** GO enrichment for the clusters identified for the DEGs in L3. The gene ratio is the proportion of genes per cluster that belong to a GO category. **C)** Hormone content measured in L3 from plants at the V4 stage, grown in axenic conditions (NB, no bacteria) *versus* the SynCom treatment (+SynCom). Letters represent the statistical significance assessed by the ANOVA test with Tukey's *post hoc* test (α=0.05). DEG, differentially expressed gene; GO, gene ontology; CR, cluster; *iP*, isopentenyladenine; *tZ*, trans-zeatin; *DHZ*, dihydrozeatin.

4 Discussion

We have revealed a molecular circuit that controls the mechanisms that plants use to incorporate the effect of resident microbiota into the leaf growth program. This mechanism requires microbiota-driven repression of a set of defence-related genes that modulate the growth-defence trade-off in a non-hormonal manner to prioritise leaf growth over defence in certain soils. Furthermore, we demonstrate that the amplitude of this response changes with leaf age/maturation stage, preserving the endogenous developmental trajectory of each leaf. Therefore, we have uncovered a leaf microbiota- and age-dependent plant program that could explain the beneficial effect of leaf microbiota on plant growth, with recognized importance for plant adaptation to poor soils.

The inverse relationship between growth and defence is generally the result of the redistribution of available resources in response to external stresses (Engelsdorf et al., 2013; Ullmann-Zeunert et al., 2013). Our work shows that in response to low nutrient availability, the distribution of mineral nutrients and trace elements in the leaves did not influence the growth promoting effect observed in the leaves of maize plants. Furthermore, we demonstrate that highly abundant bacteria inhabiting individual leaves can affect the balance between leaf growth and defence without changing leaf typical ion distribution.

We found that the distinct bacterial populations colonizing individual leaves are responsible for the observed beneficial effect on leaf growth in response to nutrient scarcity. This effect was only verified in young leaves, demonstrating that older leaves (leaf 1 and leaf 2, in our case) no longer respond to the bacterial growth effect, thus remaining blind to environmental perturbations. We found that this bacterial mechanism enhances leaf growth through the repression of a set of genes related to plant immunity. This microbiota-driven response of young maize leaves that makes them prioritise growth over defence was mainly independent from changes in the leaf hormonal profiles that are commonly identified as key regulators of the growth-defence trade-off in plants (Huot *et al.*, 2014; He *et al.*, 2022). Interestingly, we could establish that this microbiota-driven mechanism enhancing leaf growth operates in a single-leaf-specific manner, presumably to preserve the differential ability of

individual leaves to respond to biotic and abiotic stresses. Importantly, the data gathered here provides the molecular explanation of how the plant leaf integrates the local microbiota effect in the endogenous leaf developmental program that Berens et al. suggested in 2019. Further, this mechanism is essentially different from the long-distance systemic control exerted by root microbiota on leaf growth in response to light (Hou *et al.*, 2021).

Our work demonstrates that this leaf-microbiota- and age-driven mechanism balances the growth of the different leaves while respecting their endogenous differential activation of growth-defence trade-off. In a multi-stress context such as in natural ecosystems, we predict that this mechanism intersects with other branches of the leaf growth regulatory network to establish a hierarchy of biotic or abiotic stress responses to ensure plant survival. These findings demonstrate that the healthy microbiota inhabiting the leaf is also a key component of the regulatory network that controls the balance between growth and defence in the plant. Therefore, it might now be possible to manipulate endogenous growth and defence trade-off mechanisms in crops such as maize through optimization of the leaf microbiota sprayed or used as biofertilizer without compromising the plant's response to environmental cues including pathogens. It should allow us to develop crops with an optimised microbiota that can grow and adapt better to the increasingly common challenging conditions without additional penalties for the plants.

5 References

Alonge, M., Lebeigle, L., Kirsche, M., et al. (2022) Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol*, **23**, 1–19.

Bai, Y., Müller, D.B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., Dombrowski, N., Munch, P.C., Spaepen, S., Remus-Emsermann, M., Huttel, B., McHardy, A.C., Vorholt, J.A., Schulze-Lefert, P. (2015) Functional overlap of the Arabidopsis leaf and root microbiota. *Nature*, **528**, 364–369.

Bankevich, A., Nurk, S., Antipov, D., et al. (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, **19**.

Berens, M.L., Wolinska, K.W., Spaepen, S., Ziegler J, Nobori T, Nair A, Krüler V, Winkelmüller TM, Wang Y, Mine A, Becker D, Garrido-Oter R, Schulze-Lefert P, Tsuda K. (2019) Balancing trade-offs between biotic and abiotic stress responses through leaf age-dependent variation in stress hormone cross-talk. *Proc Natl Acad Sci U S A*, **116**, 2364–2373.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A. and Holmes, S.P. (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*.

Chaumeil, P.A., Mussig, A.J., Hugenholtz, P. and Parks, D.H. (2020) GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*, **36**, 1925–1927.

Chen, S., Zhou, Y., Chen, Y. and Gu, J. (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, **34**, i884–i890.

Custódio, V., Gonin, M., Stabl, G., Bakhoum, N., Oliveira, M.M., Gutjahr, C. and Castrillo, G. (2022) Sculpting the soil microbiota. *Plant J*, **109**, 508–522.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**, 2460–2461.

Eichmann, R., Richards, L. and Schäfer, P. (2021) Hormones as go-betweens in plant microbiome assembly. *The Plant Journal*, **105**, 518.

Engelsdorf, T., Horst, R.J., Pröls, R., Pröschel, M., Dietz, F., Hückelhoven, R. and Voll, L.M. (2013) Reduced carbohydrate availability enhances the susceptibility of Arabidopsis toward Colletotrichum higginsianum. *Plant Physiol*, **162**, 225–238.

Finkel, O.M., Salas-González, I., Castrillo, G., Conway JM, Law TF, Teixeira PJPL, Wilson ED, Fitzpatrick CR, Jones CD, Dangl JL. (2020) A single bacterial genus maintains root growth in a complex microbiome. *Nature 2020 587:7832*, **587**, 103–108.

Hacquard, S., Garrido-Oter, R., González, A., Spaepen S, Ackermann G, Lebeis S, McHardy AC, Dangl JL, Knight R, Ley R, Schulze-Lefert P. (2015) Microbiota and Host Nutrition across Plant and Animal Kingdoms. *Cell Host Microbe*, **17**, 603–616.

He, Z., Webster, S. and He, S.Y. (2022) Growth-defense trade-offs in plants. *Curr Biol*, **32**, R634–R639.

Hou, S., Thiergart, T., Vannier, N., Mesny, F., Ziegler, J., Pickel, B. and Hacquard, S. (2021) A microbiota-root-shoot circuit favours Arabidopsis growth over defence under suboptimal light. *Nat Plants*, **7**, 1078–1092.

Huot, B., Yao, J., Montgomery, B.L. and He, S.Y. (2014) Growth-defense tradeoffs in plants: A balancing act to optimize fitness. *Mol Plant*, **7**, 1267–1287.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T. and Aluru, S. (2018) High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications 2018 9:1*, **9**, 1–8.

Letunic, I. and Bork, P. (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res*, **49**, W293–W296.

Logemann, J., Schell, J. and Willmitzer, L. (1987) Improved method for the isolation of RNA from plant tissues. *Anal Biochem*, **163**, 16–20.

Ma, K.W., Niu, Y., Jia, Y., Ordon J, Copeland C, Emonet A, Geldner N, Guan R, Stolze SC, Nakagami H, Garrido-Oter R, Schulze-Lefert P. (2021) Coordination of microbe–host homeostasis by crosstalk with plant innate immunity. *Nature Plants 2021 7:6*, **7**, 814–825.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*.

Massoni, J., Bortfeld-Miller, M., Widmer, A. and Vorholt, J.A. (2021) Capacity of soil bacteria to reach the phyllosphere and convergence of floral communities despite soil microbiota variation. *Proc Natl Acad Sci U S A*, **118**, e2100150118.

Murillo-Roos, M., Abdullah, H.S.M., Debbar, M., Ueberschaar, N. and Agler, M.T. (2022) Cross-feeding niches among commensal leaf bacteria are shaped by the interaction of strain-level diversity and resource availability. *The ISME Journal 2022 16:9*, **16**, 2280–2289.

Nakayama, H., Leichty, A.R. and Sinha, N.R. (2022) Molecular mechanisms underlying leaf development, morphological diversification, and beyond. *Plant Cell*, **34**, 2534–2548.

Oksanen, J. (2007) vegan : Community Ecology Package. R package version 1.8-5.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. and Kingsford, C. (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*, **14**, 417–419.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res*, **41**.

Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*, **9**, 671–675.

Singkaravanit-Ogawa, S., Kosaka, A., Kitakura, S., Uchida, K., Nishiuchi, T., Ono, E., Fukunaga, S. and Takano, Y. (2021) Arabidopsis CURLY LEAF functions in leaf immunity against fungal pathogens by concomitantly repressing SEPALLATA3 and activating ORA59. *Plant J*, **108**, 1005–1019.

Soneson, C., Love, M.I. and Robinson, M.D. (2015) Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res*, **4**.

Teixeira, P.J.P., Colaianni, N.R., Fitzpatrick, C.R. and Dangl, J.L. (2019) Beyond pathogens: microbiota interactions with the plant immune system. *Curr Opin Microbiol*, **49**, 7–17.

Ullmann-Zeunert, L., Stanton, M.A., Wielsch, N., Bartram, S., Hummert, C., Svatoš, A., Baldwin, I.T. and Groten, K. (2013) Quantification of growth-defense trade-offs in a common currency: nitrogen required for phenolamide biosynthesis is not derived from ribulose-1,5-bisphosphate carboxylase/oxygenase turnover. *Plant J*, **75**, 417–429.

Vogel, C.M., Potthoff, D.B., Schäfer, M., Barandun, N. and Vorholt, J.A. (2021) Protective role of the Arabidopsis leaf microbiota against a bacterial pathogen. *Nat Microbiol*, **6**, 1537–1548.

Wagner, M.R., Lundberg, D.S., Rio, T.G. Del, Tringe, S.G., Dangl, J.L. and Mitchell-Olds, T. (2016) Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat Commun*, **7**.

Wallace, J.G., Kremling, K.A., Kovar, L.L. and Buckler, E.S. (2018) Quantitative genetics of the maize leaf microbiome. *Phytobiomes J*, **2**, 208–224.

Walter, A. and Schurr, U. (2005) Dynamics of leaf and root growth: endogenous control versus environmental impact. *Ann Bot*, **95**, 891–900.

Yu, G., Wang, L.-G., Han, Y. and He, Q.-Y. (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*, **16**, 284–287.

Zhang, Jingying, Liu, Y.X., Guo, X., Qin, Y., Garrido-Oter, R., Schulze-Lefert, P. and Bai, Y. (2021) High-throughput cultivation and identification of bacteria from the plant root microbiota. *Nature Protocols 2021 16:2*, **16**, 988–1012.

Zhang, Janie, Cook, J., Nearing, J.T., Zhang, Junzeng, Raudonis, R., Glick, B.R., Langille, M.G.I. and Cheng, Z. (2021) Harnessing the plant microbiome to promote the growth of agricultural crops. *Microbiol Res*, **245**, 126690.

Chapter V

Final conclusions and Future perspectives

Chapter V – Final conclusions and Future perspectives

The accumulation of greenhouse gases in the atmosphere is growing, temperatures are rising, precipitation patterns are changing, and extreme weather events are happening more frequently due to the effects of climate change (Stott, 2016). These effects accounted for 25% of average yield losses for low-latitude maize (Rosenzweig *et al.*, 2014). Several strategies have been developed to tackle the reduction of productivity, such as altering planting and harvesting time, sowing crops with a short life cycle, applying crop rotation and irrigation techniques, and introducing variation in cropping systems (Duku *et al.*, 2018; Marcinkowski and Piniewski, 2018; Teixeira *et al.*, 2018). Besides these strategies, a knowledge-driven approach to managing the soil microbiota rather than merely applying microorganisms was described as promising to sustainably increase crop yield (French *et al.*, 2021).

The main aim of this thesis was to elucidate microbiota functions that alter maize development under nutrient-deficiency conditions. For that purpose, a holistic approach was employed that combines the characterization of a collection of agricultural soils, including soil bacterial composition, plant microbiota, and nutrition availability. For two years, we monitored free-living soil bacteria communities, soil mineral content, and climatic conditions in 24 agricultural soils in the Sahel region (Cabo Verde) (Chapter II). Since plants recruit their symbionts from the surrounding soil free-living microbial communities, we have further evaluated the factors driving microbiota assembly in root and maize single leaves testing three different soils, two from Cabo Verde different regions (São Domingos and Tarrafal) and one from UK (Sutton Bonington) (Chapter III). Then we isolated a collection of bacterial strains and used it to design synthetic communities (SynCom) to evaluate the effects of the microbiota composition on the growth of single leaves (Chapter IV).

Our first approach (Chapter II) was to characterize the climate and soil properties factors that affect maize productivity in Santiago island. By using a survey applied to maize producers in Santiago island, we confirmed that maize production in this location follows traditional practices that rely on natural resources (soil nutrient composition and precipitation) rather than on chemical fertilization and irrigation. Indeed, when analysing the data on climate and soil properties of the regions of the different local producers, we found that soil mineral content was a good integrator of water-related (sand content, precipitation, solar radiation, water vapour pressure, and aridity) and nutrient-related (cation exchange, soil organic carbon, and nitrogen concentration) factors affecting maize production. These findings were consistent with a previous study, that showed that continuous increase in aridity in global ecosystems abruptly reduces soil nutrient composition (Berdugo *et al.*, 2020). We further selected a group of farms, representative of the variation in climate and soil features identified in the survey, to collect soil samples for soil mineral content and microbial analysis. The results from these studies demonstrated that soil mineral content and soil microbiota differs among regions. To better understand whether these variations affect maize productivity and to validate our findings, we selected two field sites that are geographically distant but have similar climate characteristics. We imposed short-term changes in soil mineral content with an expected effect on the soil microbiota through two agricultural practices commonly used to increase maize productivity: fertilization and cropping system. We quantified the height of maize plants as a proxy for plant productivity. We found that the field site locations and the agricultural practices explained small variations observed in maize height, which allowed us to reason that soil mineral content, soil microbiota or interactions between both may affect maize productivity in the field.

In Chapter III, the effect on maize growth and development, of soil mineral composition and bacterial community structure was further explored. Our aim was to identify the main drivers controlling maize leaves growth under controlled conditions. To demonstrate the soil effect on plant growth, we start by selecting two arable soils from Santiago island (Cabo Verde) with an irregular and scarce rainy season and with only few heavy rains (Baptista *et al.*, 2015) and a rich soil (free of pesticides and fertilizers) with a history of high precipitation exposure, from Sutton Bonington (Nottingham, UK). We examined differences in soil mineral content in the different soils using Inductively Coupled Plasma Mass Spectrometry (ICP-MS). This analysis showed that the soils from Santiago island have similar mineral composition but very distinct compared to the Sutton Bonington soil. To identify the main drivers of maize leaf growth under controlled conditions, we grew four maize genotypes (a model variety, B73, and three other, Matuba, CVSt and Zm523) regularly used in maize production in Africa in the three soils selected. Overall, in Chapter III, we assessed the effects on maize growth of natural nutrient availability and microbiome composition. As a proxy for plant growth, we measured total shoot length and we found that maize plants grown in São Domingos soil showed no significant differences in shoot length compared to those grown in Sutton Bonington soil, despite having a clear contrasting amount of mineral nutrients. As expected, plants grown in the Tarrafal soil were significantly smaller than in the other soils tested. In general, we found no significant differences in root dry weight among soils, allowing us to rule out the root-associated functions related to soil exploration as a causal factor for the observed differences in plant shoot growth. One of the most important observations in Chapter III, was that the differences found in the total leaves length across soils did not translate to the level of individual leaves. In all cases, old leaves (L1 and L2) showed no differences in length across the soils. On the other hand, young leaves

(L3, L4 and L5) showed significant differences in their length that recapitulated those observed in the total leaf length. Our results suggest that the leaf development program integrates soil properties and other biotic and abiotic environmental factors.

Still, in Chapter III, we investigated whether the accumulation of soil elements in leaves can explain the plant growth variability found in the different soils. For that, we determined the concentration of 17 mineral elements in the leaves of maize plants. We found that the individual leaves have distinct ionomic profiles that were conserved in the different soils. Furthermore, we observed that in all soils, nutrients such as P, K, Mg and Ca significantly correlated positively and negatively with leaves age, suggesting that, in general, the endogenous nutrient allocation in leaves is robust to changes in soil mineral nutrient composition. Nevertheless, we found that plants grown in the two poor soils from Santiago Island exhibited similar single-leaf ionomes, despite having significant differences in leaf growth. As expected, the leaves of plants grown in Sutton Bonington soil showed the most distinct ionomic profiles. Although their leaf growth was similar to that of plants grown in São Domingos soil. One of the main conclusions of Chapter III is that the accumulation of mineral nutrients in leaves of plants grown under ecologically relevant conditions, with the microbiota present, may not be enough to explain the differences found in leaf growth across soils.

In Chapter III, we also examined whether plant microbiota properties contribute to the differences observed in plant growth across soils. By profiling bacterial communities established in plant roots, leaves and soil using DNA sequencing, we found that significant differences in the root bacterial community composition are explained by the geographical localization of the soils. These differences were less obvious in the case of leaf microbiota, in which the differences in composition were also explained by leaf age. As observed for the leaf ionome, we found that

bacterial communities are distinct in different individual leaves and that this arrangement is conserved in all soils, indicating that each leaf represents a unique environment. These results suggest that leaf-specific bacteria groups could explain differences in leaf growth. Then, we hypothesized that bacterial traits necessary to colonize the different environments in individual leaves could be more frequently found in highly abundant bacteria. To test the hypothesis, we first identified highly abundant bacterial amplicon sequence variants (ASVs) per individual leaf in the three soils. We observed that the number of highly abundant bacterial ASVs shared among leaves of similar length, was in general higher than that found when comparing plants with contrasting leaf sizes growth in Tarrafal and Sutton Bonington soils. This suggests that small groups of highly abundant bacteria may influence the growth of maize leaves. This hypothesis was then tested in Chapter IV.

In Chapter IV, to demonstrate how highly abundant bacteria inhabiting the individual leaves can modulate the leaf growth in maize, we first generated a bacterial culture collection, that we later used to build a synthetic community with whole-genome sequenced members. Thereafter, we designed a clay-sand-based system to reduce the soil system complexity and determine leaf growth in axenic conditions (in the absence of microbiota). We tested different proportions of clay and sand to find the right proportion that reproduces the water retention capacity found in Sutton Bonington (the soil with the highest SWC among the three tested ones). By using this system, under axenic conditions, we determined leaf length in four different nutrient concentrations (no nutrient, ¼ Hoagland, ½ Hoagland, and Full Hoagland). The maize wild-type B73 plant leaves grew following a linear trajectory pattern, driven by the concentration of the mineral nutrients present in the mixture of clay and sand. This effect was also observed at the single-leaf resolution. The leaves of plants exposed to an

increasing gradient of nutrient concentrations followed a linear growth pattern. This indicates that under axenic conditions, increasing non-toxic concentrations of some mineral nutrients in the soil positively influences plant leaf growth. Based on this observation, we hypothesize that the leaf growth promotion we observed in plants grown in São Domingos soil (Chapter III), may be determined by the population of microbes inhabiting the leaves. To test this hypothesis, we inoculated bacterial synthetic communities in pots with B73 maize seeds where the clay-sand mixture was supplemented with two concentrations of Hoagland solution, low and full, to approximately simulate the low nutrient concentrations of the two Santiago island regions, and the nutrient-rich soil of UK. We observed a growth promotion effect in the SynCom-inoculated plants as compared to those not inoculated. Consistent with our previous experiment using natural soils (Chapter III), we found that young leaves respond positively to the microbiome inoculation and that, although such a positive effect occurred in both nutrient concentrations, it was more obvious in plants grown in low-nutrient soil.

Furthermore, to understand microbiota's role in leaf growth modulation in maize, we identified groups of bacterial isolates showing high abundance and prevalence in the colonization of specific maize leaves. These bacteria could therefore be classified as leaf markers. To establish a causal relationship between bacterial leaf markers and leaf growth, we still compared the length of maize leaves in plants inoculated with the full synthetic community as compared with those inoculated with synthetic communities (four) designed to sequentially drop out the individual leaf markers. We found that the positive effect of the synthetic community in L3, L4, and L5 growth was abolished when individual leaf markers from the L3, L4 and L5 markers were dropped out. We also tested the inoculation of single bacterial markers, but found that these alone were insufficient to induce growth promotion in maize leaves and

that an adequate microbial context is imperative. Finally, to define the mechanism underlying the beneficial bacterial effect on leaf growth, we analysed the transcriptional response of the single leaf to the SynCom. To identify groups of genes specifically involved in the leaf-growth promotion mechanism, we filtered the full SynCom-responsive genes by comparing their expression with that found in plants inoculated with the four drop-out synthetic communities (that lost the growth promotion effect although keeping similar bacterial background). Using this approach, we could identify a cluster of genes with repressed expression in leaf 3, as a response to the full SynCom. The gene ontology analysis of this cluster revealed categories related to leaf development and plant defence, such as regulation of cytoskeleton organization and ethylene and salicylic acid response, suggesting that individual leaf microbiomes influence the growth of the leaves through specific modulation of the growth-defence trade-off mechanisms.

We found that the distinct bacterial populations colonizing individual leaves justify the leaf growth beneficial effect observed in response to nutrient scarcity. However, further studies are still needed to understand the individual mechanisms operating in each leaf. One possible experiment would be to obtain homozygous mutants for the genes from the cluster with the repressed genes in response to the bacterial synthetic community, and analyse if their reduced activity enhances maize leaf growth promotion in response to the bacterial SynCom. It is also necessary to evaluate if the synthetic community exerts differential repression of the cluster genes in a leaf age-dependent manner, which could explain growth promotion phenotypes of plants grown in a variety of soils.

Finally, this PhD thesis contributed to disclose further biotic and abiotic cues regulating maize development, generating knowledge that will help to design synthetic communities with a predictable effect on the

plant host. In this way, we hope to have contributed to supporting the development of more efficient and environmentally-friendly agricultural practices, up to the challenges that humanity is currently facing.

1 References

Baptista, I., Ritsema, C. and Geissen, V. (2015) Effect of Integrated Water-Nutrient Management Strategies on Soil Erosion Mediated Nutrient Loss and Crop Productivity in Cabo Verde Drylands P. C. Abhilash, ed. *PLoS One*, **10**, e0134244.

Berdugo, M., Delgado-Baquerizo, M., Soliveres, S., Hernández-Clemente R., Zhao Y., Gaitán J.J., Gross N., Saiz H., Maire V., Lehmann A., Rillig M.C., Solé R.V., Maestre F.T. (2020) Global ecosystem thresholds driven by aridity. *Science (1979)*, **367**, 787–790.

Duku, C., Zwart, S.J. and Hein, L. (2018) Impacts of climate change on cropping patterns in a tropical, sub-humid watershed. *PLoS One*, **13**, e0192642.

French, E., Kaplan, I., Iyer-Pascuzzi, A., Nakatsu, C.H. and Enders, L. (2021) Emerging strategies for precision microbiome management in diverse agroecosystems. *Nat Plants*, **7**, 256–267.

Marcinkowski, P. and Piniewski, M. (2018) Effect of climate change on sowing and harvest dates of spring barley and maize in Poland**. *Int. Agrophys*, **32**, 265–271.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane AC, Müller C, Arneth A, Boote KJ, Folberth C, Glotter M, Khabarov N, Neumann K, Piontek F, Pugh TA, Schmid E, Stehfest E, Yang H, Jones JW. (2014) Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proc Natl Acad Sci U S A*, **111**, 3268–3273.

Stott, P. (2016) How climate change affects extreme weather events: Research can increasingly determine the contribution of climate change to extreme events such as droughts. *Science (1979)*, **352**, 1517–1518.

Teixeira, E.I., Ruiter, J. de, Ausseil, A.G., Daigneault, A., Johnstone, P., Holmes, A., Tait, A. and Ewert, F. (2018) Adapting crop rotations to climate change in regional impact modelling assessments. *Sci Total Environ*, **616–617**, 785–795.

ITQB NOVA, Avenida da República, 2780-157 Oeiras, Portugal | Tel +351 214 469 100 | itqb.unl.pt