
# INSTITUTO POTOSINO DE INVESTIGACIÓN CIENTÍFICA Y TECNOLÓGICA, A.C.

## POSGRADO EN CIENCIAS AMBIENTALES

# Effect on the variation of root exudates from *Agave lechuguilla* Torr. over the diversity and activity of nutrient-improvement rhizobacteria

Tesis que presenta

**Guadalupe Medina de la Rosa**

Para obtener el grado de

**Doctora en Ciencias Ambientales**

Directora de la Tesis:

**Dra. Nguyen Esmeralda López Lozano**

Miembros del comité tutorial:

**Dr. Ángel Gabriel Alpuche Solís**

**Dra. María de Lourdes Berenice Celis García**

**Dr. Felipe García Oliva**

**Dra. Lucía López Reyes**

**San Luís Potosí, S.L.P., marzo de 2022**

## Constancia de aprobación de tesis

**CARTA DE AUTORIZACIÓN DEL COMITÉ TUTORAL**

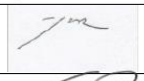
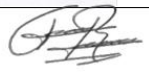
División de Ciencias Ambientales
Doctorado en Ciencias Ambientales

San Luis Potosí, S.L.P., a 16 de febrero de 2022

**Dr. Leonardo Chapa Vargas**
Coordinador Académico del Programa de
Doctorado en Ciencias Ambientales
PRESENTE

Por medio de la presente, le informamos que después de haber revisado el trabajo de Tesis de Doctorado del (la) estudiante Guadalupe Medina de la Rosa, no encontramos inconveniente alguno para que pueda llevarse a cabo el Examen de Grado ante un jurado.

Atentamente,

| COMITÉ TUTORAL | |
|---|---|
| **Nombre de Investigador** | **Firma de Aceptación** |
| Dr. Ángel Gabriel Alpuche Solís |  |
| Dra. María de Lourdes Berenice Celis García |  |
| Dra. Nguyen Esmeralda López Lozano |  |
| Dr. Felipe García Oliva |  |
| Dra. Lucía López Reyes |  |

*Este formato será entregado por el alumno en el Departamento del Posgrado.


## Créditos Institucionales

Esta tesis fue elaborada en los Laboratorios de Ecología Microbiana y Ecología de Suelos de la División de Ciencias Ambientales del Instituto Potosino de Investigación Científica y Tecnológica, A.C., bajo la dirección de la Dra. Nguyen Esmeralda López Lozano.

Los resultados de esta investigación son parte del proyecto de Ciencia Básica SEP-CONACyT 254406 cuyo responsable técnico fue la Dra. Nguyen Esmeralda López Lozano.

Durante la realización del trabajo la autora recibió una beca académica del Consejo Nacional de Ciencia y Tecnología (332648) y del Instituto Potosino de Investigación Científica y Tecnológica, A. C.

El análisis de datos fue apoyado por los recursos computacionales del Centro Nacional de Supercómputo del Instituto Potosino de Investigación Científica y Tecnológica, mediante la beca TKII-NELL001.

La identificación de las bacterias a través de la secuenciación de los genes *nifH* y *phoD*, fue por medio de la colaboración con el Dr. Cesaré Moisés Ovando Vázquez, del Centro Nacional de Supercómputo-IPICyT.

Una parte de los análisis de suelo (Capítulo 3) fueron realizados en el Instituto de Investigaciones en Ecosistemas y Sustentabilidad de la Universidad Nacional Autónoma de México, por el grupo de investigación del Dr. Felipe García Oliva.

## Institutional support

The experimental procedures were performed in the laboratories of the Environmental Sciences' Division from the Instituto Potosino de Investigación Científica y Tecnológica, A.C. and directed by Dr. Nguyen Esmeralda López Lozano.

Data analysis was supported by the computational resources of the National Supercomputing Center at IPICYT through the grant TKII-NELL001.

Identification of bacteria trough *nifH* and *phoD* sequencing was through collaboration with Dr. Cesaré Moisés Ovando Vázquez, from the National Supercomputing Center at IPICyT.

Some of the soil analyses in chapter three were performed at the Instituto de Investigaciones en Ecosistemas y Sustentabilidad from the Universidad Nacional Autónoma de México by Dr. Felipe García Oliva.

This project was supported by the funding of SEP-CONACyT Basic Science 254406 to Dr. Nguyen Esmeralda López Lozano.

During the realization of this work, the author received scholarship grants form CONACyT (332648) and Instituto Potosino de Investigación Científica y Tecnológica A. C.

IPICYT

# Instituto Potosino de Investigación Científica y Tecnológica, A.C.

## Acta de Examen de Grado

El Secretario Académico del Instituto Potosino de Investigación Científica y Tecnológica, A.C., certifica que en el Acta 030 del Libro Primero de Actas de Exámenes de Grado del Programa de Doctorado en Ciencias Ambientales está asentado lo siguiente:

En la ciudad de San Luis Potosí a los 18 días del mes de marzo del año 2022, se reunió a las 10:00 horas en las instalaciones del Instituto Potosino de Investigación Científica y Tecnológica, A.C., el Jurado integrado por:

| | | |
|---|---|---|
| **Dra. María de Lourdes Berenice Celis García** | **Presidenta** | **IPICYT** |
| **Dr. Ángel Gabriel Alpuche Solís** | **Secretario** | **IPICYT** |
| **Dr. Felipe García Oliva** | **Sinodal externo** | **UNAM** |
| **Dra. Lucia López Reyes** | **Sinodal externo** | **BUAP** |
| **Dra. Nguyen Esmeralda López Lozano** | **Sinodal** | **IPICYT** |

a fin de efectuar el examen, que para obtener el Grado de:

DOCTORA EN CIENCIAS AMBIENTALES

sustentó la C.

**Guadalupe Medina de la Rosa**

sobre la Tesis intitulada:

***Effect on the variation of root exudates from Agave lechuguilla Torr. over the diversity and activity of nutrient-improvement rhizobacteria***

que se desarrolló bajo la dirección de

**Dra. Nguyen Esmeralda López Lozano**

El Jurado, después de deliberar, determinó

APROBARLA

Dándose por terminado el acto a las 12:05 horas, procediendo a la firma del Acta los integrantes del Jurado. Dando fe el Secretario Académico del Instituto.

A petición de la interesada y para los fines que a la misma convengan, se extiende el presente documento en la ciudad de San Luis Potosí, S.L.P., México, a los 18 días del mes de marzo de 2022.

**Dr. Marcial Bonilla Marín**
Secretario Académico

**Mtra. Ivonne Lizette Cuevas Vélez**
Jefa del Departamento del Posgrado

IPICYT
SECRETARÍA ACADÉMICA
INSTITUTO POTOSINO DE
INVESTIGACIÓN CIENTÍFICA
Y TECNOLÓGICA, A.C.

## Dedicatorias

A mi familia, la mayor inspiración en mi vida. Porque siempre han estado conmigo en todo momento. Porque estando juntos, no hay adversidad que no podamos superar. Los quiero mucho.

## Agradecimientos

A la Dra. Esmeralda por haberme dado la oportunidad de trabajar con ella. Por sus valiosos consejos y su apoyo en los momentos difíciles. Porque con su ayuda, este trabajo no sólo fue enriquecedor, sino también divertido. Muchas gracias, Doctora.

A los miembros de mi comité tutorial Dr. Ángel Alpuche, Dra. Berenice Celis, Dr. Felipe y la Dra. Lucy, porque cada uno aportó invaluables recomendaciones a este proyecto. Muchas gracias por regalarme un poco de su tiempo y siempre tener una gran disposición para ayudarme a resolver mis dudas.

Al Dr. Cesaré Ovando por su gran paciencia para ayudarme con mi análisis de datos. A los Doctores que me impartieron clases, muchas gracias a todos por compartir sus conocimientos conmigo.

A los técnicos que me ayudaron y orientaron a lo largo de este proyecto: El Mto. Juan Pablo Rodas, I. Q. Ma. Del Carmen Rocha, Mto. Guillermo Vidriales y por supuesto, el personal del LANBAMA. A todos ellos muchas gracias por su inagotable paciencia y apoyo con el análisis de mis muestras.

A mi grupo de trabajo, Ernesto, Ángel, Rosaura y María, porque me recibieron con los brazos abiertos y me ayudaron no sólo con el trabajo sino también a sentirme en confianza y disfrutar esta etapa. Mis compañeros de generación, Jacky, Kristell, Mine, Maximino y Francisco. A todos ustedes muchas gracias por su compañía, compartir sus conocimientos y brindarme su amistad.

A mis amigas, Bety, Karina, Laura y Liliana, por su amistad y apoyo.

A mis padres, por su cariño y por siempre creer en mí. Siempre han sido mi ejemplo a seguir. Mis hermanos y mis sobrinos, porque sé que cuento con ustedes en todo momento.

A ti abuelito, porque tus bendiciones siempre me daban fuerza para seguir. Por gracia de Dios ya descansas con mi abuelita, pero sé que siempre verás por nosotros.

# Index

## List of tables

## List of figures

## Abstract

The wide distribution of *Agave lechuguilla* Torrey has drawn attention to the mechanisms that enable its survival in the Chihuahuan Desert. This success could be attributed to the association of lechuguilla with plant-growth promoting rhizobacteria (PGPR) that, through diverse mechanisms, increase nutrient availability, improve the tolerance of plants to stressful conditions, aid in the control of pathogen microorganisms, and produce phytohormones, overall favoring plant growth. This work encompasses several factors that may influence the complex interaction plant-soil-bacteria and the possible ways that direct the recruitment of the rhizobacterial community of lechuguilla.

In the first chapter, we explored the effect of contrasting soil properties from the Central and Meseta subregions from the Chihuahuan Desert, over the bacterial diversity associated to the rhizosphere of young and mature plants of *A. lechuguilla* T. It was found that the soil properties related to nutrient availability were associated with the composition, enzymatic activity, and abundance of the bacterial communities, however as we consider later, these properties may also be the consequence of the presence of the plants and the bacterial activity.

For the second chapter its was addressed the function of the present amino acids in the exudates from the roots of lechuguilla over the abundance and enzymatic activity of PGPR. In addition, the variations in the functional diversity of the rhizobacterial communities from young and mature lechuguilla plants were explored. It was found that mature lechuguillas, above young plants or bulk soil, stimulated functions related to biological control of pathogens, phytohormones production, tolerance against abiotic stress, and improvement of the nutrient availability especially by the action of nitrogen-fixing bacteria (NFB) and organic phosphorus mineralizing bacteria (OPMB). Furthermore, the enhanced enzymatic activity and abundance of NFB and OPMB reached the highest values in the rhizosphere of mature lechuguillas. Probably, due to the release of arginine and methionine, which are energy reserves and contribute with essential elements for the biosynthesis of the proteins, whether they are incorporated directly or after mineralization.

Finally, in the third chapter, the variations in the abundance and enzymatic activity of PGPR in soil samples as response to different amino acids, organic acids, and sugars identified in the rhizosphere of lechuguilla were studied. It was found that lechuguilla root exudates seemed to boost the priming effect of NFB and OPMB communities, that may lead to the accumulation of nutrients such as total organic carbon, ammonium, nitrite, and nitrate. This increase could explain, the well-off state of the rhizosphere and the tolerance of lechuguilla to the harsh conditions of the Chihuahuan Desert.

This work suggests that the root exudates in *A. lechuguilla* T. can change according to the plant stage and, consequently, the recruitment of beneficial bacteria and their enzymatic activity. However, further research is needed, to understand the complex processes in the interaction plant-soil-bacteria under arid conditions.

**Keywords:** *Agave lechuguilla* Torrey, rhizosphere, bacterial communities, plant-soil-microorganism interactions, co-occurrence networks, root exudates, amino acids, nitrogen-fixing bacteria, organic phosphorus mineralizing bacteria, functional prediction.

## Resumen

La amplia distribución de *Agave lechuguilla* Torrey. ha llamado la atención sobre los mecanismos que permiten su sobrevivencia en el Desierto Chihuahuense. Este éxito puede estar relacionado con la asociación de lechuguilla con rizobacterias promotoras de crecimiento vegetal (RPCV) que, mediante diversos mecanismos, aumentan la disponibilidad de nutrientes, ayudan a la planta a tolerar el estrés abiótico, controlan microorganismos patógenos y producen hormonas vegetales que favorecen el crecimiento de la planta. Este trabajo engloba diferentes factores que pueden influir en la compleja interacción planta-suelo-bacterias, así como en los posibles métodos que dirigen el reclutamiento de la comunidad bacteriana en la rizósfera de lechuguilla.

En el primer capítulo, se explora el efecto de las propiedades contrastantes del suelo en las subregiones Central y Meseta del Desierto Chihuahuense, sobre la diversidad bacteriana asociada a la rizósfera de plantas jóvenes y adultas de *A. lechuguilla* T. Se encontró que las propiedades del suelo relacionadas con la disponibilidad de nutrientes estuvieron asociadas con la composición, actividad enzimática y abundancia de las comunidades bacterianas, y a su vez, dichas propiedades pueden ser la consecuencia de la presencia de las plantas de agave y la actividad bacteriana.

El segundo capítulo se aborda el papel de los aminoácidos presentes en los exudados de la raíz de lechuguilla sobre la abundancia y actividad enzimática de las RPCV. Además, exploramos las variaciones en la diversidad funcional de las comunidades de rizobacterias de plantas jóvenes y adultas de lechuguilla. Encontramos que las plantas adultas de lechuguilla, más que las jóvenes o el suelo libre de raíces, estimularon la presencia de funciones relacionadas con el control biológico de patógenos, producción de hormonas vegetales, tolerancia contra estrés abiótico y el aumento de la disponibilidad de nutrientes, especialmente por medio de la actividad de bacterias fijadoras de nitrógeno (BFN) y mineralizadoras de fósforo orgánico (BMFO). Adicionalmente, el aumento de la actividad enzimática y la abundancia de BFN y BMFO, alcanzaron los valores más altos en la rizósfera de

lechuguillas adultas. Probablemente debido a la liberación de arginina y metionina, los cuales son reservas de energía y contribuyen con elementos esenciales para la biosíntesis de proteínas, ya sea que se utilicen directamente o después de su mineralización.

Finalmente, en el tercer capítulo, se estudiaron las variaciones en la abundancia y actividad enzimática de las RPCV en muestras de suelo, como respuesta a diferentes aminoácidos, ácidos orgánicos y azúcares identificados en la rizósfera de lechuguilla. Al respecto, se encontró que los exudados de la raíz de lechuguilla parecen estimular el efecto de imprimación de las comunidades de BFN y BMFO, que conducen a la acumulación de nutrientes tales como el carbono orgánico total, amonio, nitrito y nitrato. Este incremento podría explicar, hasta cierto punto, la mayor fertilidad de la rizósfera, así como la tolerancia de lechuguilla a las condiciones más agrestes del Desierto Chihuahuense.

Con este trabajo se sugiere que la etapa de crecimiento influye en la composición de los exudados de la raíz de *A. lechuguilla* T. y, por consiguiente, en el reclutamiento de bacterias benéficas en la rizósfera, así como de su actividad enzimática. Sin embargo, más estudios son necesarios para entender el complejo proceso de interacción entre planta-bacterias bajo condiciones áridas.

**Palabras clave:** *Agave lechuguilla* Torrey, rizósfera, interacción planta-suelo-microorganismo, redes de co-ocurrencia, exudados de la raíz, aminoácidos, bacterias fijadoras de nitrógeno, bacterias mineralizadoras de fósforo orgánico, predicción funcional.

## General introduction

The soil microbiome can establish close communication with plants, so these organisms are of particular interest among other microbial communities (Lau and Lennon 2011). Either as pathogens or growth-promoters (Scheuring and Yu 2012; Dutta and Podile 2017), the relationship between plant-microorganisms is not totally random but the consequence of a lengthy selection process sheltered in the soil.

The relationship between plants and microorganisms may have originated even from the evolution of the first terrestrial plants, which supposedly were able to leave the aquatic ecosystems with the aid of arbuscular mycorrhizal fungi (Vandenkoornhuyse *et al.* 2015). Said interactions between plant and microorganisms are known as holobiont, a unit with interdependent and complex dynamics (Vandenkoornhuyse *et al.* 2015), that begin their assemble in the first stages of the plant as a seedling and change as the microorganisms settle in different plant compartments (Fonseca-García *et al.* 2018). The soil composition is of great importance as most of these interactions occur in this media, and it is both, the primary source of microbial diversity or the buffer that regulates the availability of nutrients and the communication between the symbionts through metabolites (de Souza *et al.* 2016; Sánchez-Canizares *et al.* 2017). Thus, soil physicochemical properties like the parent material, pH, organic matter concentration, and soil moisture had been reported as factors that shape the bacterial community (Bulgarelli *et al.* 2012; Qiao *et al.* 2017).

In the soil occurs one of the most important interactions between plants and microorganisms. The rhizosphere is the space created by the roots deposition of metabolites released in the soil (Hartmann *et al.* 2008). Several of these plant exudates have been analyzed and amid the most common and abundant compounds are organic acids (acetic, citric, malic, and oxalic), sugars (glucose, maltose, and xylose), and amino acids (alanine, glycine, tryptophan, and proline) (Eilers *et al.* 2010; Badri *et al.* 2013; Anandyawati *et al.* 2017).

It has been shown that through root exudates, the plants can favor the selection of specific microbiomes, such as the phyla Acidobacteria, Actinobacteria,

Bacteroidetes, Firmicutes, and Proteobacteria, which had been reported as the most common and abundant in the rhizosphere (Schreiter *et al.* 2014); and include bacteria that promote plant growth through phytohormone production, enhancing nutrient availability, protecting the plants against pathogens, or reducing the effects produced by abiotic stress (Kinkel *et al.* 2011; Bakker *et al.* 2013).

Therefore, microorganisms fulfill essential functions for plant growth, especially in environments with high temperatures, low water availability, and oligotrophic conditions such as arid lands (Perry and Goodall 1979). Nonetheless, it is considered that these conditions may influence, for example, the nutrition mechanisms of the organisms or tolerance to abiotic stress and, as such, arid lands can sustain a great diversity (López-Lozano *et al.* 2013). In these ecosystems, plants with photosynthesis C3 or C4 may find their adaptation difficult, while those with crassulacean acid metabolism (CAM) like agaves have shown mechanisms that aids in the tolerance to extreme conditions (Stewart 2015). For example, its metabolism permits the $CO_2$ uptake and accumulation of organic acids at night, while the Calvin cycle and sugar synthesis are performed in the daylight with closed stomata, which diminishes the loss of water by transpiration (Matiz *et al.* 2013). Other features that aid in the conservation of water are the internal water storage in the leaves, osmoregulation with fructans, thick cuticles, and even the smaller size and number of stomata (Matiz *et al.* 2013).

Even with this, the association between plant and microorganism may favor the tolerance of plant species in demanding ecosystems, where its growth would otherwise be limited (Soussi *et al.* 2016). Consequently, the study of microorganisms from arid lands takes relevance in a scenario where desertification and alteration of the rainy season are increasingly common (Reynolds *et al.* 2005). The limited availability of water and the process of desertification suffered by the ecosystems urges us to search for biotechnological alternatives to maintain food production and to promote conservation or reforestation of degraded lands. However, this use requires studies that help us understand how factors influence microbiome assemblage, especially the communities with plant growth-promoting mechanisms.

This project proposed to assess the influence of abiotic and biotic conditions on the composition and activity of the bacterial community associated with the rhizosphere of *A. lechuguilla* T. to understand its dynamics and the role of the factors in the environmental filtering of the microbiota. Biotic conditions include the root exudates, which may be influenced along the growth of the plants. In addition, to explore the behavior of plant growth-promoting bacteria and the effect of the metabolites contained in the root exudates, we analyzed nitrogen-fixing bacteria and organic phosphorus mineralizing bacteria associated with the rhizosphere of lechuguilla. To cover all the objectives, this work was divided into three chapters, which respond to different complementary questions: I) At a regional level, which soil properties, and how, influence the rhizobacterial communities associated to *A. lechuguilla* T. at different growth stages? II) What is the role of the amino acids from lechuguilla root exudates over the selection of the bacterial community? and III) At an experimental level, which variations promote the presence of the root exudates from *A. lechuguilla*, over plant-growth promoting rhizobacteria? Each question is addressed at depth in its respective chapter.

## References


Anandyawati, Anwar S, Nugroho B *et al*. Study of Root Exudate Organic Acids and Microbial Population in the Rhizosphere of Oil Palm Seedling. J Trop Soils 2017;22:29–36.

Badri D V., Chaparro JM, Zhang R *et al*. Application of natural blends of phytochemicals derived from the root exudates of arabidopsis to the soil reveal that phenolic-related compounds predominantly modulate the soil microbiome. J Biol Chem 2013;288:4502–12.

Bakker PAHM, Berendsen RL, Doornbos RF *et al*. The rhizosphere revisited: root microbiomics. Front Plant Sci 2013;4:1–7.

Bulgarelli D, Rott M, Schlaeppi K *et al*. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature 2012;488:91–5.

Dutta S, Podile AR. Plant Growth Promoting Rhizobacteria (PGPR): the bugs to debug the root zone. Crit Rev Microbiol 2017;36:232–44.

Eilers KG, Lauber CL, Knight R *et al*. Shifts in bacterial community structure associated with inputs of low molecular weight carbon compounds to soil. Soil Biol Biochem 2010;42:896–903.

Fonseca-García C, Desgarennes D, Flores-Núñez VM *et al*. The Microbiome of Desert CAM Plants: Lessons From Amplicon Sequencing and Metagenomics. In: Nagarajan M (ed.). Metagenomics. Perspectives, Methods, and Applications. Academic Press, 2018, 231–54.

Hartmann A, Rothballer M, Schmid M. Lorenz Hiltner, a pioneer in rhizosphere microbial ecology and soil bacteriology research. Plant Soil 2008;312:7–14.

Kinkel LL, Bakker MG, Schlatter DC. A Coevolutionary Framework for Managing Disease-Suppressive Soils. Annu Rev Phytopathol 2011;49:47–67.

Lau JA, Lennon JT. Evolutionary ecology of plant-microbe interactions: Soil microbial structure alters selection on plant traits. New Phytol 2011;192:215–24.

López-Lozano NE, Heidelberg KB, Nelson WC *et al*. Microbial secondary succession in soil microcosms of a desert oasis in the Cuatro Cienegas Basin, Mexico. PeerJ 2013;1:e47.

Matiz A, Tamaso P, Yepes A *et al*. CAM Photosynthesis in Bromeliads and Agaves: What Can We Learn from These Plants? Photosynthesis. InTech, 2013, 91–134.

Perry RA, Goodall DW. Arid Land Ecosystems : Structure, Functioning, and Management. London: Cambridge University Press, 1979.

Qiao Q, Wang F, Zhang J *et al*. The Variation in the Rhizosphere Microbiome of Cotton with Soil Type, Genotype and Developmental Stage. Sci Rep 2017;7:3940.

Reynolds JF, Maestre FT, Huber-Sannwald E *et al*. Aspectos socioeconómicos y biofísicos de la desertificación. Ecosistemas 2005;14:3–21.

Sánchez-Cañizares C, Jorrín B, Poole PS *et al*. Understanding the holobiont: the interdependence of plants and their microbiome. Curr Opin Microbiol 2017;38:188–96.

Scheuring I, Yu DW. How to assemble a beneficial microbiome in three easy steps. Ecol Lett 2012;15:1300–7.

Schreiter S, Ding GC, Heuer H *et al*. Effect of the soil type on the microbiome in the rhizosphere of field-grown lettuce. Front Microbiol 2014;5:1–13.

Soussi A, Ferjani R, Marasco R *et al*. Plant-associated microbiomes in arid lands: diversity, ecology and biotechnological potential. Plant Soil 2016;405:357–70.

de Souza RSC, Okura VK, Armanhi JSL *et al*. Unlocking the bacterial and fungal communities assemblages of sugarcane microbiome. Sci Rep 2016;6:28774.

Stewart JR. Agave as a model CAM crop system for a warming and drying world. Front Plant Sci 2015;6:684.

Vandenkoornhuyse P, Quaiser A, Duhamel M *et al*. The importance of the microbiome of the plant holobiont. New Phytol 2015;206:1196–206.

# Chapter 1. The nutrient-improvement bacteria selected by *Agave lechuguilla* Torr. and their role in the rhizosphere community

## 1.1 Introduction

Drylands are widely spread ecosystems covering about 45% of the surface of the earth (Prăvălie 2016). Besides the low water availability and high evapotranspiration rate, arid soils yield a low biomass production which is highly related to nitrogen (N) and phosphorus (P) (Pointing and Belnap 2012). With the scant deposition of organic matter in the arid soils (Augusto *et al*. 2017), the primary source of this nutrient depends on *nitrogen-fixing bacteria (*NFB) such as the genera *Bryobacter* (Acidobacteria), *Succinispira* (Firmicutes), *Burkholderia*, *Microvirga*, and *Rhizobium* (Proteobacteria) (Karp *et al*. 2019). Phosphorus availability, in turn, depends directly on the soil parent material, weathering process, and, once it is incorporated into the biomass, the recycling of the organic matter (Crain *et al*. 2018). This way, *organic phosphorus mineralizing bacteria (*OPMB), through the action of phosphatases, are essential for phosphorus availability in soils (Turrion *et al*. 2010). Bacteria from the phyla Actinobacteria, Cyanobacteria, Deinococcus-Thermus, and Proteobacteria had been recorded as producers of these enzymes (Ragot, Kertsez and Bünemann 2015). Together NFB and OPMB increase the availability of N and P, the most essential nutrients for plant development, and thus, may help to ameliorate the abiotic stress and even increase the resistance against pathogens (Wang *et al*. 2019), which is the reason why these nutrient-improvement bacteria are some of the most crucial *plant growth-promoting rhizobacteria* (PGPR) in drylands (Rashid *et al*. 2016). Among the biotic factors that affect PGPR, root exudates have an essential role in microbiome selection and enzymatic activity. These compounds include a mix of sugars, organic acids, amino acids, phenolic compounds, and secondary

---

A modified version of this chapter was published as:

Medina-de la Rosa, G., García-Oliva, F., Alpuche-Solís, Á.G., Ovando-Vázquez, C., and López-Lozano, N.E. (2021) The nutrient-improvement bacteria selected by *A. lechuguilla* T. and their role in the rhizosphere community. FEMS Microbiol Ecol 97: 1–14.

https://doi.org/10.1093/femsec/fiab137

metabolites (Naylor and Coleman-Derr 2018), which fluctuate depending on abiotic conditions and plant genotype compartment, and growth stage (Coleman-Derr *et al*. 2016). For example, mature plants of carrots, cabbage, and potatoes that do not require as many nutrients as younger plants (Hennion *et al*. 2019) produce root exudates with a higher carbon concentration that favor a microbiome with potential for nitrogen fixation, biological control, and metal detoxifiers (Dechassa and Schenk 2004). This change has also been studied in the dominant bacteria in the rhizosphere of leguminous shrubs, where Proteobacteria is replaced by beneficial taxa from the phyla Acidobacteria, Bacteroidetes, Chloroflexi and Firmicutes in the mature stages of the plant (Na *et al*. 2018). Nonetheless, this behavior has yet to be explored on plants from natural ecosystems and not only for economic interest.

Ecological interactions such as competition, depredation, mutualism, etcetera that occur within microbial communities can be represented as networks based on correlations between the relative abundance of the community members. These networks help us to explore microbe-microbe interactions with greater precision (Shi *et al*. 2016), as well as to identify the taxa with more connections (hubs or keystone genera) that maintain the network structure, and also the factors involved in its delineation like edaphic conditions and plant-microbe interactions (Kinkel, Bakker and Schlatter 2011). In this regard, beneficial bacteria in the rhizosphere could aid in plant development even under stressful conditions from arid lands (Barton and Northup 2011). Furthermore, studying these ecosystems could lead us to understand how organisms are adapted to live in those harsh conditions and the complexity of the ecological processes behind their functioning.

The Chihuahuan desert is the largest dryland in Mexico and is the second in worldwide biodiversity (Granados-Sánchez *et al*. 2011). Among its subregions, Central and Meseta have contrasting characteristics, such as lower water availability in the former, which increases the accumulation of sodium in the soil. Regardless, both subregions are relevant conservation sites due to the diversity of organisms and the high number of endemism registered (Villarreal-Quintanilla *et al*. 2017). *A. lechuguilla* T. has considerable ecological and biotechnological importance (Castillo-

Quiroz, Cano-Pineda and Berlanga-Reyes 2012) and is one of the most common shrubs inside the Chihuahuan desert, where covers a surface of around 20 million ha of poorly developed and limestone-derived soils (García-Arévalo 2002). The wide distribution of *A. lechuguilla* T. is associated with its crassulacean acid metabolism, genetic variability, and its participation in soil formation (Narcia *et al*. 2012). However, it has been hypothesized that plant-microbiome interactions may also have a critical role in lechuguilla tolerance to arid conditions. In a previous study, López-Lozano *et al*. (2020) reported that lechuguillas from Cuatro Ciénegas Valley in Central Chihuahuan Desert recruit specific rhizosphere microbes such as the nutrient-improvement bacteria Bradyrhizobia, Microvirga, Rhizobia, order Sphingomonadales (Proteobacteria), family Gaiellaoceae and, Microlunatus (Actinobacteria), with functional traits that increase nutrient availability, especially nitrogen and phosphorus. Besides, seasonality and soil properties such as pH, electrical conductivity, and the C:N ratio were the main factors influencing the bacterial community structure in the rhizosphere of lechuguilla (López-Lozano *et al*. 2020). This study highlights the association of nutrient-enhancing bacteria and their host under oligotrophic conditions. However, whether the stage of the plant development may affect rhizosphere communities and how their interactions change in different soils and climatic conditions is yet to be answered.

To understand the complex interaction soil-plant-microbiome, we investigated the bacterial community of *A. lechuguilla* T. at different growth stages in two contrasting subregions from the Chihuahuan Desert. In this regard, we hypothesized that mature plants will select and enhance the enzymatic activity and abundance of nutrient improvement bacteria and that these groups will be of great importance in the microbial interaction networks within the rhizosphere. The specific objectives were: 1) Determine the abiotic factors (physicochemical soil properties) influencing the bacterial community composition of *A. lechuguilla* T. from two subregions of the Chihuahuan Desert, 2) Identify whether the bacterial community of young and mature plants of *A. lechuguilla*, as well as bulk soil, has a different composition by sequencing the 16S rRNA gene, 3) Study the networks of these bacterial

communities, in order to identify the keystone genera in the rhizosphere of young and mature lechuguilla, and bulk soil communities and 4) finally, as nutrient availability is essential in these environments, we determined the effect of the plant stage of lechuguilla and which soil properties influence the enzymatic activity and abundance of NFB and OPMB, quantifying the nitrogenase and alkaline phosphomonoesterase activities, as well as the copy numbers of their respective genes (*nifH* or *phoD*), with real-time polymerase chain reaction (qPCR).

## 1.2 Methods

### 1.2.1 Site description and sampling collection

Sampling was made at the end of the dry season in 2018 in two subregions of the Chihuahuan Desert. The Central subregion is located at an altitude between 700 to 1,400 m, has a predominant hot desert climate (Bwh), annual precipitation of 313 mm, and temperatures ranging from 10.1°C to 28.7°C. Meseta subregion has an altitude of 1,500 to 2,100 m. This land has a hot semi-arid climate (Bsh), annual precipitation of 667 mm, and temperatures between 10.5°C to 24.6°C. Both subregions have calcaric soils; however, in the Central subregion are lithosols, while in the Meseta are gypsic xerosols and lithosols.

Two sampling sites were selected for each subregion: Tres Coronas (26°36'35.92"N, 103°56'46.53"W) and Álamos (26°38'29.07"N, 104° 3'20.95"W) in Central subregion, and Charco Blanco (22°38'26.80"N, 100°30'39.25"O) and Los Amoles (22°44'8.53"N, 100°29'46.56"W) in the Meseta subregion (Figure 1.2.1).

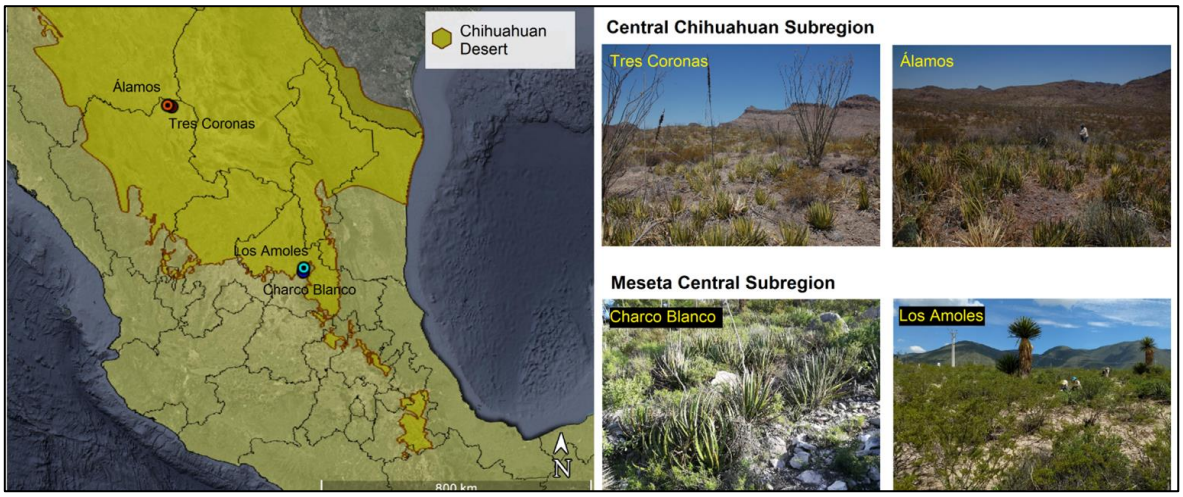



Figure 1.2.1.Sampling sites inside the Chihuahuan Desert. Each subregion had two experimental units: Álamos (26°38'29.07"N, 104°3'20.95"W) and Tres Coronas (26°36'35.92"N, 103°56'46.53"W) for Central; Amoles (22°44'8.53"N, 100°29'46.56"W) and Charco Blanco (22°38'26.80"N, 100°30'39.25"O) for Meseta (Map from Google Earth).

The sampling was made in "lechuguillales", large and almost pure patches of lechuguilla, which are characteristic of the agave species, to avoid the influence of other plants (Granados-Sánchez *et al*. 2011). We delimited 20 $m^2$ within these

patches, and four mature and four young specimens of *A. lechuguilla* T. were chosen. Mature agaves were near reproductive maturity, with at least 40 leaves (Freeman and Reid 1985) but without inflorescence. Young plants were those still joined to the mature specimen by the rhizome and with about 20 leaves. There had to be at least 2 m between them to avoid sampling the same specimen (Trame, Coddington and Paige 1995). As the rhizosphere is challenging to delimit (Philippot *et al*. 2013), we considered the soil attached to the roots for molecular analyses. To complete the amount of soil required for enzymatic activity and physicochemical analysis, we also used the soil surrounding the roots without exceeding the leaves' circumference span. We collected *in situ* rhizosphere samples of each plant in four equidistant points around the root system within the first 10 cm below the surface. These samples were mixed and homogenized to form one sample. We collected bulk soil samples from four patches of soil with no plants or roots at the same depth. Samples were taken with a sterile spatula to avoid contamination. About 1 kg of soil was kept at 4°C in dark plastic bags for enzymatic activity and soil properties analysis, while 2 g were kept in Eppendorf tubes at -80°C for molecular procedures.

### 1.2.2 Determination of physicochemical soil properties

The percentage of soil moisture (SM) was quantified according to constant weight using the gravimetric method. The soil was oven-dried at 60°C and sieved (2mm). The method of Bouyoucos was applied to determine texture. Electric conductivity and pH were obtained through the electrometric method with deionized water (1:2 ratio). Total organic carbon (TOC) and nitrogen (TN) were determined with Schubert & Nielsen method in a combustion elemental analyzer COSTECH (ECS4010) (Schubert and Nielsen 2000). Extractable phosphorus, nitrite + nitrate ($NO_2$+$NO_3$), and ammonium ($NH_4$) concentrations were determined according to Bray-Kurtz (Bray and Kurtz 1945), vanadium reduction (García-Robledo, Corzo and Papaspyrou 2014), and a variation to the Berthelot method (Nelson 1983), respectively. Lastly, the extraction of micronutrients iron (Fe), magnesium (Mn), zinc (Zn), and copper (Cu); and exchangeable cations calcium ($Ca^{2+}$), magnesium ($Mg^{2+}$), sodium ($Na^{+}$), and potassium ($K^{+}$), was made with diethylene triamine penta-acetic

acid and ammonium acetate, respectively (Lindsay and Norvell 1978). Quantification of these elements was performed by Laboratorio Nacional de Biotecnología Agrícola, Médica y Ambiental (LANBAMA-IPICyT), with inductively coupled plasma atomic emission spectroscopy (ICP-OES).

### 1.2.3 DNA extraction, 16S rRNA gene sequencing, and sequence analysis

Genomic material was extracted from 250 mg of soil using Quick-DNA Fecal/Soil Microbe Miniprep kit® from Zymo Research (D6010) and quantified by NanoDrop 2000 (Table 0-1 and 0-2). Bacterial community composition was determined by 16S rRNA gene sequencing, which was performed by Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO), on the Illumina MiSeq platform (2x300), with primers 357F-5'CTCCTACGGGAGGCAGCAG-3' (Turner *et al*. 1999) and CDR-5'-CTTGTGCGGGCCCCCGTCAATTC-3' (Rudi *et al*. 1997), for V3-V5 regions. Reagent-only and mock community samples (ZymoBIOMICS™ Microbial Community Standard II, Log Distribution) were included for quality control. We confirmed with the mock community that taxa abundance and identification accuracy were acceptable above 0.01% of relative abundance. The bacterial composition of the reagent-only sample was statistically different from the soils, with a high percentage of phyla Cyanobacteria (2.25%) and Firmicutes (10.99%), whereas, in contrast, soil samples were around 0.15% and 4.57%, respectively. Besides, the number of reads was lower than that obtained for the samples (Table 0-1 and 0-2), and the similarity with the samples was tested with a non-metric multidimensional scaling (NMDS), where we can observe a clear separation between the blank and soil samples (Figure 1.2.2). Hence, we consider that the contamination during the handling of the samples and their processing was negligible.

The sequences analysis was performed with Mothur (V. 1.35.1) (Schloss *et al*. 2009). Firstly, the reads were filtered according to several parameters, I. e. Quality (Q-value ≥ 25), number of homopolymers (<8), and the ambiguities were omitted. Chimeric sequences were identified and excluded with the UCHIME algorithm. After that, sequences were aligned with the database SILVA 16S rRNA gene (V. 132), using the

Nearest alignment space termination algorithm (NAST), and trimmed to ensure optimal alignment. No redundant sequences were obtained, and a distance matrix was computed for operational taxonomic unit (OTU) assignation (97% similarity). Finally, taxonomic categorization was made with Mothur Bayesian classifier and SILVA database (https://www.arb-silva.de/). Alpha diversity indexes Chao 1, Shannon, and Simpson, were calculated and rarefaction curves were obtained to assess the species accumulation at the same sampling effort.

The reads obtained from Illumina sequencing are in the GenBank Sequence Read Archive under the Bioproject accession PRJNA672793.

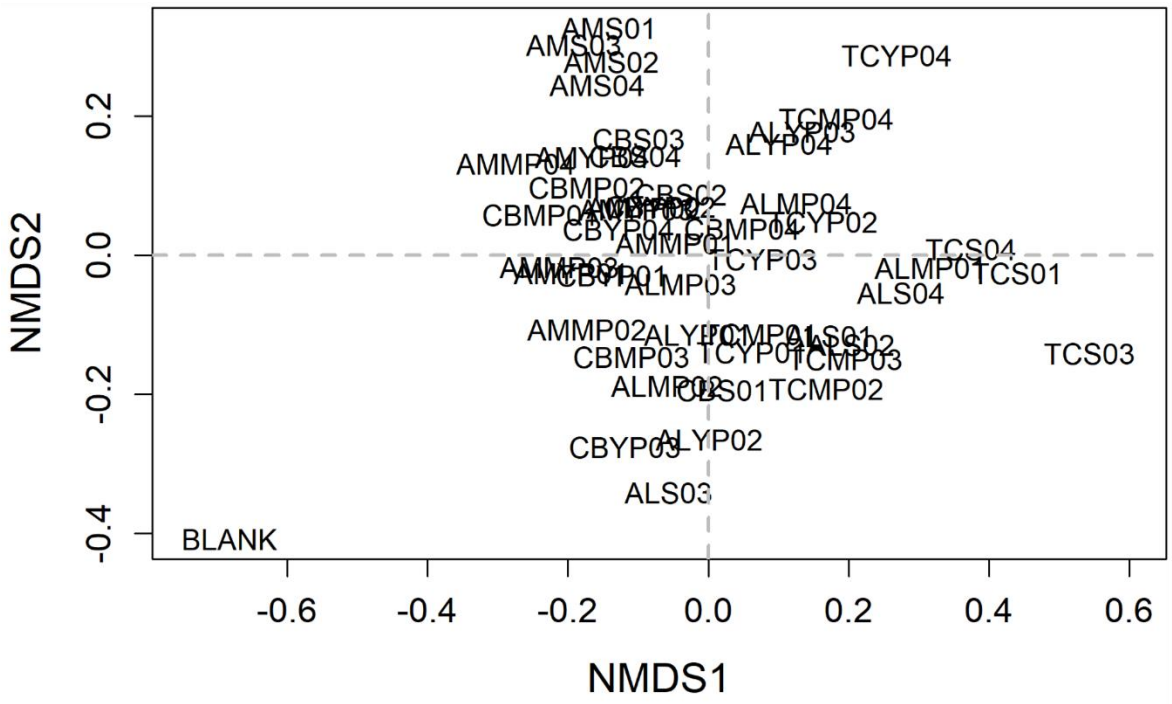



Figure 1.2.2. NMDS of bacterial communities in soil samples against blank. Sample ID provides information about locations (AL-Los Álamos, TC-Tres Coronas, CB-Charco Blanco, AM-Amoles), type (MP- Mature plant, YP-Young plant, and S- Bulk soil), and replicates (Stress=0.162).

### 1.2.4 Co-occurrence networks analysis

Bacterial community networks were obtained in R (V. 4.0.2) with the package igraph (V. 1.2.4.2) (Csárdi and Nepusz 2006). Networks were constructed with Spearman correlations of genera above 0.01% of relative abundance (Shi *et al*. 2019) and False Discovery rate (FDR) correction (Benjamini and Hochberg 1995). The cutoff values

were FDR<0.05 and ρ>0.7 (Garcia and Kao-Kniffin 2019). The networks were generated considering the correlations as the weights of the network. Modules were computed with the cluster walk-trap method, which finds the densely connected subgraphs within the network via random walks over the weighted edges (Hoffman *et al*. 2018). Hub score of nodes was calculated with the Kleinberg logarithm, based on the Eigen centrality value of the vectors, which is related to the number of connections of the nodes in the network, and thus, the nodes with high values are those with a high number of connections (Kleinberg 1999). Minimum spanning trees (MST) or forest for disconnected networks, were calculated with Prim's algorithm for weighted graphs to identify their top nodes (Prim 1957). For MST, the distance between nodes was defined as the inverse of the weights (correlations). Topology network properties and centrality measures were determined for an undirected graph, such as diameter, mean distance, eigenvalue, betweenness, modularity, and degree. Scale-free distribution was tested to fit the degrees of the network to the power-law model (Judd, Small and Stemler 2013; Broido and Clauset 2019). Lastly, it was confirmed that empiric networks were different from random networks (n = 1,000).

### 1.2.5 NFB and OPMB abundance

The abundance of NFB and OPMB on soil samples were quantified with real-time quantitative polymerase chain reaction (qPCR) in a PikoReal 96 Real-Time PCR System (TCR0096, Thermo Fisher Scientific Inc.). The reaction for *nifH* gene had a final volume of 10 µl: 5 µl SYBR Green 2X, 0.3 mM primer PolR (5'-ATSGCCATCATYTCRCCGGA-3') (Poly, Monrozier and Bally 2001), 0.15 mM FPGH19 (5'-TACGGCAARGGTGGNATHG-3') (Simonet *et al*. 1991), and 1 µl DNA (1:20). Amplification conditions were: Initial denaturalization of 10 min at 95°C, 40 cycles of denaturalization at 94°C, alignment at 57°C, and extension 72°, one minute each, with a final extension of 5 min at 72°C (Simonet *et al*. 1991). *phoD* gene reaction had a final volume of 20 µl: 10 µl SYBR Green 2X, 0.4 mM from each primer (F733 5'-TGGGAYGATCAYGARGT-3' and R1083 5'-CTGSGCSAKSACRTTCCA-3') and 1 µl DNA (1:10). Amplification was a three-step reaction: Initial denaturalization of 10 min at 95°C, followed by 40 cycles of 15 s at 95°C, alignment at 58°C, and

extension of 72°C (one minute each), with a final extension of 5 min at 72°C (Ragot, Kertsez and Bünemann 2015). Both reactions included a final melting step from alignment temperature to 95°C. Standard curves were run alongside the samples using 100 to 108 copies of gene clones obtained with pGEM-T easy Vector system (PROMEGA) from *Bacillus subtilis* (*phoD*) and *Geobacter sulfurreducens* (*nifH*). The standard curve was optimized close to 100% efficiency. Inhibition in the qPCR reactions was tested by mixing serial dilutions of DNA extracted from the soil against a known standard DNA before qPCR. The Ct values of the standard DNA did not change in the diluted soil DNA, indicating the absence of severe inhibition.

### 1.2.6 Enzymatic activity

Potential nitrogenase activity in soils was quantified with the acetylene reduction method (López-Lozano, Carcaño-Montiel, and Bashan 2016). Briefly, serologic flasks (14 ml) were filled with 6 g of soil and sealed with rubber stoppers and aluminum caps. About 20% of headspace air was replaced with acetylene. After seven days of incubation at 37°C, ethylene production was measured in 250 µl of headspace using gas chromatograph Agilent 6890 GC N.05.04 with a flame ionization detector, with air and hydrogen as carriers. A standard curve of ethylene (0.3-2.1 nmol) was run for quantification. The results are reported as nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$.

Alkaline phosphomonoesterase (ALP) activity was determined with the modified colorimetric method developed by Tabatabai and Bremner (1969). About 0.5 g of soil was mixed with 2.375 ml Buffer Tris 0.5 M (pH 11) and 0.125 ml of pNPP (p-Nitrophenol phosphate, Sigma-Aldrich P7998) as substrate. The mix was incubated for 30 min at 37°C (150 rpm). After that, 2 ml of NaOH 0.5 M and 0.5 ml of $CaCl_2$ 0.5 M were added. Samples were centrifuged for 5 min at 2,500 rpm, and their absorbance was read at 405 nm. Blanks without pNPP were run alongside the samples. For the quantification, a standard curve with p-nitrophenol up to 180 µg $ml^{-1}$ pNP was made. The results are reported as µg pNP $g^{-1}$ $h^{-1}$.

### 1.2.7 Statistical analysis

All the analyses were performed with the software R (V. 4.0.2) (R Core Team 2020). Firstly, the difference in soil properties between subregions and sample type (bulk and rhizosphere from mature or young plants) was evaluated with a two-way analysis of variance (ANOVA), variances homogeneity of the residuals was verified. To visualize the difference between samples, a principal component analysis (PCA) was performed with the packages FactoMineR (V. 3.5.3) and Factoextra (V. 3.5.3) (See supplementary material d). Soil properties with correlation $\rho > 0.8$ (Spearman) were omitted. The influence of soil properties over the composition of bacterial communities' composition was identified with a non-metric multidimensional scaling (NMDS) and environmental fit ($p<0.01$), with the package vegan (V. 2.5-7), and differences in the bacteria community composition were tested by analysis of similarities (ANOSIM) (See supplementary material c). The heatmap of relative abundances was obtained from bacterial genera (>1%), and the samples were clustered according to Bray Curtis distance (pheatmap V. 3.5.3) (See supplementary material b). Differences in relative abundance between subregion and sample types were determined with the Kruskal test and Dunn *post hoc*. We used the sparse partial least square discriminant analysis (sPLS-DA), a supervised learning method (Paliy and Shankar 2016), to identify the characteristic taxa between sample types with the same taxa used for network construction (mixOmics V. 6.6.2.). To validate the model, we calculated the mean of the Receiver operating characteristic (ROC) and considered as an acceptable accuracy a value above 40% (Westerhuis *et al*. 2008) (See supplementary material h).

For the results of enzymatic activity and abundance of NFB and OPMB, two-way ANOVA was performed to identify differences between subregions and sample type. Linear regression models with soil properties were obtained to determine their influence over the bacterial groups. The order of the variables in the model followed the mean decrease accuracy (%IncMSE), obtained from linear regression with the classification algorithm within the package randomForest (V. 4.6-14) (Table 0-3, Supplementary material e). Models were constructed with the backward stepwise

selection method. Finally, the correlation between abundance and enzymatic activity from NFB and OPMB against the relative abundance of bacterial genera was calculated (corrplot V. 0.84). Those correlations with $\rho>0.4$ ($p<0.01$) were kept, and the enzymatic activity or gene presence was verified in the literature.

## 1.3 Results

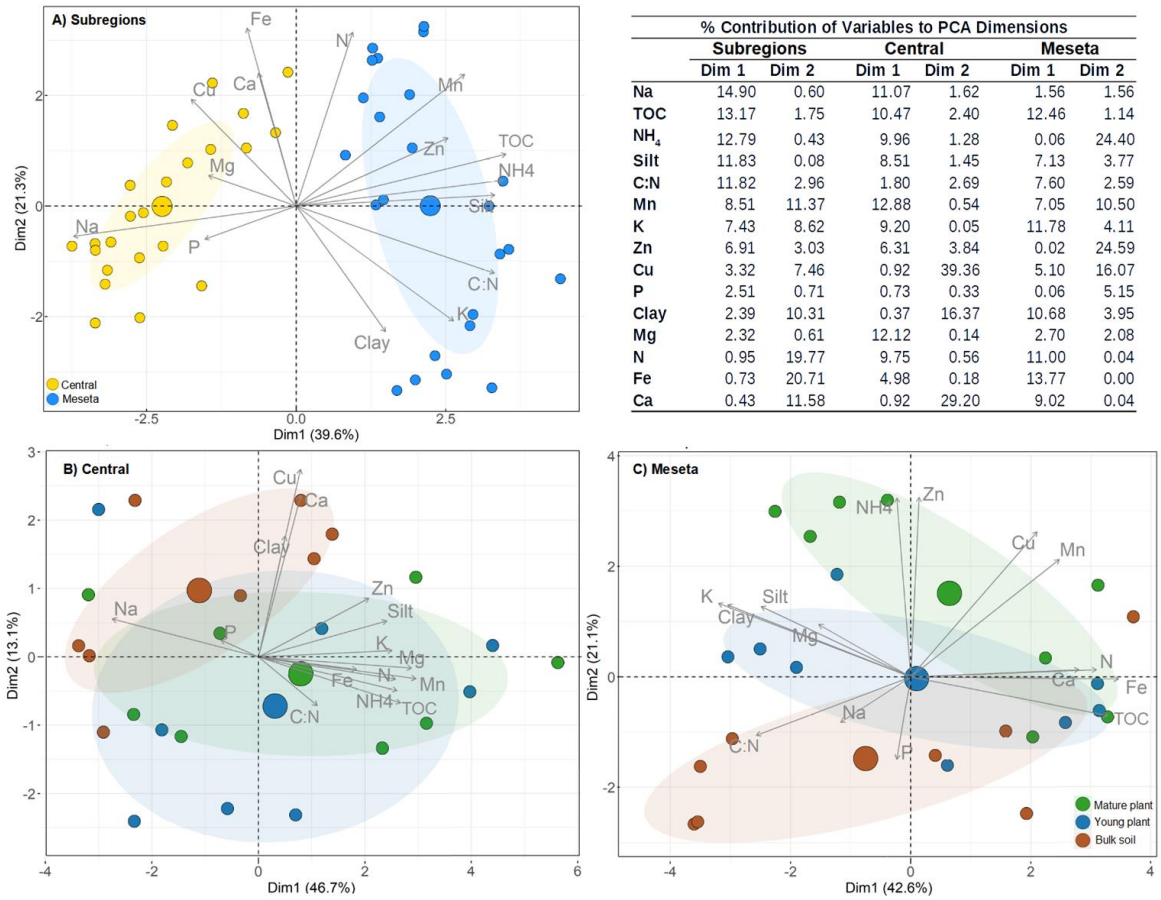
### 1.3.1 The nutrient concentration in soils between the two subregions of the Chihuahuan Desert.

The main difference between the Chihuahuan desert subregions was found in the nutrient availability, e.g., The Meseta subregion had a higher concentration of TOC, $NH_4$, and $NO_2$+$NO_3$ than the Central subregion, which is reflected on a C:N ratio up to 45.7 ± 3.7 in the former, against 11.4 ± 0.8 in the latter ($p<0.001$) (Table 1.3-1). Besides, among exchangeable cations and micronutrients, Na concentration in the Central subregion (27.74 ± 1.16 g $kg^{-1}$) was about threefold the concentration in Meseta (9.36 ± 0.20 g $kg^{-1}$). Mg and Cu also had higher concentrations in the Central subregion, 0.15 ± 0.01 and 0.47 ± 0.02 g $kg^{-1}$, respectively ($p<0.001$) (Table 1.3-1). Thus, the concentration of Na, $NH_4$, Mn, TOC, and C:N ratio held the main variation between subregions and were relevant for sample differentiation in the PCA (Variance Dim 1 + Dim 2 = 60.90%) (Figure 1.3.1).

The nutrient concentration was also higher in soils under the influence of plants, a pattern found in both subregions (Table 1.3-1). For example, in bulk soil from the Central subregion, TOC was about 1.78 ± 0.20%, while the rhizosphere from mature and young agaves had 2.94 ± 0.47% and 3.04 ± 0.51%, respectively ($p < 0.05$). Meanwhile, in the Meseta subregion, the bulk soil had 9.64 ± 0.45% TOC in contrast to the rhizosphere of mature (10.19 ± 0.45%) and young agaves (10.45 ± 0.44%) ($p < 0.05$). Besides, all nitrogen forms were higher in the rhizosphere of both stages of lechuguilla (Table 1.3-1). A PCA was performed for each subregion (Figure 1.3.1B–C). We found that high concentration of Na for bulk soil samples and nutrients for both rhizospheres were the principal soil properties for clustering in the Central subregion (Variance Dim 1 + Dim 2 = 59.80%). In the Meseta, this pattern was similar, though the difference between young and mature rhizosphere samples was accentuated by dimension 2, with $NH_4$, Zn, Mn, and Cu as the variables with more contribution to the model (Variance Dim 1 + Dim 2 = 63.70%).

Table 1.3-1. Physicochemical properties from bulk and rhizosphere from *A. lechuguilla* T., from Central and Meseta subregions of Chihuahuan Desert. Values depict mean and standard error (n=8). Two-way ANOVA was performed for subregion and sample type (bulk soil, mature or young plant) variables (*** p <0.001, ** p <0.01, * p <0.05).

| | Central | | | Meseta | | | Two-way ANOVA (F-value) | | |
|---|---|---|---|---|---|---|---|---|---|
| **Soil properties** | **Mature plant** | **Young plant** | **Bulk Soil** | **Mature plant** | **Young plant** | **Bulk Soil** | **Sample type** | **Subregion** | **Sample type * Subregion** |
| **% Total C** | 2.94±0.47 | 3.04±0.51 | 1.78±0.20 | 10.19±0.45 | 10.45±0.44 | 9.64±0.45 | 3.236* | 449.212*** | 0.265 |
| **% Total N** | 0.32±0.02 | 0.26±0.03 | 0.21±0.03 | 0.37±0.04 | 0.35±0.04 | 0.23±0.03 | 7.668** | 3.872 | 0.582 |
| **C:N ratio** | 8.80±0.80 | 11.41±0.83 | 9.30±1.00 | 29.50±2.00 | 32.70±3.40 | 45.70±3.70 | 6.792** | 193.208*** | 7.410** |
| **mg $NH_4$ $kg^{-1}$** | 5.85±1.22 | 5.60±1.29 | 3.55±0.30 | 22.10±2.31 | 18.87±2.02 | 10.22±1.23 | 11.587*** | 92.497*** | 5.098* |
| **mg $NO_2$+$NO_3$ $kg^{-1}$** | 9.95±1.98 | 12.00±3.33 | 9.02±0.84 | 86.54±11.77 | 72.86±7.01 | 42.37±656 | 6.68** | 118.46*** | 5.84** |
| **mg P $kg^{-1}$** | 1.91±0.73 | 2.86±1.10 | 1.77±0.64 | 0.63±0.05 | 0.50±0.03 | 2.07±0.92 | 0.428 | 3.73 | 1.796 |
| **g Ca $kg^{-1}$** | 0.91±0.01 | 0.87±0.03 | 0.91±0.03 | 0.89±0.02 | 0.89±0.02 | 0.88±0.02 | 0.408 | 0.562 | 0.542 |
| **g K $kg^{-1}$** | 0.25±0.02 | 0.27±0.02 | 0.27±0.01 | 0.44±0.07 | 0.43±0.08 | 0.40±0.08 | 0.033 | 12.346** | 0.156 |
| **g Mg $kg^{-1}$** | 0.15±0.01 | 0.14±0.01 | 0.12±0.01 | 0.10±0.01 | 0.11±0.01 | 0.09±0.01 | 3.309* | 28.998*** | 0.212 |
| **g Na $kg^{-1}$** | 24.41±1.29 | 24.93±1.22 | 27.74±1.16 | 9.36±0.20 | 8.95±0.16 | 9.36±0.68 | 2.144 | 487.492*** | 1.767 |
| **g Cu $kg^{-1}$** | 0.41±0.03 | 0.39±0.03 | 0.47±0.02 | 0.39±0.02 | 0.31±0.02 | 0.27±0.03 | 1.774 | 23.064*** | 6.560** |
| **g Fe $kg^{-1}$** | 4.83±0.43 | 4.26±0.32 | 3.74±0.41 | 4.14±0.42 | 3.5±0.46 | 3.65±0.62 | 1.613 | 1.799 | 0.302 |
| **g Mn $kg^{-1}$** | 8.94±1.58 | 8.41±1.31 | 6.23±1.04 | 15.55±0.89 | 12.03±0.64 | 11.64±1.67 | 3.602* | 26.367*** | 0.734 |
| **g Zn $kg^{-1}$** | 0.45±0.13 | 0.40±0.14 | 0.31±0.12 | 1.06±0.17 | 0.64±0.07 | 0.61±0.15 | 2.64 | 12.091** | 1.138 |
| **dS $m^{-1}$** | 0.20±0.02 | 0.17±0.02 | 0.12±0.01 | 0.26±0.02 | 0.24±0.01 | 0.19±0.01 | 8.672*** | 20.925*** | 0.016 |
| **pH** | 8.4±0.1 | 8.4±0.10 | 8.6±0.1 | 8.0±0.2 | 8.1±0.2 | 8.2±0.2 | 1.338 | 14.00*** | 0.052 |
| **% Soil moisture** | 3.35±1.51 | 1.47±0.27 | 1.23±0.58 | 23.19±1.73 | 23.89±3.32 | 18.67±2.01 | 2.563 | 242.586*** | 1.268 |
| **% Sand** | 56.35±3.83 | 58.29±2.91 | 56.04±2.58 | 40.10±3.93 | 42.20±3.32 | 43.05±3.17 | 0.19 | 30.984*** | 0.153 |
| **% Clay** | 13.90±1.16 | 13.86±1.14 | 15.36±0.84 | 16.20±1.96 | 17.90±2.83 | 14.75±1.58 | 0.155 | 1.85 | 0.935 |
| **% Silt** | 29.75±4.17 | 27.84±2.47 | 28.59±2.24 | 43.68±2.05 | 39.85±1.34 | 42.18±1.74 | 0.66 | 41.680*** | 0.083 |




| % Contribution of Variables to PCA Dimensions | | | | | | |
|---|---|---|---|---|---|---|
| | Subregions | | Central | | Meseta | |
| | Dim 1 | Dim 2 | Dim 1 | Dim 2 | Dim 1 | Dim 2 |
| Na | 14.90 | 0.60 | 11.07 | 1.62 | 1.56 | 1.56 |
| TOC | 13.17 | 1.75 | 10.47 | 2.40 | 12.46 | 1.14 |
| $NH_4$ | 12.79 | 0.43 | 9.96 | 1.28 | 0.06 | 24.40 |
| Silt | 11.83 | 0.08 | 8.51 | 1.45 | 7.13 | 3.77 |
| C:N | 11.82 | 2.96 | 1.80 | 2.69 | 7.60 | 2.59 |
| Mn | 8.51 | 11.37 | 12.88 | 0.54 | 7.05 | 10.50 |
| K | 7.43 | 8.62 | 9.20 | 0.05 | 11.78 | 4.11 |
| Zn | 6.91 | 3.03 | 6.31 | 3.84 | 0.02 | 24.59 |
| Cu | 3.32 | 7.46 | 0.92 | 39.36 | 5.10 | 16.07 |
| P | 2.51 | 0.71 | 0.73 | 0.33 | 0.06 | 5.15 |
| Clay | 2.39 | 10.31 | 0.37 | 16.37 | 10.68 | 3.95 |
| Mg | 2.32 | 0.61 | 12.12 | 0.14 | 2.70 | 2.08 |
| N | 0.95 | 19.77 | 9.75 | 0.56 | 11.00 | 0.04 |
| Fe | 0.73 | 20.71 | 4.98 | 0.18 | 13.77 | 0.00 |
| Ca | 0.43 | 11.58 | 0.92 | 29.20 | 9.02 | 0.04 |

Figure 1.3.1. Principal Component Analysis (PCA) of physicochemical properties of *A. lechuguilla* T. rhizosphere. A) PCA shows division between Central and Meseta Subregions. With each subset, a new PCA was made, points in B and C represent bulk soil, mature and young plants. The size of the vectors is proportional to percentage of contribution in the two components (Contribution to each dimension is summarized in table).

### 1.3.2 Sodium, carbon, and nitrogen are the main drivers of the bacterial community composition

Sequencing of 16S rRNA gene regions from V3 to V5 reported a mean of 113,844 reads in the libraries of each sample, and due to a low number of reads (1,744), one bulk soil sample from the Central subregion was omitted (Table 0-1). However, the remaining seven samples were enough to perform the analysis. The mean Shannon index per sample was 5.17, and 0.02 for Simpson index (Table 0-1 and 0-2). There were no significant differences in the diversity indexes and estimators between the subregions or sample types. In addition, the rarefaction curves showed adequate coverage of the diversity (Figure 1.3.2).

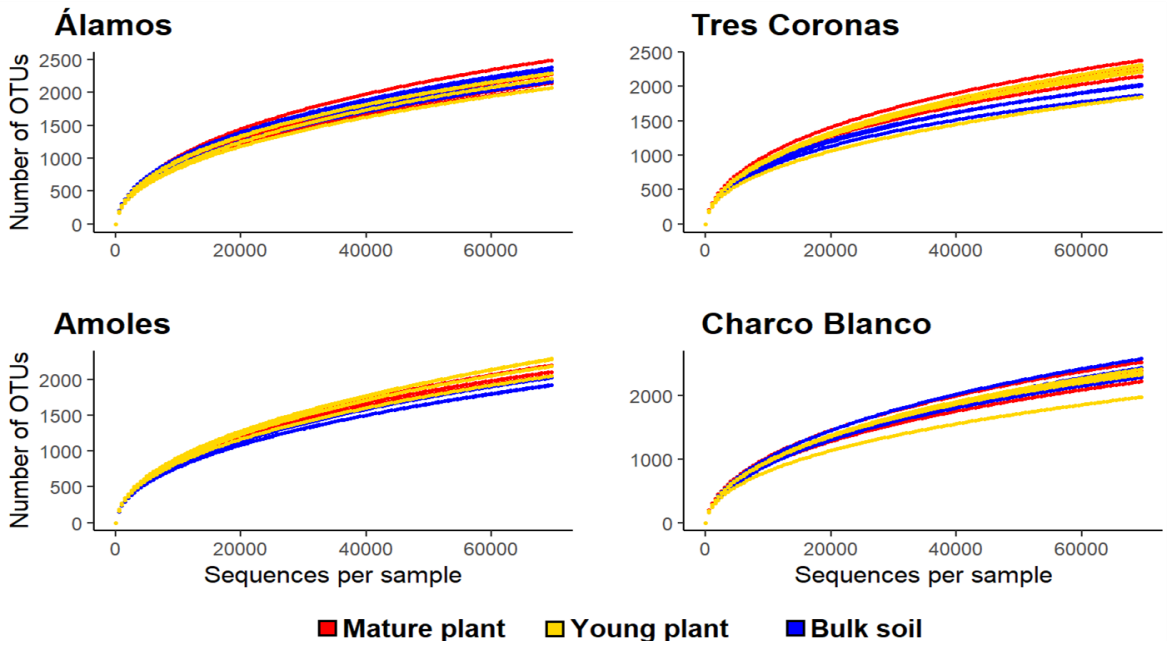



Figure 1.3.2. Rarefaction curves of bacterial gene 16S rRNA by sampling site. Rarefaction curves of bacterial OTUs in the rhizosphere of *A. lechuguilla* T. and bulk soil from the Chihuahuan Desert. Data was obtained with a 97% threshold of similarity.

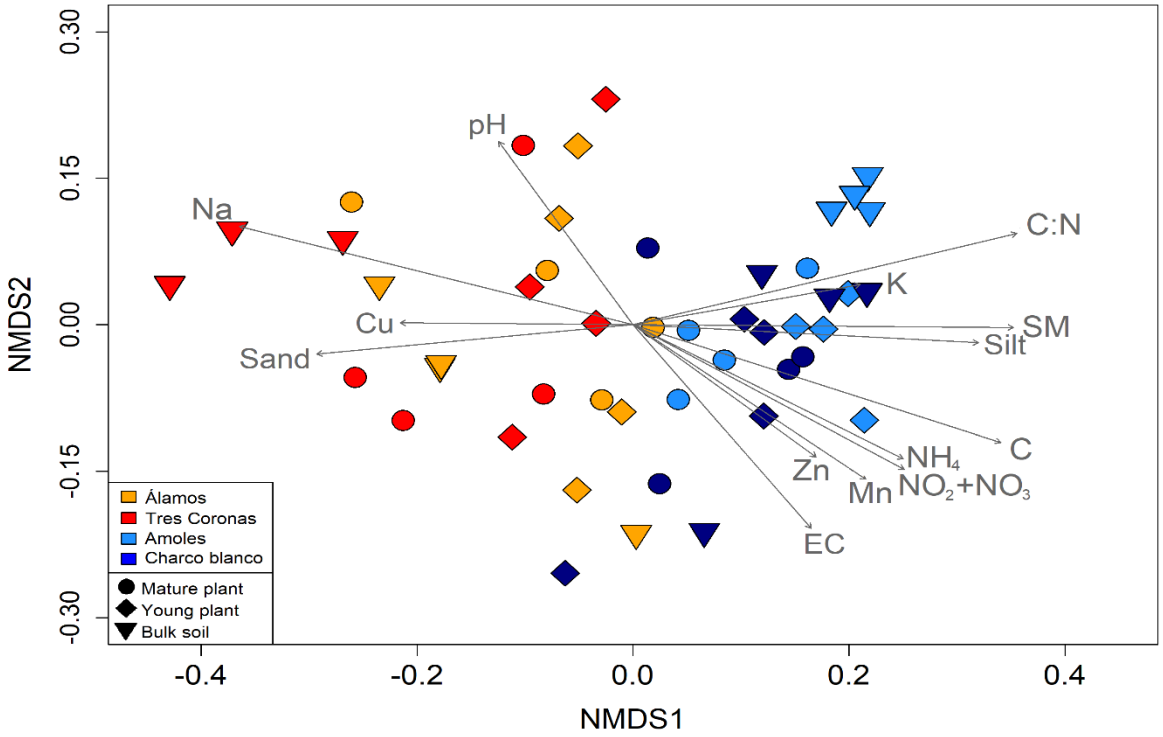



Figure 1.3.3. Non-metric multidimensional scaling (NMDS) and environmental fit. NMDS was performed with the relative abundance of bacterial genera identified in soils from Chihuahuan Desert Central and Meseta subregions (Bray Curtis dissimilarity, Stress = 0.1561). Only significant physicochemical soil properties are displayed as vectors ($p<0.01$), arrow size is proportional to correlation with ordination axes.

In order to explore the composition of the bacterial communities and their similarities between samples, a non-metric multidimensional scaling (NMDS) with the environmental fit was performed (Figure 1.3.3). We could identify high similarity between bacterial communities from the same subregion (ANOSIM R=0.033, $p<0.05$); samples from Álamos and Tres Coronas were clustered and related to higher concentration of Na, Cu, Sand, and pH. In contrast, higher availability of nutrients such as $NO_2$+$NO_3$, $NH_4$, and TOC was associated with the bacterial communities from Los Amoles and Charco Blanco in the Meseta subregion (Stress= 0.1561, $p<0.01$).

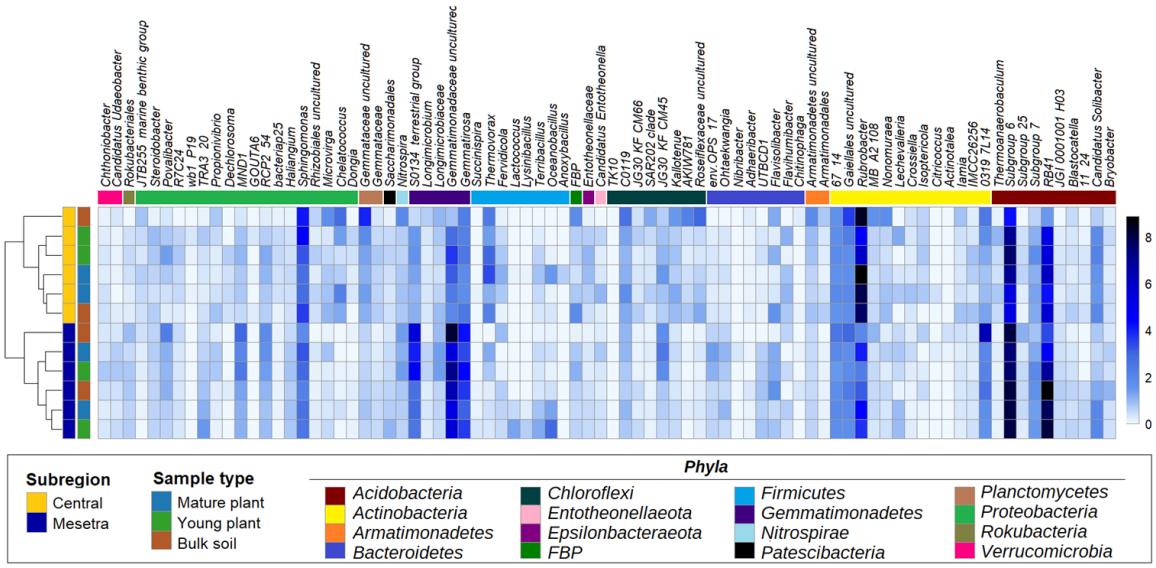



Figure 1.3.4. Bacterial community composition of soil samples from the Chihuahuan Desert. Heatmap based on the relative abundance of the genera. Clustering is based on Bray Curtis distance (Tree on the left). Color bars on the left denote subregion (Central or Meseta) and sample type (bulk soil, rhizosphere of mature or young plants). Phyla are shown on the horizontal color bar.

### 1.3.3 Differences in the rhizobacterial community of young and mature lechuguillas are subtle and within Actinobacteria

In general, the bacterial community found in the samples was composed of 16 phyla above 1% of relative abundance (Figure 1.3.4). The most abundant phyla were Acidobacteria, Actinobacteria, Firmicutes, Chloroflexi, Gemmatimonadetes, Planctomycetes, and Proteobacteria. However, each subregion had a characteristic composition at deeper taxonomic levels. For example, at the genera level, in the

Central subregion *Rubrobacter* (Actinobacteria), *Thermovorax* (Firmicutes), *Chelatococcus*, and *Sphingomonas* (Proteobacteria) showed higher relative abundance (Kruskal p<0.05). Meanwhile, *Subgroup 6* (Acidobacteria), *S0134 terrestrial group*, *Gemmatirosa*, and *uncultured Gemmatimonadaceae* (Gemmatimonadetes) were more abundant in the Meseta (Kruskal p<0.05) (Figure 1.3.4 and 1.3.5A).

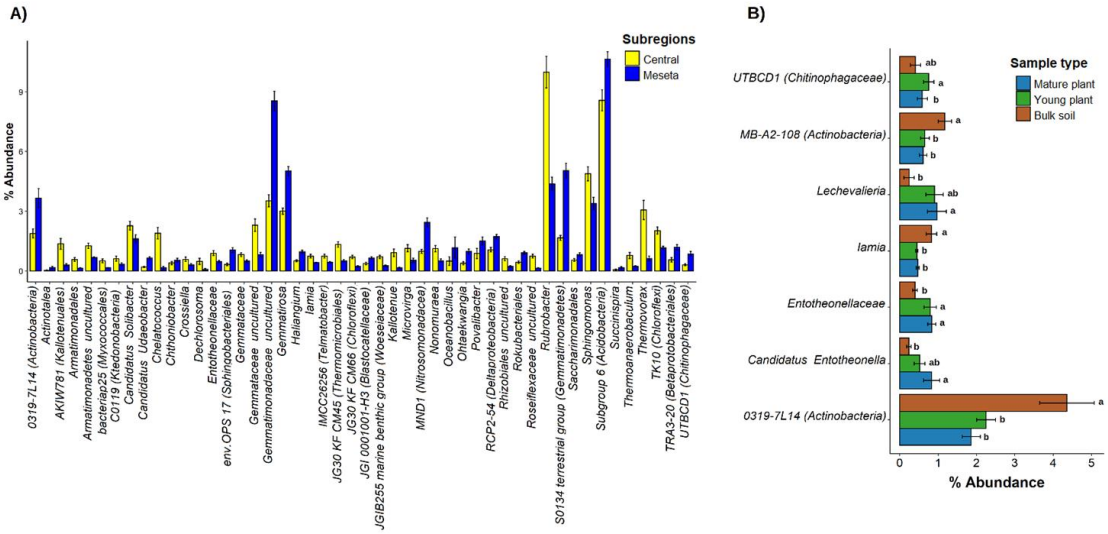
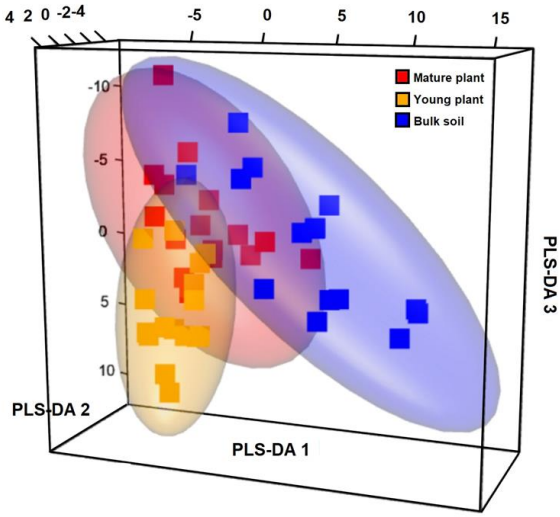



Figure 1.3.5.Differences between the relative abundance of bacterial genera. A) Central and Meseta subregions. B) Sample type: bulk and rhizosphere from mature or young plants of *A. lechuguilla* T. Kruskal-Wallis and post Hoc Dunn tests were performed. Different letters indicate difference between the relative abundance of the genera (p<0.05).

According to the sparse Partial Least Discriminant Analysis (sPLS-DA), the discrimination between bacterial communities of bulk soil and rhizosphere from mature and young plants is better observed in three dimensions (accuracy = 55.53%). The first component showed discrimination based on plant presence (ROC = 0.81, p<0.001) (Figure 1.3.6). For example, bulk soil had a higher relative abundance for the genera: *0319-7L14, Iamia*, and *MB-A2-108* (Actinobacteria); this was corroborated with the Kruskal test and *post hoc* Dunn (p<0.05) (Figure 1.3.5).

Only *Rubrobacter* (Actinobacteria) had a significant difference in relative abundance between young and mature plants (Figure 1.3.7A). Even though the discrimination by growth stages showed a gradual increase in the relative abundance of genera

from bulk soil to the rhizosphere of mature plants, in *Deinococcus* (Deferribacteres) and *Ilumatobacter* (Actinobacteria), as well as a reduction in the relative abundance of the taxa in the rhizosphere of mature and young plants, like *Sinosphorangium*, *0319-7L14* (Actinobacteria), *Symbiobacterium* (Firmicutes), *661256* (Nitrospirae), and *Limnobacter* (Proteobacteria) (Figure 1.3.6).




| Phyla | Genera | comp1 | comp2 | comp3 | MP | YP | S |
|---|---|---|---|---|---|---|---|
| *Acidobacteria* | *Vicinamibacter* | 4.53 | 3.31 | 3.01 | | | |
| | *Acidipila* | 2.48 | 3.26 | 2.96 | | | |
| *Actinobacteria* | *0319 7L14* | 5.86 | 4.29 | 3.90 | | | |
| | *Ilumatobacter* | 5.79 | 4.23 | 3.85 | | | |
| | *Microtrichaceae uncultured* | 5.38 | 3.93 | 3.57 | | | |
| | *Iamia* | 5.29 | 3.87 | 3.52 | | | |
| | *Euzebya* | 4.41 | 3.22 | 2.93 | | | |
| | *Micromonospora* | 4.27 | 3.12 | 2.84 | | | |
| | *MB A2 108* | 4.23 | 3.09 | 2.81 | | | |
| | *Sinosporangium* | 4.01 | 2.94 | 2.67 | | | |
| | *Coriobacteriales uncultured* | 0.00 | 7.78 | 7.07 | | | |
| | *Rubrobacter* | 0.00 | 3.61 | 3.28 | | | |
| *Chlamydiae* | *Parachlamydiaceae unculture* | 0.00 | 10.00 | 9.09 | | | |
| *Chloroflexi* | *Chloronema* | 5.84 | 4.27 | 3.88 | | | |
| *Deferribacteres* | *Deinococcus* | 4.51 | 3.29 | 2.99 | | | |
| *Firmicutes* | *Aquibacillus* | 4.14 | 3.03 | 2.75 | | | |
| | *Symbiobacterium* | 4.10 | 3.00 | 2.90 | | | |
| | *M55 D21* | 3.79 | 2.77 | 2.52 | | | |
| *Nitrospirae* | *661256* | 0.00 | 6.71 | 6.10 | | | |
| *Proteobacteria* | *Polyangiaceae uncultured* | 6.20 | 4.53 | 4.12 | | | |
| | *Reyranella* | 5.73 | 4.19 | 3.80 | | | |
| | *Limnobacter* | 5.10 | 3.73 | 3.39 | | | |
| | *Stakelama* | 4.69 | 3.43 | 3.12 | | | |
| | *Rubellimicrobium* | 4.49 | 3.28 | 2.98 | | | |
| | *Hephaestia* | 2.25 | 3.30 | 3.00 | | | |
| | *Hydrogenophilaceae* | 0.00 | 8.98 | 8.16 | | | |
| | *Spirillum* | 0.00 | 6.80 | 6.18 | | | |
| | *Bradymonadales* | 0.00 | 3.99 | 3.63 | | | |
| *WS2* | *WS2* | 4.21 | 3.08 | 2.80 | | | |



Figure 1.3.6. Sparse partial least squares – discriminant analysis (sPLS-DA) and variable importance in projection (VIP). The analysis was performed with the relative abundance of bacterial genera identified in bulk soil (S) and rhizosphere of young (YP) and mature (MP) plants of *A. lechuguilla* T. from Chihuahuan desert (Accuracy = 55.32%, ROC component 1 = 0.81 with $p<0.05$). Plot of the first three components (left) and VIP scores from the top 20 most discriminant genera (right). The color bar reflects the mean of the relative abundance of taxa on each class.

Interestingly, the rhizobacterial composition differed between subregions. In the rhizosphere of *A. lechuguilla* T. from the Central subregion, we found a higher abundance of the genera *Subgroup 6, Subgroup 25* (Acidobacteria), *Entotheonellaceae* (Epsilonbacteraeota), *Dongia*, *Microvirga*, *MND1*, and *R7C24* (Proteobacteria) (Figure 1.3.7A). In comparison, *Citrococcus* (Actinobacteria), *JG30 KF CM66* (Choloroflexi), *Candidatus Entotheonella* (Entotheonellaeota), *Longimicrobium* (Gemmatimonadetes), and *Thermovorax* (Firmicutes) were more abundant in the Meseta subregion (Figure 1.3.7B). In this region, *Rubrobacter* has a significant difference in its abundance between mature (5.68%) and young agaves (4.02%) (Figure 1.3.7B).

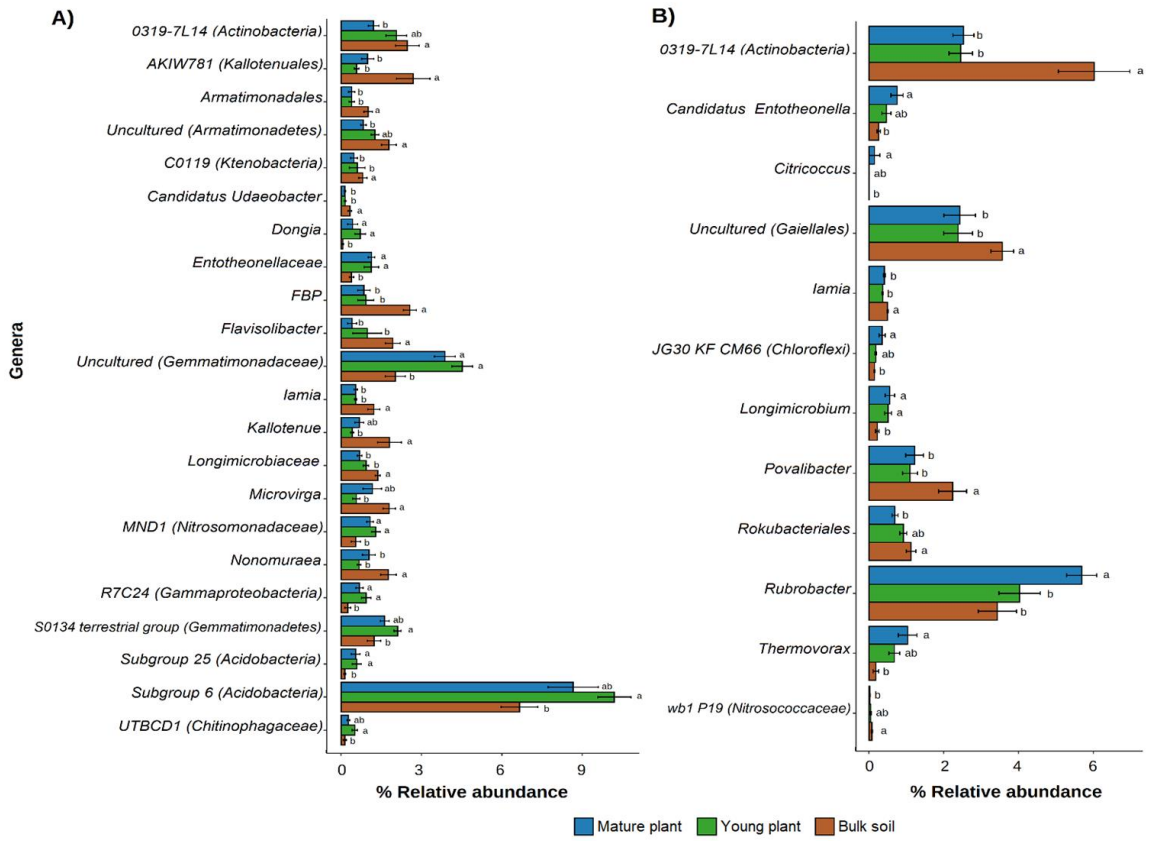



Figure 1.3.7. Differences in the relative abundance of bacterial genera identified in bulk soil and rhizosphere from mature and young plants of *A. lechuguilla* T., from Central (A) and Meseta (B) subregions. Kruskal-Wallis and post Hoc Dunn tests were performed. Different letters indicate difference between the relative abundance of the genera ($p<0.05$).

### 1.3.4 Ecological interaction networks change according to the subregion and plant growth stage

Co-occurrence networks and Minimum Spanning Trees (MST) were constructed to analyze the differences in the ecological interaction patterns between conditions (Figure 1.3.8). The networks presented a module-like structure (Table 0-4 and 0-5), where the nodes form clusters or communities tightly connected. Another characteristic of these networks was a scale-free degree distribution, verified with linear regression ($R^2 > 0.7$) (Table 0-4 and 0-5). This distribution shows that only a few nodes are highly connected within the network. Both characteristics suggested that these co-occurrence networks are models that depict real interactions in highly organized communities, where only some taxa are the critical components of the network.

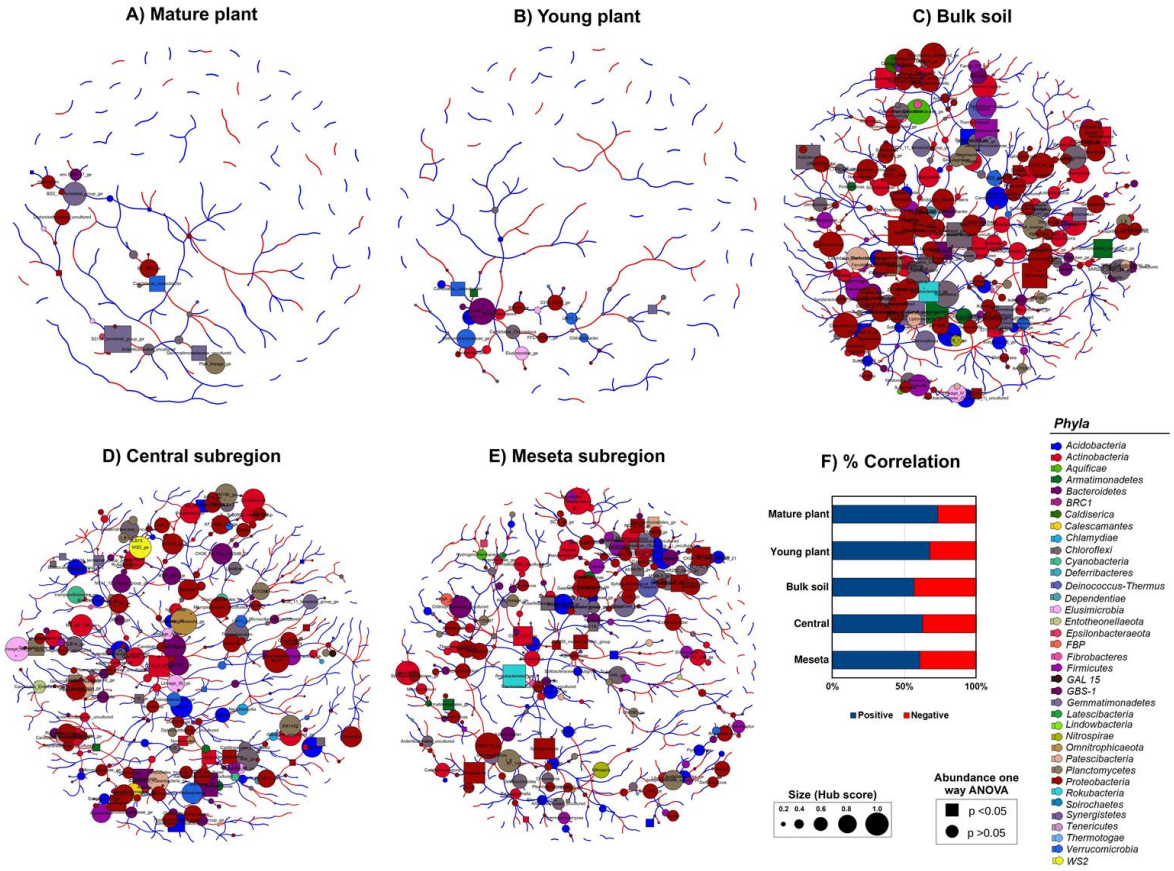



Figure 1.3.8. Minimal spanning trees from networks of bacterial communities associated to the rhizosphere of *A. lechuguilla* T. and bulk soil from Central and Meseta subregions of the Chihuahuan desert. Nodes represent bacterial genera and color edges display significant Spearman correlation between the genera ($\rho > 0.7$, FDR <0.05), positive in blue and negative in red. The size of the nodes depicts the hub score (log) and the colors the phyla. The square nodes are those which had different abundances between the three types of networks (A-C) or the two subregions (D, E) (Kruskal-Wallis, $p<0.05$) (See Figure 1.3.5 and 1.3.7). F) Percentage of positive and negative correlations in the networks.

Community networks from Central and Meseta subregions (n = 23 and 24, respectively) showed similarities in their structure (Figure 1.3.8D-E). For example, the former network was composed of 868 nodes with a mean of connections per node (degree) of 29.26 ± 1.20; meanwhile, the latter had 785 nodes with a mean degree of 16.88 ± 0.73 (Table 0-4). The network obtained from bulk soil (n=15) had nearly the same number of nodes (869), with a mean degree of 39.76. In contrast, the rhizosphere networks have a small size, composed of 238 and 248 nodes, in young and mature plants, respectively. The number of interactions was 2.07 ± 0.11 and 2.26 ± 0.53, and these rhizobacterial networks had a modularity value of 0.7 to

0.8, which shows a high number of modules with a small number of interactions (Figure 1.3.8D-E, Table 0-5).

Important taxa in the networks (hubs) could represent “key species” for the community structure. These strongly interconnected taxa were identified by a hub score above the mean of the hub values (0.5), a value based on the number of interactions between nodes. In the Central subregion network, we found hubs, amid genera *Conexibacter*, *Iamia*, *Uncultured Gaiellales, 0319 7L14* (Actinobacteria), *Thermobaculum* (Chloroflexi), *Thermovorax* (Firmicutes), *Blyi10*, *R7C24*, *Rhodovibrio* (Proteobacteria), and the phylum Omnitrophicaeota (Table 0-4). As for the Meseta subregion, we identified hubs in the genera *Atopobium*, *Paraerskovia* (Actinobacteria), *Uncultured Gemmatimonadaceae*, *S1034 terrestrial group, Gemmatimonas* (Gemmatimonadetes), *MBNT15*, *Caulobacteraceae*, *Povalibacter*, and *Sphingomonas* (Proteobacteria). Most of them also show a difference in their relative abundance and are more abundant within the subregion (Figure 1.3.5A and 1.3.8 D-E).

In bulk soil, the hubs were mainly among the phyla Actinobacteria, Chloroflexi, and Proteobacteria, besides other phyla, while hubs from the rhizosphere of agave were just a few (Figure 1.3.8A-C). For instance, in the network of the young plants, the genera *env. OPS17* (Bacteroidetes), *Chthoniobacteraceae* (Verrucomicrobia), and *0319 6G20* (Proteobacteria) had the highest hub score (1.00, 0.70, and 0.59, respectively). While hubs in the networks of the mature plants were all included in the most extensive module composed of 46 nodes, belong to genera from phyla Gemmatimonadetes: *Gemmatimonadaceae uncultured, S0134 terrestrial group* and *BD2 11 terrestrial group* (HS = 0.56, 1.00, and 0.88, respectively), Planctomycetes (*Pla4 lineage*, HS = 0.74), and Proteobacteria (*B1 7BS, Diplorickettsiaceae uncultured*, and *Haliangium*; hubs scores of 0.69, 0.59, and 0.51, respectively) (Table 0-5). *Candidatus Udaeobacter* (Verrucomicrobia), genera present in both lechuguilla networks, showed a high number of interactions, despite its low abundance in the rhizosphere (Figure 1.3.7A).

It is noticeable that all networks had a higher proportion of positive correlations than negative ones (Figure 1.3.8F). In both subregions, the networks had around 60% positive correlations; however, the proportion among the taxa changed. Genera from the phylum Actinobacteria showed a substantial increase in the negative interactions in the Central subregion. For example, from phylum Actinobacteria: *0319-7L14, Atopobium, Gaiellales Uncultured*, and *MB A2 108* had over 70% of negative interactions while in the Meseta subregion, this percentage was below 45%. Positive interactions in the rhizosphere networks increased. For younger plants, positive interactions were 67.48% and 73.04% for mature agaves; however, bulk soil had 56.60% (Figure 1.3.8F, Table 1.3-2).

Table 1.3-2. Interactions in networks from *A. lechuguilla* T. in Central and Meseta subregions from the Chihuahuan Desert. Percentage of positive and negative interactions on each network and mean of edges between nodes, with standard error. Letters show the difference between networks according to T test ($p<0.05$), pair comparison was made separately for subregion networks and bulk and rhizosphere of agaves.

| | **Positive edges (out)** | | **Negative edges (out)** | |
|---|---|---|---|---|
| | **%** | **Edges per node** | **%** | **Edges per node** |
| **Central** | 62.59 | 18.31 ± 0.80 a | 37.41 | 10.94 ± 0.53 a |
| **Meseta** | 60.55 | 10.22 ± 0.44 b | 39.45 | 6.66 ± 0.35 b |
| **Mature plant** | 73.04 | 1.65 ± 0.10 b | 26.96 | 0.61 ± 0.06 b |
| **Young plant** | 67.48 | 1.39 ± 0.08 c | 32.52 | 0.67 ± 0.06 b |
| **Bulk soil** | 56.60 | 22.50 ± 0.82 a | 43.40 | 17.26 ± 0.77a |

### 1.3.5 *A. lechuguilla* T. promotes the activity and abundance of NFB and OPMB

In order to understand which environmental factors affect the enzymatic activity and abundance of NFB and OPMB, nitrogenase, ALP and the copy numbers of their respective genes (*nifH* or *phoD*), were quantified.

With the analysis of variance, it was found that the enzymatic activity of NFB and OPMB was higher in the rhizosphere of mature lechuguillas and that each group had different values according to the subregion ($p<0.001$). For example, in the Central

subregion, nitrogenase activity was 193.72 ± 15.42 and 181.84 ± 9.27 nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$ for mature and young plants. While in bulk soil, it was 152.19 ± 6.32 nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$. However, in the Meseta, the highest activity was 107.32 ± 7.05 nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$ for mature plants (Table 1.3-3). As for ALP activity, in the Meseta subregion, mature rhizosphere of agaves reported 447.65 ± 15.35 µg pNP $g^{-1}$ $h^{-1}$, 415.04 ± 16.42 µg pNP $g^{-1}$ $h^{-1}$ the young plants and bulk soil 371.31 ± 21.22 µg pNP $g^{-1}$ $h^{-1}$. However, ALP activity only reached 112.14 ± 21.08 µg pNP $g^{-1}$ $h^{-1}$ in mature plants of the Central subregion (Table 1.3-3).

Table 1.3-3. Enzymatic activity and abundance of nitrogen fixing bacteria and organic phosphorus mineralizing bacteria. Mean and standard error are displayed (n = 8). At the bottom of the table, result of the two-way ANOVA is shown (*** p <0.001, ** p <0.01, * p <0.05).

| Subregion | Sample type | Organic phosphorus mineralizing bacteria | | Nitrogen fixing bacteria | |
|---|---|---|---|---|---|
| | | µg pNP $g^{-1}$ $h^{-1}$ soil | log copies *phoD* $g^{-1}$ soil | nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$ | log copies *nifH* $g^{-1}$ soil |
| **Central** | **Mature plant** | 112.14 ± 21.08 | 8.69 ± 0.09 | 193.72 ± 15.42 | 6.31 ± 0.04 |
| | **Young plant** | 52.57 ± 8.69 | 8.48 ± 0.09 | 181.84 ± 9.27 | 5.78 ± 0.06 |
| | **Bulk soil** | 41.56 ± 5.06 | 8.18 ± 0.09 | 152.19 ± 6.32 | 5.71 ± 0.07 |
| **Meseta** | **Mature plant** | 447.65 ± 15.35 | 7.74 ± 0.41 | 107.32 ± 7.05 | 6.87 ± 0.29 |
| | **Young plant** | 415.07 ± 16.42 | 7.92 ± 0.20 | 78.14 ± 3.13 | 7.15 ± 0.20 |
| | **Bulk soil** | 371.31 ± 21.22 | 7.53 ± 0.19 | 77.33 ± 5.02 | 6.80 ± 0.06 |
| **Two-way ANOVA (F-value)** | | | | | |
| **Sample type** | | 11.017*** | 1.57 | 11.07*** | 3.450* |
| **Subregion** | | 703.70*** | 18.53*** | 209.18*** | 73.06*** |
| **Sample type*Subregion** | | 0.61 | 0.5 | 2.341 | 4.15* |

The gene copy numbers of these functional groups revealed that the Central subregion had a higher number of OPMB with *phoD* and the Meseta subregion had more *nifH* than the northern subregion (p<0.001). Besides, differences among soil sample types were only significant for *nifH* (p<0.05), and the presence of the plant (mature or young) seems to promote the number of NFB. For instance, OPMB with *phoD* from Central were above 8.18 ± 0.09 log copies *phoD* $g^{-1}$ $soil^{-1}$, and they only reached 7.92 ± 0.20 log copies *phoD* $g^{-1}$ $soil^{-1}$ in the Meseta. *nifH* gene in the Central subregion was 6.31 ± 0.04 log copies *nifH* $g^{-1}$ $soil^{-1}$ and 5.78 ± 0.06 log copies *nifH* $g^{-1}$ $soil^{-1}$ for mature and young plants, respectively, while bulk soil had 5.71 ± 0.07

log copies *nifH* $g^{-1}$ $soil^{-1}$. In contrast, young agaves' rhizosphere had 7.15 ± 0.20 log copies *nifH* $g^{-1}$ $soil^{-1}$, while mature plants and bulk soil were around 6.80 ± 0.06 log copies *nifH* $g^{-1}$ $soil^{-1}$.

Table 1.3-4. Linear regression models of potential enzyme activities and abundance of nitrogen fixing bacteria and organic phosphorus mineralizing bacteria, from bulk and rhizosphere of mature and young plants of *A. lechuguilla* T. Models were made with the backward stepwise selection method (*** p <0.001, ** p <0.01, * p <0.05). The variable order in the table matches the order in the model.

| | Variable | F value | Percentage contribution |
|---|---|---|---|
| **Organic phosphorus mineralizing bacteria** | | | |
| **Potential alkaline phosphomonoesterase activity** | Na | 1306.14*** | 88.68 |
| | C:N ratio | 6.341181* | 0.43 |
| | $NO_2$+$NO_3$ | 40.99*** | 2.78 |
| | Total Org. C | 52.42*** | 3.56 |
| | Mg | 10.51*** | 0.71 |
| | Sample type | 8.25*** | 1.12 |
| | | Adjusted $R^2$ = 0.9681, p <0.001 | |
| **Log *phoD*** | Total Org. C | 20.19*** | 27.76 |
| | Sample type | 4.27* | 11.74 |
| | | Adjusted $R^2$ = 0.3537, p <0.001 | |
| **Nitrogen fixing bacteria** | | | |
| **Potential nitrogenase activity** | C:N ratio | 141.52*** | 62.11 |
| | Total Org. C | 15.59*** | 6.84 |
| | Sample type | 5.92* | 5.20 |
| | Fe | 16.90*** | 7.42 |
| | | Adjusted $R^2$ = 0.793, p <0.001 | |
| **Log *nifH*** | Na | 34.87*** | 58.51 |
| | | Adjusted $R^2$ = 0.5761, p <0.001 | |

#### 1.3.5.1 Sodium and organic carbon concentration are limiting factors for NFB and OPMB activity and abundance

Linear models constructed with stepwise regression showed that the presence of the plant and seven soil properties were related to the enzymatic activity and abundance of NFB and OPMB (Table 1.3-4). In this sense, sodium concentration explained most of the variation in the ALP activity (88.68%). Moreover, TOC, C:N ratio, $NO_2$+$NO_3$,

and Mg positively affected the ALP activity and explained about 7.48% of the variation. As for NFB, C:N ratio was the main factor related to the nitrogenase activity (62.11%), followed by Fe concentration (7.42%), TOC (6.84%), and the presence of the plant (5.20%). The variation in the abundance of the corresponding functional genes was also mainly explained by sodium (58.51%) for *nifH* and TOC (27.76%) with sample type (11.74%) for *phoD* (Table 1.3-4). Interestingly, the main factors in all the models (sodium, TOC, and C:N ratio) negatively affect the enzymatic activity or abundance of NFB and OPMB (Figure 1.3.9).

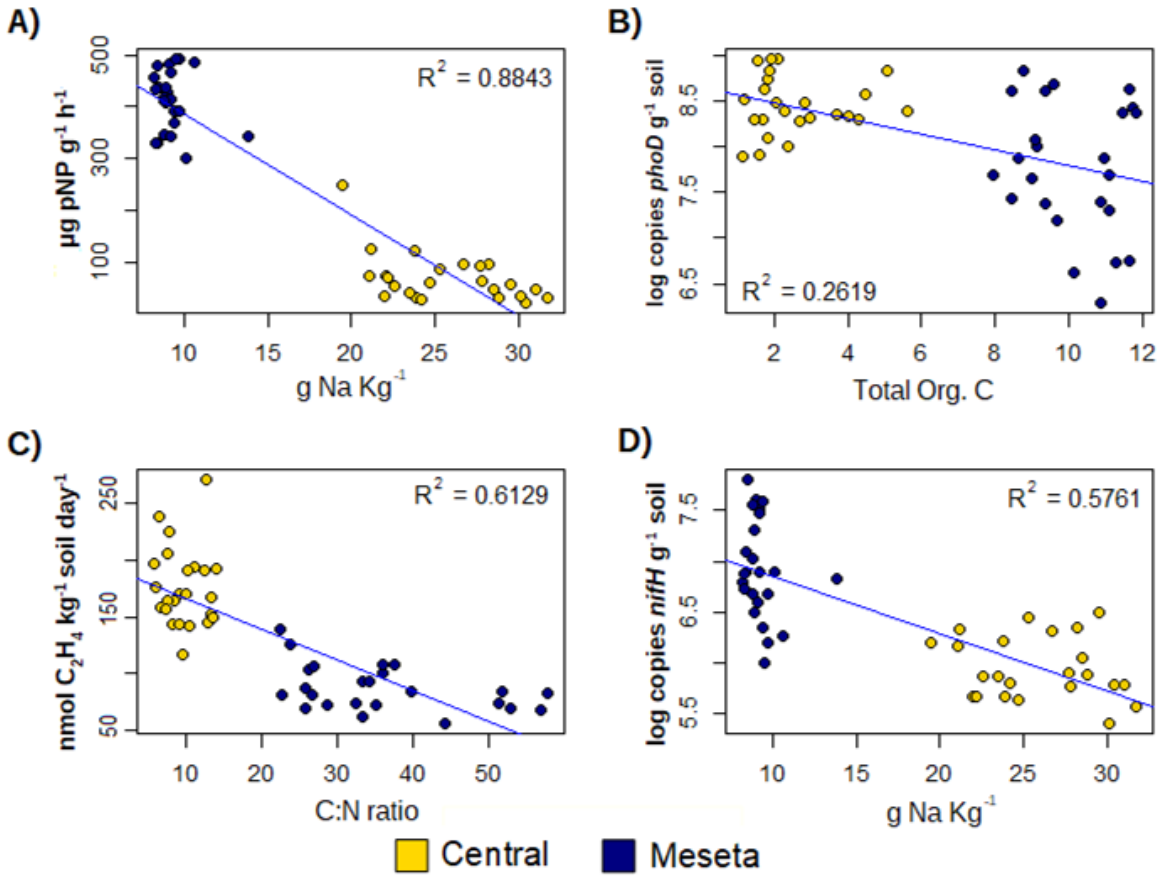



Figure 1.3.9. Linear regression between enzymatic activity and abundance of bacteria with soil properties from Central and Meseta subregions. A & B) Organic phosphate mineralizing bacteria. C & D) Nitrogen fixing bacteria.

### 1.3.5.2 Correlation between enzymatic activity and bacterial taxa

Correlation between genera over 1% of relative abundance and activity/abundance of NFB and OPMB was tested by the Spearman method. In the Central subregion,

positive correlations mainly were related to nitrogenase activity and *nifH* abundance (ρ ≥ 0.4, p<0.01), and these traits were verified by literature revision considering the possession of genes or the detected activity under laboratory conditions for *Bryobacter* (Acidobacteria), *Crossiella*, *Nonomuraea, Rubrobacter* (Actinobacteria), *AKIW781* (Chloroflexi), *Microvirga*, and *GOUTA6* (Proteobacteria) (Figure 1.3.10A). In the Meseta subregion, *Actinotalea*, *Lechevalieria* (Actinobacteria), and *GOUTA6* (Proteobacteria) were significantly correlated with OPMB attributes (ρ ≥ 0.4, p<0.01), while *Dechlorosoma* and *Haliangium* (Proteobacteria) (ρ ≥ 0.4, p<0.01), had been reported as both NFB and OPMB (Figure 1.3.10B).

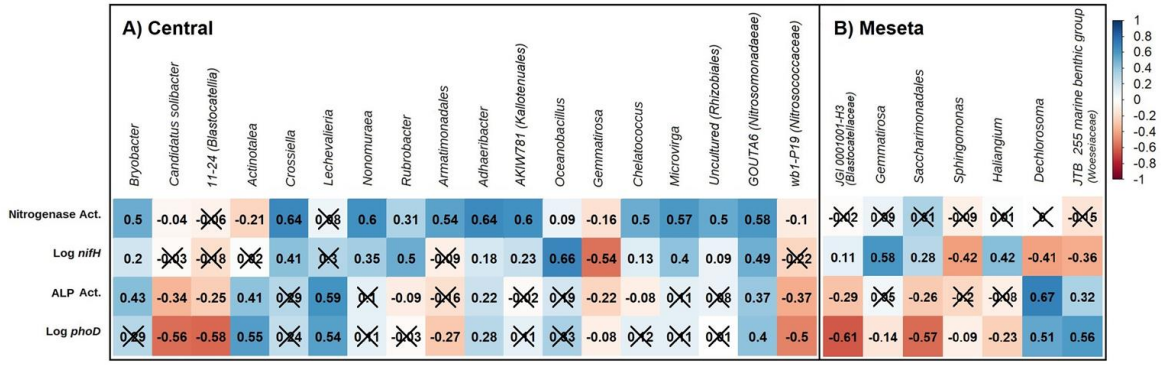



Figure 1.3.10.Correlogram between bacterial genera and enzymatic activity/abundance of nutrient-improvement bacteria. The bacterial groups include: Nitrogen fixing bacteria (Nitrogenase, *nifH*) and organic phosphate mineralizing bacteria (Alkaline phosphomonoesterase -ALP-, *phoD*). A) Central and C) Meseta subregions. Spearman method was applied. The number and color intensity on the correlogram correspond to rho value of significant correlations (p<0.01).

## 1.4 Discussion

The objective of this project was to understand how soil properties on a regional scale and plant growth stage influence the composition and interactions of bacterial communities, as well as over the enzymatic activity and abundance of NFB and OPMB. In this work, we found that both factors are important at different levels. While soil properties affected the composition of the bacterial communities, mature lechuguillas stimulated the activity and abundance of nutrient-improvement rhizobacteria. Besides, taxa associated with these functional groups had a crucial role in the rhizobacteria community, reflected in the high number of positive interactions within the bacterial networks. This interrelationship between abiotic conditions, host, and microbiome could explain, to some extent, the wide distribution of *A. lechuguilla* T. under the harsh conditions in the Chihuahuan Desert.

### 1.4.1 *A. lechuguilla* T. ameliorated the effect of climate change and parent material in their rhizosphere

The soils from the Central and Meseta subregions provide different environments, resulting from the climate conditions, topography, and parent material (Dinerstein *et al*. 2001; Augusto *et al*. 2017). The scarce rain and high evaporation rates in these subregions lead to the accumulation of salts that limit nutrient acquisition and productivity (Osman 2013), and this condition is exacerbated by the limestone soils from the Chihuahuan Desert (García-Arévalo 2002). Furthermore, according to the criteria from Gamalero *et al*. (2020), the soil from the Central subregion could reach the grade of “sodic”, which implies electrical conductivity <4 $dSm^{-1}$, pH>8.5, and the absorption of sodium by clay. This condition hinders the acquisition of nutrients and soil productivity (Granados-Sánchez *et al*. 2011), by hindering, in the long run, the vegetative development as consequence of the toxic effect of the hyperaccumulation of sodium cations in the cell (Gamalero *et al*. 2020). Nevertheless, *A. lechuguilla* T. has a high production rate for these environments, 0.38 kg $m^2$ $year^{-1}$ (Stewart 2015), and the resulting accumulation of organic matter is reflected on the C:N ratio. This relation rules the decomposition process and its exploitation by both plants and

bacteria (Brust 2019). Besides, the accumulation of organic matter from both lechuguilla development stages was reflected over lower C:N ratio levels. For example, the values for the Central subregion in young agaves (11.41 ± 0.83) are similar to the 15-20 reported in the northern subregions of the Chihuahuan desert in New Mexico (Hodge, Robinson and Fitter 2000). In contrast, the higher concentration of organic matter in the Meseta, derived in a C:N ratio of 29.50 in the mature lechuguillas, enough to enable normal nutrient incorporation (Moorhead and Reynolds 1991). Therefore, even with the high levels of radiation, extreme temperatures, and the low water availability distinctive of arid environments (Prăvălie 2016), lechuguilla could shape soil conditions for microorganisms able to overcome the limited availability of nutrients and high sodium concentration, which were also significantly harsher in Central Subregion.

### 1.4.2 Bacterial communities form the Chihuahuan Desert were supported by taxa tolerant to abiotic stress and nutrient intake versatility

The microorganisms with the highest abundance in Central and Meseta subregions were phyla Acidobacteria, Actinobacteria, Gemmatimonadetes, and Proteobacteria, common in Mexican arid lands, such as The Sonoran Desert, Cuatro Ciénegas Basin, and Tehuacán-Cuicatlán Valley (Nagy, Pérez and García-Pichel 2005; Torres-Cortés *et al*. 2012; López-Lozano *et al*. 2013). The wide distribution of these phyla has been related to abiotic stress tolerance and nutrient acquisition (DeBruyn *et al*. 2011). Remarkably, the most interconnected genera within the co-occurrence networks (hubs) from Central and Meseta subregions could also relate to these traits. For instance, in the Central subregion, *Iamia* and *Conexibacter* (Actinobacteria), some of the most abundant genera, had been associated with organic matter decomposition (Xia *et al*. 2016), can perform organic phosphorus mineralization or nitrogen fixation, and thus probably helping with the nutrient intake within the network (Karp *et al*. 2019). While, genera from Chloroflexi phylum: *Candidatus Chloroploca, Kouleothrix,* and *FFCH7168* can perform carbon fixation (Garrity *et al*. 2001; Grouzdev *et al*. 2018), a mechanism that may be key to tolerate the low

concentration of TOC in this subregion or *Thermobaculum*, a non-phototrophic thermotolerant bacteria with different phosphatases (Botero *et al*. 2004). Another example is the genera *Omnitrophicaeota* (Omnitrophicaeota) found in saline estuaries and is supposed to have a role in carbon and sulfur cycles (Baricz *et al*. 2019).

Interestingly, with higher nutrient availability than the Central subregion, the Meseta subregion supports a different bacterial community, where the most representative bacteria were the phyla Gemmatimonadetes and Proteobacteria. These taxa are usually found in soils rich in organic matter and have photosynthetic members able to produce their energy and substrates (Ullah *et al*. 2019); or are involved in carbon, nitrogen, and sulfur cycles (Spain Krumholz and Elshahed 2009). For example, the hubs *S0134 terrestrial group* and *BD2 11 terrestrial group* (Gemmatimonadetes) have been isolated from the rhizosphere, and some of them seem favor the mineralization of organic phosphorus (Karp *et al*. 2019). Among the hubs from phylum Proteobacteria, the genera *Sphingomonas* is tolerant to saline conditions and can incorporate heavy metals (Asaf *et al*. 2017), and genera *MBNT15* is a specialist with mechanisms for nitrate reduction and energy generation through acetate and hydrogen oxidation (Chen *et al*. 2020).

Some bacteria showed a difference in their behavior according to nutrient availability in each subregion. For example, genera *0319-7L14, Atopobium, Gaiellales uncultured,* and *MB A2 108* (Actinobacteria) had fewer negative interactions with other genera in the network of the Meseta subregion; while in the meager soils of the Central subregion, the negative correlations of the same taxa were almost twice as in the Meseta. A predator behavior has been recorded in some strains of the phylum Actinobacteria. For example, the genera *Streptomyces* produce antibacterial metabolites when cultivated in low nutrient media, suggesting that these bacteria can modulate their behavior to improve their survival (Ibrahimi *et al*. 2020). As we have seen, nutrient availability seems essential for the interaction network structure, although we do not rule out other important factors not analyzed in this work.

### 1.4.3 Beneficial bacteria were selected gradually during the development of *A. lechuguilla* T.

Along with abiotic conditions, plants can modify the bacterial community in their surrounding soil (Qiao *et al*. 2017; Naylor and Coleman-Derr 2018). A common effect in oligotrophic soils, known as *fertility islands*, is generated due to the nutrient and moisture accumulation around the plants, which can shelter different organisms from bulk soil exposed to harsher conditions (Garcia-Moya and McKell 2006). Nutrient accumulation in fertility islands could be derived from root exudates and expected to increase gradually during plant development. This effect has been reported in *Agave angustifolia* (Bautista-Cruz *et al*. 2007) and *A. lechuguilla* T. (López-Lozano *et al*. 2020).

In bulk soil, the most abundant genera were *Iamia, 0319-7L14,* and *MB-A2-108* (Actinobacteria), with tolerance to arid conditions (Barka *et al*. 2016). However, the rhizobacteria associated with lechuguilla could decompose organic matter or use carbon sources secreted by plants. For instance, genera *Lechevalieria* (Actinobacteria) has been isolated from hyper arid soils in the Atacama Desert and can use sugars and organic acids (Okoro *et al*. 2010); genera *Vicinamibacter* (Acidobacteria), *M55-D21*, and *Baillaceae* (Firmicutes) are susceptible to organic matter or root exudates and could be enhancing the plant nutrition through nitrogen fixation or phosphate solubilization, respectively (Mandic-Mulec, Stefanic and van Elsas 2015; Huber *et al*. 2016; Jin *et al*. 2020).

The growth stage of *A. lechuguilla* T. was another factor in selecting the bacterial community. sPLS-DA revealed *Ilumatobacter* (Actinobacteria) and *0319-7L14* (Proteobacteria) as genera related to mature agaves (Shange *et al*. 2012). Moreover, this recruitment could favor the host as these taxa are known for their ability to produce phosphatases (Matsumoto *et al*. 2009).

Besides, the data suggested that not only the bacterial composition change during plant development, but also the interactions within the co-occurrence networks, which can be explained by the selection of specialized bacteria with the capacity to support the host in exchange for shelter and nutrients provided by plants through the

metabolites released to the rhizosphere (Granados-Sánchez *et al*. 2011). The hubs found in the young plant network, such as *Chthoniobacteraceae*, and *Candidatus Udaeobacter* (Verrucomicrobia), can be related to nutrients availability and antibiotics resistance (Willms *et al*. 2020). *Thermoanaerobaculum* (Acidobacteria), another hub in the network of the young lechuguillas, is associated with sulfur cycles and produces siderophores; this last one, a mechanism for biological control and Fe acquisition (Tschoeke *et al*. 2020), which could explain the high proportion of negative interactions with other bacteria. Instead, the principal hubs from the mature lechuguilla network were *among Gemmatimonadaceae Uncultured, S0134 terrestrial group,* and *BD211 terrestrial group* (Gemmatimonadetes), which have been either reported as OPMB or NFB (DeBruyn *et al*. 2011). These mechanisms are energetically expensive for the microorganisms and require high amounts of carbon that mature agaves can fulfill by root exudates since they do not require the same levels of carbon for their growth as young plants (Hildebrandt *et al*. 2015). *Cupriavidus* (Proteobacteria), genera only present in the mature plant network, has been isolated from alkaline soils in the northwest of Mexico, near the Meseta subregion; one of its attributes is the ability to produce biofilms, siderophores, and bactericides (Rojas-Rojas *et al*. 2016). These traits could be improving the protection of the host against pathogens.

As we have seen, plant growth promotion mechanisms were features from the bacterial hubs in the networks. However, OPMB and NFB were the most common and were among the hubs from lechuguilla networks. This potential favors the establishment of mutualistic relationships between plant-bacteria and bacteria-bacteria (Omirou, Fasoula, and Ioannides 2016), which was corroborated by the increase of positive interactions inside the communities of young and mature agave, in contrast with bulk soil. The presence and age of lechuguilla affected the soil properties, which, in turn, were closely related to the activity and abundance of OPMB and NFB. For OPMB, Na concentration was the main limiting factor for the alkaline phosphomonoesterase activity. It is known that high concentrations of sodium compounds such as $NaNO_3$, $NaClO_4$, and NaSCN, have an inhibitory effect

(Pandey and Banik 2011). Likewise, as an enzyme that requires organic matter as substrate, it is understandable that TOC was a limiting factor for ALP activity, which explains the lower activity in the sodic and poor soils from the Central subregion. In contrast, the higher levels of organic matter from the Meseta enable the incorporation of phosphorus into the recycling process by microorganisms with calcium-based phosphatases, codified in genes like *phoX* and *phoD* (Ragot *et al*. 2017). Another element important for ALP activity was Mg, which can substitute the $Fe^{2+}$ cofactor (Gomez and Ingram 1995).

The correlation between bacteria genera and ALP activity and *phoD* abundance was tested ($p<0.01$). For instance, *SAR 202 clade* (Chloroflexi), *Adhaeribacter* (Bacteroidetes), and *Dechlorosoma* (Proteobacteria) were correlated with OPMB activity though the *phoD* gene has not been reported in their genome, which indicates that these genera may use different phosphatases (Karp *et al*. 2019).

Likewise, sodium was also a limiting factor for nitrogenase activity and *nifH* gene abundance. Together, sodium concentration and alkaline conditions are known to have a negative effect on the communities due to the decrease in carbon intake efficiency (Keshri *et al*. 2015). Nevertheless, in the harsher conditions of the Central subregion, nitrogenase activity was higher than in the Meseta. The low availability of nitrogen in the soil may encourage the microorganisms to obtain this nutrient through nitrogenase activity (Bahuguna and Pal 2011). Furthermore, the cost of the enzymatic process could be compensated in the rhizosphere by the root exudates, even more in the case of mature agaves, which may provide the bacteria with the carbon source needed for the reaction (Anandyawati *et al*. 2017). While in the Meseta subregion, the nitrogenase activity may be affected by the higher concentrations of ammonium from the mineralization of the organic matter, which negatively affects nitrogenase activity (Zhang *et al*. 2019). Besides, the relation between nitrogenase activity and Fe may be explained by its function as a cofactor for nitrogen fixation (Inglett *et al*. 2009).

Remarkably, the main part of the taxa correlated with nitrogenase activity or abundance of *nifH* were from the Central subregion, where the highest activity was

recorded. For example, the genera *Bryobacter, Candidatus solibacter* (Acidobacteria), *Succinispira* (Firmicutes), *Crossiella*, *Nonomuraea* (Actinobacteria), *Microvirga*, *Rhizobiales Uncultured*, and *Haliangium* (Proteobacteria) had reports either about nitrogen-fixing activity, the presence of the *nifH* gene, or both (Karp *et al*. 2019). Moreover, for the genera with significant correlations but no records of either activity, such as *AKIW781* (Chloroflexi), *GOUTA*, *wb1-P19*, and *JTB 255 marine benthic group* (Proteobacteria), the high correlation could indicate that these taxa may have relevance in arid soils, and further research is needed on uncultured taxa described only by sequencing.

## 1.5 Conclusions

We have explored the composition of the bacterial communities associated with the rhizosphere of *A. lechuguilla* T. Although this plant has a high tolerance to the harsh conditions of the Chihuahuan Desert, the cooperation with PGPR, such as NFB and OPMB, may contribute substantially to its vast establishment. Along with their capacity to improve the available nutrients, mechanisms of tolerance to arid conditions, salinity and alkaline soils were among the most abundant and interconnected hubs in the bacterial networks. These findings led us to infer the importance of these microorganisms for both the host and the bacterial communities (Figure 1.5.1).

The growth stage of *A. lechuguilla* T. was also a significant factor shaping the bacterial microbiome. Even though the rhizobacterial community in young lechuguillas shared similarities with bulk soil, these plants, like mature ones, showed already a selection that favored NFB and OPMB.

In addition, abiotic conditions are relevant for the composition of bacterial communities. Different climate conditions and parent material are reflected in the availability of nutrient concentration, and some of them, like sodium and micronutrients, affect the activity and abundance of bacterial communities.

These findings bring insights and new questions about lechuguilla mechanisms for the selection of the bacterial communities and the identity of the NFB and OPMB that support its growth under the arid conditions of the Chihuahuan Desert.

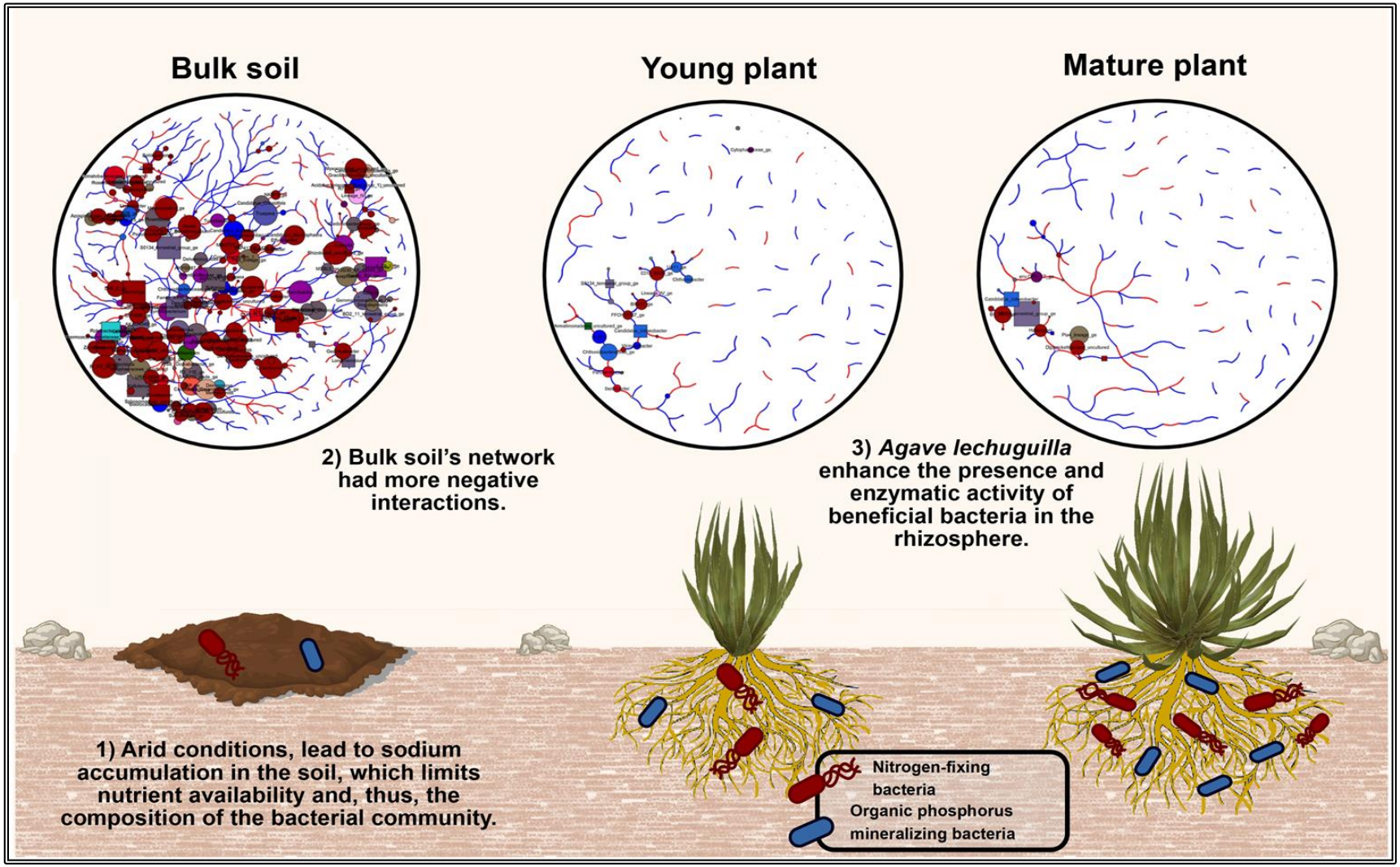



Figure 1.5.1. Bacterial communities and *A. lechuguilla* T. interactions. Plant growth promoting rhizobacteria are selected over the growth of *A. lechuguilla* T., favoring those taxa tolerant to the arid conditions of the Chihuahuan desert and with mechanisms for nutrient improvement.

## 1.6 References


Anandyawati, Anwar S, Nugroho B *et al*. Study of root exudate organic acids and microbial population in the rhizosphere of oil palm seedling. J Trop Soils 2017;22:29–36. DOI:10.5400/jts.2017.22.1.29.

Asaf S, Khan MA, Khan AL *et al*. Bacterial endophytes from arid land plants regulate endogenous hormone content and promote growth in crop plants: an example of *Sphingomonas* sp. and *Serratia marcescens*. J Plant Interact 2017;12:31–8. DOI:10.1080/17429145.2016.1274060.

Augusto L, Achat DL, Jonard M *et al*. Soil parent material—A major driver of plant nutrient limitations in terrestrial ecosystems. Glob Chang Biol 2017;23:3808–24. DOI:10.1111/gcb.13691

Bahuguna RN, Pal M. Physiology of nitrogen fixation in legumes under elevated $CO_2$. In: Singh MP, Khetarpal S, Pandey R, *et al*. (eds.). Climate Change: Impacts and Adaptations in Crop Plants. Vol 2. New Delhi: Today and Tomorrow, 2011, 213–28.

Baricz A, Chiriac C, Bulzu P-A *et al*. Patterns of prokaryotic diversity in sediments of brackish to hypersaline salt lakes in Romania. Stud Univ Babes-Bolyai, Biol 2019;64:96.

Barka EA, Vatsa P, Sanchez L *et al*. Taxonomy, physiology, and natural products of Actinobacteria. Microbiol Mol Biol Rev 2016;80:1–43. DOI:10.1128/MMBR.00019-15.

Barton LL, Northup DE. Microbe-microbe interactions. Microbial Ecology. 1st ed. John wiley & Sons, Inc., 2011, 159–82.

Bautista-Cruz A, Carrillo-González R, Arnaud-Viñas MR *et al*. Soil fertility properties on Agave angustifolia Haw. plantations. Soil Tillage Res 2007;96:342–9. DOI:10.1016/J.STILL.2007.08.001.

Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Ser B 1995;57:289–300. DOI:10.1111/j.2517-6161.1995.tb02031.x.

Botero LM, Brown KB, Brumefield S *et al*. *Thermobaculum terrenum* gen. nov., sp. nov.: a non-phototrophic gram-positive thermophile representing an environmental clone group related to the Chloroflexi (green non-sulfur bacteria) and Thermomicrobia. Arch Microbiol 2004;181:269–77. DOI:10.1007/s00203-004-0647-7.

Bray RH, Kurtz LZ. Determination of total, organic and available forms of phosphorus in soils. Soil Sci 1945;59:39–45. DOI:10.1097/00010694-194501000-00006.

Broido AD, Clauset A. Scale-free networks are rare. Nat Commun 2019;10:1017. DOI:10.1038/s41467-019-08746-5.

Brust GE. Management strategies for organic vegetable fertility. Safety and Practice for Organic Food. Elsevier, 2019, 193–212. DOI:10.1016/B978-0-12-812060-6.00009-X.

Castillo-Quiroz D, Cano-Pineda A, Berlanga-Reyes CA. Establecimiento y Aprovechamiento de La Lechuguilla (*Agave lechuguilla* Torr.). 1st ed. México: Comisión Nacional Forestal-Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, 2012.

Chen Y-J, Leung PM, Wood JL *et al*. Metabolic flexibility allows bacterial habitat generalists to become dominant in a frequently disturbed ecosystem. ISME J 2020, DOI:10.1038/s41396-021-00988-w.

Coleman-Derr D, Desgarennes D, Fonseca-Garcia C *et al*. Plant compartment and biogeography affect microbiome composition in cultivated and native Agave species. New Phytol 2016;209:798–811. DOI:10.1007/s13205-018-1317-9.

Crain G, McLaren J, Brunner B *et al*. Biologically available phosphorus in biocrust-dominated soils of the Chihuahuan Desert. Soil Syst 2018;2:56. DOI:10.3390/soilsystems2040056.

Csárdi G, Nepusz T. The igraph software package for complex network research. Interjournal 2006;Complex Sy:9.

DeBruyn JM, Nixon LT, Fawaz MN *et al*. Global biogeography and quantitative seasonal dynamics of Gemmatimonadetes in Soil. Appl Environ Microbiol 2011;77:6295–300. DOI:10.1128/AEM.05005-11.

Dechassa N, Schenk MK. Exudation of organic anions by roots of cabbage, carrot, and potato as influenced by environmental factors and plant age. J Plant Nutr Soil Sci 2004;167:623–9. DOI:10.1002/jpln.200420424.

Dinerstein E, Olson D, Atchley J *et al*. eds. Ecoregion-based conservation in the Chihuahuan Desert. A Biological Assessment. 2nd ed. CONABIO-The nature Coservancy-PRONATURA-TEC, 2001.

Freeman CE, Reid WH. Aspects of the reproductive biology of *Agave lechuguilla* Torr. Desert Plants 1985;7:75–80.

Gamalero E, Bona E, Todeschini V *et al*. Saline and arid soils: Impact on bacteria, plants, and their interaction. Biology (Basel) 2020;9:116. DOI:10.3390/biology9060116.

Garcia J, Kao-Kniffin J. Can dynamic network modelling be used to identify adaptive microbiomes?. Funct Ecol 2019:1–10. DOI:10.1111/1365-2435.13491.

García-Arévalo A. Vascular plants of the Mapimí biosphere reserve, México: A checklist. BRIT Press 2002;20:797–807.

Garcia-Moya E, McKell CM. Contribution of shrubs to the nitrogen economy of a desert-wash plant community. Ecology 2006;51:81–8. DOI:10.2307/1933601.

García-Robledo E, Corzo A, Papaspyrou S. A fast and direct spectrophotometric method for the sequential determination of nitrate and nitrite at low concentrations in small volumes. Mar Chem 2014;162:30–6. DOI:10.1016/j.marchem.2014.03.002.

Garrity GM, Holt JG, Castenholz RW *et al*. Phylum BVI. Chloroflexi phy. nov. Bergey's Manual® of Systematic Bacteriology. New York, NY: Springer New York, 2001, 427–46.

Gomez PF, Ingram LO. Cloning, sequencing, and characterization of the intracellular invertase gene from Zymomonas mobilis. FEMS Microbiol Lett 1995;125:237–46. DOI:10.1080/00021369.1991.10870787.

Granados-Sánchez D, Sánchez-González A, Granados-Victorino RL *et al*. Ecología de la vegetación del Desierto Chihuahuense. Rev Chapingo Ser Ciencias For y del Ambient 2011;XVII:111–30. DOI:10.5154/r.rchscfa.2010.10.102.

Grouzdev DS, Rysina MS, Bryantseva IA *et al*. Draft genome sequences of '*Candidatus Chloroploca* asiatica' and '*Candidatus Viridilinea mediisalina'*, candidate representatives of the Chloroflexales order: phylogenetic and taxonomic implications. Stand Genomic Sci 2018;13:24. DOI:10.1186/s40793-018-0329-8.

Hennion N, Durand M, Vriet C *et al*. Sugars en route to the roots. Transport, metabolism and storage within plant roots and towards microorganisms of the rhizosphere. Physiol Plant 2019;165:44–57. DOI:10.1111/ppl.12751.

Hildebrandt TM, Nunes Nesi A, Araújo WL *et al*. Amino acid catabolism in plants. Mol Plant 2015;8:1563–79. DOI:10.1016/j.molp.2015.09.005.

Hodge A, Robinson D, Fitter A. Are microorganisms more effective than plants at competing for nitrogen? Trends Plant Sci 2000;5:304–8. DOI:10.1016/S1360-1385(00)01656-3.

Hoffman M, Steinley D, Gates KM *et al*. Detecting clusters/communities in social networks. Multivariate Behav Res 2018;53:57–73. DOI:10.1080/00273171.2017.1391682.

Huber KJ, Geppert AM, Wanner G *et al*. The first representative of the globally widespread subdivision 6 Acidobacteria, *Vicinamibacter silvestris* gen. nov., sp. nov., isolated from

subtropical savannah soil. Int J Syst Evol Microbiol 2016;66:2971–9.DOI:10.1099/ijsem.0.001131.

Ibrahimi M, Korichi W, Hafidi M *et al*. Marine Actinobacteria: Screening for predation leads to the discovery of potential new drugs against multidrug-resistant bacteria. Antibiot (Basel, Switzerland) 2020;9, DOI:10.3390/antibiotics9020091.

Inglett PW, D'Angelo EM, Reddy KR *et al*. Periphyton nitrogenase activity as an indicator of wetland eutrophication: spatial patterns and response to phosphorus dosing in a northern Everglades ecosystem. Wetl Ecol Manag 2009;17:131–44. DOI:10.1007/s11273-008-9095-5.

Jin Z, Ci M, Yang W *et al*. Sulfate reduction behavior in the leachate saturated zone of landfill sites. Sci Total Environ 2020;730:138946. DOI:10.1016/j.scitotenv.2020.138946.

Judd K, Small M, Stemler T. What exactly are the properties of scale-free and other networks? EPL (Europhysics Lett 2013;103:58004. DOI:10.1209/0295-5075/103/58004.

Karp PD, Billington R, Caspi R *et al*. The BioCyc collection of microbial genomes and metabolic pathways. Brief Bioinform 2019;20:1085–93. DOI:10.1093/bib/bbx085.

Keshri J, Yousuf B, Mishra A *et al*. The abundance of functional genes, *cbbL*, *nifH*, *amoA* and *apsA*, and bacterial community structure of intertidal soil from Arabian Sea. Microbiol Res 2015;175:57–66. DOI:10.1016/j.micres.2015.02.007.

Kinkel LL, Bakker MG, Schlatter DC. A coevolutionary framework for managing disease-suppressive soils. Annu Rev Phytopathol 2011;49:47–67. DOI:10.1146/annurev-phyto-072910-095232.

Kleinberg JM. Authoritative sources in a hyperlinked environment. J ACM 1999;46:604–32. DOI:10.1145/324133.324140.

Lindsay WL, Norvell WA. Development of a DTPA soil test for zinc, iron, manganese, and copper. Soil Sci Soc Am J 1978;42:421–8.DOI:10.2136/sssaj1978.03615995004200030009x.

López-Lozano NE, Carcaño-Montiel MG, Bashan Y. Using native trees and cacti to improve soil potential nitrogen fixation during long-term restoration of arid lands. Plant Soil 2016;403:317–29. DOI:10.1007/s11104-016-2807-3.

López-Lozano NE, Echeverría Molinar A, Ortiz Durán EA *et al*. Bacterial diversity and interaction networks of *Agave lechuguilla* rhizosphere differ significantly from bulk soil in the oligotrophic basin of Cuatro Cienegas. Front Plant Sci 2020;11:1028.DOI:10.3389/fpls.2020.01028.

López-Lozano NE, Heidelberg KB, Nelson WC *et al*. Microbial secondary succession in soil microcosms of a desert oasis in the Cuatro Cienegas Basin, Mexico. PeerJ 2013;1:e47.DOI:10.7717/peerj.47.

Mandic-Mulec I, Stefanic P, van Elsas JD. Ecology of Bacillaceae. Microbiol Spectr 2015;3:1–24. DOI:10.1128/microbiolspec.TBS-0017-2013.

Matsumoto A, Kasai H, Matsuo Y *et al*. *Ilumatobacter fluminis* gen. nov., sp. nov., a novel actinobacterium isolated from the sediment of an estuary. J Gen Appl Microbiol 2009;55:201–5. DOI:10.2323/jgam.55.201.

Moorhead DL, Reynolds JF. A general model of litter decomposition in the northern Chihuahuan Desert. Ecol Modell 1991;56:197–219. DOI:10.1016/0304-3800(91)90200-K.

Na X, Li X, Zhang Z *et al*. Bacterial community dynamics in the rhizosphere of a long-lived, leguminous shrub across a 40-year age sequence. J Soils Sediments 2018;18:76–84. DOI:10.1007/s11368-017-1745-x.

Nagy ML, Pérez A, Garcia-Pichel F. The prokaryotic diversity of biological soil crusts in the Sonoran Desert (Organ Pipe Cactus National Monument, AZ). FEMS Microbiol Ecol 2005;54:233–45. DOI:10.1016/j.femsec.2005.03.011.

Narcia M, Castillo D, Ramos V *et al*. Turno técnico de la lechuguilla (*Agave lechuguilla* Torr.) en el noreste de México. Rev Mex ciencias For 2012;3:81–8.

Naylor D, Coleman-Derr D. Drought stress and root-associated bacterial communities. Front Plant Sci 2018;8:2223. DOI:10.3389/fpls.2017.02223.

Nelson DW. Determination of ammonium in KCl extracts of soils by the salicylate method. Commun Soil Sci Plant Anal 1983;14:1051–62. DOI:10.1080/00103628309367431.

Okoro CK, Bull AT, Mutreja A *et al*. *Lechevalieria atacamensis* sp. nov., *Lechevalieria deserti* sp. nov. and *Lechevalieria roselyniae* sp. nov., isolated from hyperarid soils. Int J Syst Evol Microbiol 2010;60:296–300. DOI:10.1099/ijs.0.009985-0.

Omirou M, Fasoula DA, Ioannides IM. *Bradyrhizobium* inoculation alters indigenous AMF community assemblages and interacts positively with AMF inoculum to improve cowpea performance. Appl Soil Ecol 2016;108:381–9. DOI:10.1016/j.apsoil.2016.09.018.

Osman KT. Rocks, minerals, and soils. Forest Soils. Cham: Springer International Publishing, 2013, 1–17. DOI:10.1007/978-3-319-02541-4_1.

Paliy O, Shankar V. Application of multivariate statistical techniques in microbial ecology. Mol Ecol 2016;25:1032–57. DOI:10.1111/mec.13536.

Pandey SK, Banik RM. Extractive fermentation for enhanced production of alkaline phosphatase from *Bacillus licheniformis* MTCC 1483 using aqueous two-phase systems. Bioresour Technol 2011;102:4226–31.DOI:10.1016/j.biortech.2010.12.066.

Philippot L, Raaijmakers JM, Lemanceau P, Van Der Putten WH. Going back to the roots: the microbial ecology of the rhizosphere. Nat Rev Microbiol 2013;11(11):789-799. DOI: 10.1038/nrmicro3109

Pointing SB, Belnap J. Microbial colonization and controls in dryland systems. Nat Rev Microbiol 2012;10:654. DOI:10.1038/nrmicro2831.

Poly F, Monrozier LJ, Bally R. Improvement in the RFLP procedure for studying the diversity of *nifH* genes in communities of nitrogen fixers in soil. Res Microbiol 2001;152:95–103. DOI:10.1016/S0923-2508(00)01172-4.

Prăvălie R. Drylands extent and environmental issues. A global approach. Earth-Science Rev 2016;161:259–78. DOI:10.1016/j.earscirev.2016.08.003.

Prim RC. Shortest connection networks and some generalizations. Bell Syst Tech J 1957;36:1389–401.DOI:10.1002/j.1538-7305.1957.tb01515.x.

Qiao Q, Wang F, Zhang J *et al*. The variation in the rhizosphere microbiome of cotton with soil type, genotype and developmental stage. Sci Rep 2017;7:3940. DOI:10.1038/s41598-017-04213-7.

R Core Team. R: A language and environment for statistical computing. 2020.

Ragot SA, Kertesz MA, Bünemann EK. Diversity of the *phoD* alkaline gene in soil. Appl Environ Microbiol 2015;81:7281–9. DOI:10.1128/AEM.01823-15.

Ragot SA, Kertesz MA, Mészáros É *et al*. Soil *phoD* and *phoX* alkaline phosphatase gene diversity responds to multiple environmental factors. FEMS Microbiol Ecol 2017;93:fiw212. DOI:10.1093/femsec/fiw212.

Rashid MI, Mujawar LH, Shahzad T *et al*. Bacteria and fungi can contribute to nutrients bioavailability and aggregate formation in degraded soils. Microbiol Res 2016;183:26–41. DOI:10.1016/j.micres.2015.11.007.

Rojas-Rojas FU, Huntemann M, Clum A *et al*. Draft genome sequence of heavy metal-resistant *Cupriavidus alkaliphilus* ASC-732T, isolated from *Agave* rhizosphere in the northeast of Mexico. Genome Announc 2016;4:e01013-16. DOI:10.1128/genomeA.01013-16.

Rudi K, Skulberg OM, Larsen F *et al*. Strain characterization and classification of oxyphotobacteria in clone cultures on the basis of 16S rRNA sequences from the variable regions V6, V7, and V8. Appl Environ Microbiol 1997;63:2593–9. DOI:10.1016/j.micres.2015.11.007.

Schloss PD, Westcott SL, Ryabin T *et al*. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 2009;75:7537–41. DOI:10.1128/AEM.01541-09.

Schubert CJ, Nielsen B. Effects of decarbonation treatments on δ13C values in marine sediments. Mar Chem 2000;72:55–9. DOI:10.1016/S0304-4203(00)00066-9.

Shange RS, Ankumah RO, Ibekwe AM *et al*. Distinct soil bacterial communities revealed under a diversely managed agroecosystem. PLoS One 2012;7:e40338. DOI:10.1371/journal.pone.0040338

Shi S, Nuccio EE, Shi ZJ *et al*. The interconnected rhizosphere: High network complexity dominates rhizosphere assemblages. Ecol Lett 2016;19:926–36.DOI:10.1111/ele.12630.

Shi Y, Fan K, Li Y *et al*. Archaea enhance the robustness of microbial co-occurrence networks in Tibetan Plateau soils. Soil Sci Soc Am J 2019;83:1093–9. DOI:10.2136/sssaj2018.11.0426.

Simonet P, Grosjean MC, Misra AK *et al*. Frankia genus-specific characterization by polymerase chain reaction. Appl Environ Microbiol 1991;57:3278–86. DOI:10.1128/AEM.57.11.3278-3286.1991.

Spain AM, Krumholz LR, Elshahed MS. Abundance, composition, diversity and novelty of soil Proteobacteria. ISME J 2009;3:992–1000. DOI:10.1038/ismej.2009.43.

Stewart JR. *Agave* as a model CAM crop system for a warming and drying world. Front Plant Sci 2015;6:684. DOI:10.3389/fpls.2015.00684.

Tabatabai MA, Bremner JM. Use of p-nitrophenyl phosphate for assay of soil phosphatase activity. Soil Biol Biochem 1969;1:301–7. DOI:10.1016/0038-0717(69)90012-1.

Torres-Cortés G, Millán V, Fernández-González AJ *et al*. Bacterial community in the rhizosphere of the cactus species *Mammillaria carnea* during dry and rainy seasons assessed by deep sequencing. Plant Soil 2012;357:275–88. DOI:10.1007/s11104-012-1152-4.

Trame AA, Coddington AJ, Paige KN *et al*. Field and genetic studies testing optimal outcrossing in *Agave schottii*, a long-lived clonal plant. Oecologia 1995;104:93–100. DOI:10.1007/BF00365567.

Tschoeke DA, Coutinho FH, Leomil L *et al*. New bacterial and archaeal lineages discovered in organic rich sediments of a large tropical Bay. Mar Genomics 2020;54:100789. DOI:10.1016/j.margen.2020.100789.

Turner S, Pryer KM, Miao VPW *et al*. Investigating deep phylogenetic relationships among Cyanobacteria and plastids by small subunit rRNA sequence analysis. J Eukaryot Microbiol 1999;46:327–38. DOI:10.1111/j.1550-7408.1999.tb04612.x.

Turrion M-B, Lafuente F, Aroca M-J *et al*. Characterization of soil phosphorus in a fire-affected forest Cambisol by chemical extractions and 31P-NMR spectroscopy analysis. Sci Total Environ 2010;408:3342–8. DOI:10.1016/j.scitotenv.2010.03.035.

Ullah A, Akbar A, Luo Q *et al*. Microbiome diversity in cotton rhizosphere under normal and drought conditions. Microb Ecol 2019;77:429–39. DOI:10.1007/s00248-018-1260-7.

Villarreal-Quintanilla JA, Bartolomé-Hernández JA, Estrada-Castillón E *et al*. El elemento endémico de la flora vascular del Desierto Chihuahuense. Acta Bot Mex 2017;Enero:65–96. DOI:10.21829/abm118.2017.1201

Wang Q, Sun H, Xu C *et al*. Analysis of rhizosphere bacterial and fungal communities associated with rusty root disease of Panax ginseng. Appl Soil Ecol 2019;138:245–52. DOI:10.1016/j.apsoil.2019.03.012.

Westerhuis JA, Hoefsloot HCJ, Smit S *et al*. Assessment of PLSDA cross validation. Metabolomics 2008;4:81–9. DOI:10.1007/s11306-007-0099-6.

Willms IM, Rudolph AY, Göschel I *et al*. Globally abundant "*Candidatus Udaeobacter*" benefits from release of antibiotics in soil and potentially performs trace gas scavenging. mSphere 2020;5, DOI:10.1128/msphere.00186-20.

Xia Y, Wang Y, Wang Y *et al*. Cellular adhesiveness and cellulolytic capacity in Anaerolineae revealed by omics-based genome interpretation. Biotechnol Biofuels 2016;9:111. DOI:10.1186/s13068-016-0524-z.

Zhang J, Lv J, Dawuda MM *et al*. Appropriate ammonium-nitrate ratio improves nutrient accumulation and fruit quality in pepper (*Capsicum annuum* L.). Agronomy 2019;9, DOI:10.3390/agronomy9110683.

# Chapter 2. Amino acids in the root exudates of *Agave lechuguilla* Torr. favor the selection and enzymatic activity of nutrient-improvement rhizobacteria

## 2.1 Introduction

*A. lechuguilla* T. is a valuable water-efficient plant from the Chihuahuan Desert, an ecosystem with low nutrient availability and extreme climatic conditions (Houri and Machaka-Houri 2016). However, despite its ecological, cultural, and biotechnological importance, little is known about the plant mechanisms for tolerating arid conditions or even the interactions with microorganisms that may facilitate the plant tolerance to the harsh conditions of the desert, even as such knowledge may be invaluable under the actual pressure of the desertification process and climate change (Reynolds *et al.* 2007; Behnke and Mortimore 2016). In the rhizosphere, the root exudates recruit a beneficial microbiome at the cost of carbon and other nutrients (Schmidt *et al.* 2019), and instead, the plant gets nutrients, defense against pathogens, and phytohormones that increase its biomass or relieve the abiotic stress (Liu *et al.* 2011; Li *et al.* 2016; Vurukonda *et al.* 2016).

Most of the studies about the composition of root exudates use C3 and C4 plants as models of study, while those with crassulacean acid metabolism (CAM) are mainly descriptive and do not explore in-depth the response of the plants to abiotic conditions. For example, in *Opuntia* spp. and *Sedum alfredii*, the most abundant metabolites in the root exudates were amino acids, phenolic compounds, organic acids, and flavonoids (Li *et al.* 2013; Tao *et al.* 2016); all of which had shown a significant influence on the rhizobacteria associated to C3 and C4 plants (Abdel-Lateif, Bogusz and Hocher 2012; Ray *et al.* 2018; Rahmoune *et al.* 2019). Likewise, the factors that influence the composition of the root exudates in those plants have been examined, and soil properties are among the most common, specifically the concentration of certain elements (Naher *et al.* 2009; Ma *et al.* 2017). For instance, the deficit of nitrogen in soil showed to be essential for the amino acid structure and reduced the release of these molecules. Likewise, the lack of potassium (K) and iron (Fe) decreased and enhanced, respectively, the concentration of carbohydrates

(Carvalhais *et al.* 2013b). Moreover, Arginine (Arg) accumulated in plants as response to the deficiency of K, phosphorus (P), magnesium (Mg), sulfur (S), Fe, manganese (Mn), chlorine (Cl), and copper (Cu) (Rabe and Lovatt 1986). This far, most of the studies had been centered on C3 and C4 plants and ecosystems different to the arid lands. Even the compounds studied in the root exudates had been mostly organic acids and sugars, while others like the amino acids had been disregarded (Baziramakenga*,* Simard and Leroux 1995; Yoneyama, Terakado-Tonooka and Minamisawa 2017; Hennion *et al*. 2019).

Similarly, the soil moisture, pH, salinity, and electric conductivity had been reported as essential factors for the plant and influence the composition of root exudates since they define the availability of the nutrients (Henry and Jefferies 2002). Some of these factors are of especial relevance in arid ecosystems since plants are usually exposed to salinity, drought, or alkaline soils (Clark *et al.* 2009). Proline is one of the most studied amino acids for its role as an osmolyte to tolerate drought or salinity. This metabolite aids the plant to keep the turgor and osmotic balance in the cell (Hayat *et al.* 2012). Which hints at the importance of amino acids in the tolerance of plants and their role in the recruitment of rhizobacteria.

Moreover, the plant species and growth stage had been reported as significant factors that shape the composition of the root exudates (Gransee and Wittenmayer 2000; Badri and Vivanco 2009), and thus that of the rhizobacterial community (Qiao *et al*. 2017). For example, Hildebrandt *et al*. (2015) proposed that plants use the amino acids available in the germination stage and reduce their biosynthesis. This pattern will reverse while the plant is actively growing in its juvenile stage, but further in the mature and senescence stages, the biosynthesis of the amino acids will remain at a low pace (Hildebrandt *et al.* 2015).

Likewise, just like plants influence the composition of the bacterial community, it has been found that root exudates may change under the influence of bacteria, pathogen or beneficial alike (Carvalhais *et al.* 2013a; Yuan *et al.* 2018). This variation is reflected in the composition of the bacterial community and the functions of the microbiome, as these influence the recycling of nutrients and its symbiosis (Barret

*et al.* 2011). Thus, it is crucial to comprehend the factors involved in the shaping of the interaction between plants and microorganisms.

The plants can synthesize 20 amino acids (Hutapea *et al.* 2018), which are essential components of their tissues, where they can reach up to 30% of the plant dry biomass (Owen and Jones, 2001). Furthermore, the transport in and out from the roots of these compounds to the soil is through passive transport by a difference in the concentration gradient (Phillips *et al.* 2004), or the active influx with adenosine triphosphate (ATP)-dependent transporter (Farrar *et al.* 2003). Amino acids are also the most abundant components of the bacterial biomass (Miltner *et al.* 2009) and have a primal effect that comes not only from the carbon chains like in sugar and organic acids, but also from the nitrogen and, in some of them, sulfur (Spohn, Ermak and Kuzyakov 2013; Huo, Luo and Cheng 2017). In addition, beyond the organic nitrogen source (Jones *et al.* 2005), they also have roles as metabolic intermediates, substrates, structural components, and physiologic regulators (Jones and Darrah 1994; Moe 2013).

In the rhizosphere, the amino acids are responsible for the recruitment of beneficial bacteria with mechanisms that favor the growth of their host (Oku *et al.* 2012; Ayangbenro and Babalola 2021). Previous works suggest that one of the most important mechanisms is the improvement of nutrient availability. In the rhizosphere of lechuguilla, it has been found that nitrogen-fixing bacteria (NFB) and organic phosphorus mineralizing bacteria (OPMB) are abundant and have essential roles in the sustenance of the bacterial community (López-Lozano *et al.* 2020; Medina-de la Rosa *et al.* 2021). In the first case, the reduction of the atmospheric nitrogen ($N_2$) is possible through the activity of nitrogenases (Zheng and Dean 1994), a complex of a two-component metalloenzyme that, in its most conventional form, includes the iron protein and molybdenum-iron protein (Rees and Howard 2000). Plus, these enzymes are encoded by the *nifHDK* genes (Zehr and Turner 2001), among which the *nifH* gene is included in the alternative forms of nitrogenase. The well-conserved sequence of the *nifH* gene had allowed its use as a phylogenetic marker on a wide variety of terrestrial and aquatic ecosystems (Ueda *et al.* 1995; Zehr *et al.* 2003).

Furthermore, the enzymatic activity of the nitrogenase has been associated with the concentration of histidine (His) and glutamic acid (Glu) (Abu-Zaitoon Yousef 2012), and some examples of diazotrophic bacteria associated with the release of amino acids are *Azotobacter Azospirillum, Rhizobium, Mesorhizobium*, and *Sinorhizobium* (González-López *et al.* 2005).

Phosphorus (P) is a macronutrient highly restricted by its reactivity and, in the alkaline soils of the Chihuahuan Desert, is quickly immobilized by calcium (Ca) and Mg (Perroni *et al.* 2014). Likewise, P mineralization depends on the action of phosphatases produced by OPMB (García-Oliva *et al.* 2018), such as alkaline phosphatases encoded by the *phoD* gene that hydrolyses the organic P forms and aids to increase P availability for the microbiome and the plants alike (Chen *et al.* 2017). The deficiency of this macronutrient in plants seems to enhance the accumulation of nitrogen in the form of arginine (Arg) as a result of the delay in plant growth (Rabe and Lovatt 1986) and the release of gamma-aminobutyric acid and carbohydrates, which favor the recruitment of plant-growth promoting rhizobacteria (PGPR) that could aid to increase P availability (Carvalhais *et al.* 2013b). However, the information about the interaction between nutrient-improvement bacteria and amino acids from root exudates is scarce and even more on the rhizobacterial communities of lechuguilla.

To further understand the influence of root exudates in the rhizobacterial communities associated with *A. lechuguilla* T., the amino acids concentration and the bacteria communities in the rhizosphere of lechuguilla in different growth stages from the Meseta subregion in the Chihuahuan Desert, were characterized. In this regard, it was hypothesized that mature plants release higher doses of amino acids in the rhizosphere than young plants, which would favor the selection and enzymatic activity of PGPR. The specific objectives of this project were: 1) To determine the difference in the amino acids composition between the rhizosphere of mature and young plants and bulk soil, 2) Identify the influence of lechuguilla and its growth stage over the soil properties, 3) Study the rhizobacterial community associated to lechuguilla and identify variations in the functions of the community, as well as those

favored in the rhizosphere, and 4) Identify the influence of plant stage and the amino acids released in the root exudates over the diversity, abundance, and enzymatic activity of NFB and OPMB in the rhizosphere of lechuguilla. Since there is no previous information about the root exudates of lechuguilla, this work offers an insight into its composition, as well as an insight to their effect on the composition of the rhizobacterial communities of *A. lechuguilla* T., especially those with plant-growth promoting mechanisms.

## 2.2 Methods

### 2.2.1 Site description and sample collection

Sampling was made at the end of the dry season in 2018 at Los Amoles, a community inside the Meseta Subregion of the Chihuahuan Desert in San Luís Potosí, Mexico (22°44'8.53"N, 100°29'46.56"W) (Figure 2.2.1).


Figure 2.2.1.Sampling site of soil samples from the rhizosphere of *A. lechuguilla* T. The sampling was made in lechuguillales from the Chihuahuan Desert in Los Amoles, a community from Guadalcázar in San Luís Potosí, Mexico.

The site was located at an altitude of 1,443 m, has a hot semi-arid climate (Bsh), annual precipitation of 667.4 mm, and a range of temperatures from 17.6° to 24.6°C. The Meseta subregion is dominated by desert shrubs communities, with a high abundance of cacti, yucca woodlands, and crasicaule. *A. lechuguilla* T. grows in "lechuguillales", vegetation patches where this plant is dominant (Zavala-Hurtado and Jiménez 2020). In one of these patches and within 20 $m^2$, we selected four

mature plants of lechuguilla with about 40 leaves (Freeman and Reid 1985). Young plants (4) were those joined to the mature ones by the rhizome and a mean of 20 leaves. In addition, to avoid sampling the same specimen, each pair was at least 2 m apart (Trame*,* Coddington and Paige 1995). Rhizosphere samples were taken from soil near the roots, within the leaves' circumference, and the first 10 cm below the surface. These samples were transported to the laboratory, homogenized, and preserved at -4° C until further examination. For molecular analyses, samples were taken directly from the soil attached to the roots and kept at -80° C. Four bulk soil samples were obtained within the sampling site in root-free areas and at least 2 m apart from plants. All these procedures were made with disinfected materials to avoid contamination.

### 2.2.2 Physicochemical soil properties

First, soil moisture (SM) was determined by the gravimetric method. Previous to the analyses of the physicochemical properties, the soil was oven-dried at 60°C and sieved (2 mm). Other properties determined were texture, pH, and electric conductivity (EC) (SEMARNAT 2002). The nutrients quantified included: Total organic carbon (TOC), total organic nitrogen (TON) (Schubert and Nielsen 2000), extractable phosphorus (Bray and Kurtz 1945), ammonium ($NH_4$) (Nelson 1983), nitrite plus nitrates ($NO_2$+$NO_3$) (García-Robledo*,* Corzo and Papaspyoru 2014), micronutrients (Fe, Mn, Zn, and Cu), and exchangeable cations ($Ca^{2+}$, $Mg^{2+}$, $Na^{+}$, and $K^{+}$) (Lindsay and Norvell 1978).

### 2.2.3 Amino acids identification

The quantification of the amino acids from the rhizosphere and bulk soil was performed with the AccQ-TagTM Ultra Derivatization Kit in High-performance liquid chromatography with a fluorescence detector. Briefly, the soil samples were oven-dried at 60°C till constant weight and kept at 4°C. Next, the soil was ground, degreased with hexane, and hydrolyzed with hydrochloric acid in a dry bath at 115°C. Then, the samples were reconstituted, derivatized, and read according to the procedure described in the kit. The results were reported as mg $g^{-1}$ dry soil.

### 2.2.4 DNA extraction and 16S rRNA gene sequencing of bacterial community

The extraction of the genomic material was made according to the directions of the Quick-DNA Fecal/Soil Microbe Miniprep Kit® from Zymo Research (D6010). Region V3-V5 from the 16S rRNA gene was targeted with the primers 357F-5'CTCCTACGGGAGGCAGCAG (Turner *et al.* 1999) and CDR 5'-CTTGTGCGGGCCCCCGTCAATTC (Rudi *et al.* 1997) in the Illumina MiSeq Platform (2x300). The sequences analysis was performed with Mothur (V.1.35.1). Briefly, the reads were filtered according to Q-value (≥ 25), homopolymers (<8), and the ambiguities were omitted. The UCHIME algorithm was used to identify and exclude the chimeric sequences. After that, the sequences were trimmed and aligned with the Nearest Alignment Space Termination algorithm, with the database SILVA 16S rRNA gene (V. 132). Operational taxonomic units (OTU) were obtained from the distance matrix of no redundant sequences (97% similarity). Lastly, the Mothur Bayesian classifier and SILVA database were used for the taxonomic categorization (https://www.arb-silva.de/).

The samples reagent-only and mock community (ZymoBIOMICS™ Microbial Community Standard II, Log Distribution) were included as quality controls. With the mock community we confirmed the accuracy of the bacterial composition above 0.01% of relative abundance. In addition, in the reagent-only sample, some of the most abundant genera were *Desulfotomaculum* (8.72%), *Cutibacterium* (7.27%), *Roseiflexus* (4.93%), and *Xenococcus CRM* (3.55%), which were absent in the rest of the samples. Thus, we discarded the contamination due to the handling of the samples.

The functional prediction was made from 16S rRNA gene sequences in R (V 4.1.0.) (R Core Team, 2020) with the software Tax4fun2 (V 1.1.5). With this method we compared the OTU sequences from 16S rRNA gene and aligned them with the pre-clustered reference database from Kyoto Encyclopedia of Genes and Genomes (KEGG), mode at 99% identity (Wemheuer *et al.* 2020). This way we can have the theoretical potential functions present in the bacterial community.

### 2.2.5 Nitrogenase and phosphomonoesterase activity in soil samples

Nitrogenase activity was quantified with the acetylene reduction method (López-Lozano, Carcaño-Montiel, and Bashan 2016). Briefly, 6 g of soil were deposited in serologic flasks and sealed with rubber stoppers and aluminum caps. From the head-space, 20% of the air was swapped with acetylene. The flasks were incubated for seven days at 37°C, and the ethylene production was measured in a gas chromatograph with a flame ionization detector (Agilent 6890 GC N.05.04). A standard curve of ethylene ($C_2H_4$) from 0.3 to 2.1 nmol, was run for quantification and nitrogenase activity was reported as nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$

Alkaline phosphomonoesterase (APM) activity was determined with the colorimetric method developed by Tabatabai. Following the protocol, 0.5 g of soil were mixed with 2.375 ml Buffer Tris 0.5 M (pH 11) and 0.125 ml of pNPP (p-Nitrophenol phosphate, Sigma-Aldrich P7998) as substrate. The mix was incubated for 30 min at 37°C (150 rpm). After that period, 2 ml of NaOH 0.5 M and 0.5 ml of $CaCl_2$ 0.5 M were added. Finally, samples were centrifuged (2,500 rpm x 5 min), and their absorbance was read at 405 nm (Tabatabai and Bremner 1969). Blanks without pNPP were run alongside the samples. For the quantification, a standard curve with p-nitrophenol (pNP) up to 180 µg $ml^{-1}$ pNP was made, and the APM activity was reported as µg pNP $g^{-1}$ $h^{-1}$.

### 2.2.6 Abundance quantification of *nifH* and *phoD* genes in soil samples

The quantification of NFB and OPMB was made by real-time quantitative polymerase chain reaction (qPCR). For the *nifH* gene, PCR reactions had: 5 µl of SYBR Green 2X, 0.3 mM *PolR* (5'-ATSGCCATCATYTCRCCGGA-3') (Poly, Monrozier and Bally 2001), 0.15 mM *FPGH19* (5'-TACGGCAARGGTGGNATHG-3') (Simonet *et al.* 1991), and 1 µl DNA (1:20) in a final volume of 10 µl. The conditions for the PCR were: Initial denaturation of 10 min at 95°C, 40 cycles of 1 min each of denaturation (94°C), alignment (57°C), and extension (72°C); with a final extension of 72°C for 5 min (Simonet *et al.* 1991). For the *phoD* gene, the 10 µl reaction was composed of 5 µl SYBR Green 2X, 0.4 mM of each primer (*F733* 5'-

TGGGAYGATCAYGARGT-3’ and *R1083* 5’-CTGSGCSAKSACRTTCCA-3’), and 1 µl DNA (1:10). The amplification was made with a three-step reaction of 40 cycles of 15 s of denaturation at 95°C and one minute each of alignment (57°C) and extension (72°C) (Ragot, Kertesz and Bünemann 2015). A final melting step up to 95°C was added to both reactions. Standard curves up to $10^8$ copies were run alongside the samples. Said curves were made with DNA from clones of *Bacillus subtilis* (*phoD*) and *Geobacter sulfurreducens* (*nifH*), obtained with the pGEM-T Easy Vector system (PROMEGA). The reactions were performed in the PikoReal 96 Real-Time PCR system (TCR0096, Thermo Fisher Scientific Inc.).

### 2.2.7 *nifH* and *phoD* sequencing and data processing

The identification of NFB and OPMB was made through *nifH* and *phoD* genes sequencing. The primers used for the reactions were the same as for the qPCR method. The sequencing was performed by the Molecular Research LP MR DNA in the Illumina MiSeq platform (2x300). As quality controls, reagent-only and mock community (ZymoBIOMICS™ Microbial Community Standard II, Log Distribution) were run alongside the samples, however, the PCR reactions of both quality controls were negative maybe due to the lack of extracted DNA or because the bacterial strains in the mock did not have the respective gene. In addition, one sample from bulk soil was discarded since PCR reaction was also negative and sequencing was not viable.

For each *nifH* and *phoD* fastq sequencing file, we split fastq files into forward and reverse paired sequencing files using a custom R script (R Core Team 2020). This script splits the sequences into two sequences (left and right). The right sequence is then reversed complemented. These two sequences are the forward and reverse sequences for each sample. For *nifH* sequencing files, we obtained from 0.2 to 0.3 million sequences. For *phoD* sequencing files, we obtained between 0.1 to 0.2 million sequences.

For each gene, we concatenated all the forward and reverse files into separated individual files. We performed this task with a custom bash script (GNU, 2007).

These concatenated forward and reverse paired-end sequences are used as input to the spades assembling tool (Nurk *et al.* 2017). We used the—meta option.

Next, we utilized the Trinotate pipeline to annotate the resulting assemblies for *nifH* and *phoD* (https://github.com/Trinotate). The tools used by the Trinotate pipeline were blastx and blastp (Altschul *et al.* 1990), signalP (Petersen *et al.* 2011), tmhmm (Krogh *et al.* 2001), RNAmmer (Lagesen *et al.* 2007), and hmmer (Finn, Clements and Eddy 2011). The databases used were PFAM (Punta *et al.* 2012), KEGG (Kanehisa *et al.* 2012), GO (Ashburner *et al.* 2000), and eggnog (Powell *et al.* 2011). The results were gathered into a single data frame using the Trinotate report method with default parameters.

For each *nifH* and *phoD* assembly, we mapped the forward and reverse pairs with hisat2 v2.2.1 (Kim *et al.* 2019). We used the parameters -a -5 10 -3 10 - -very-sensitive. The mapping statistics were around 78% for *nifH* and *phoD* genes (Figure 2.2.2).

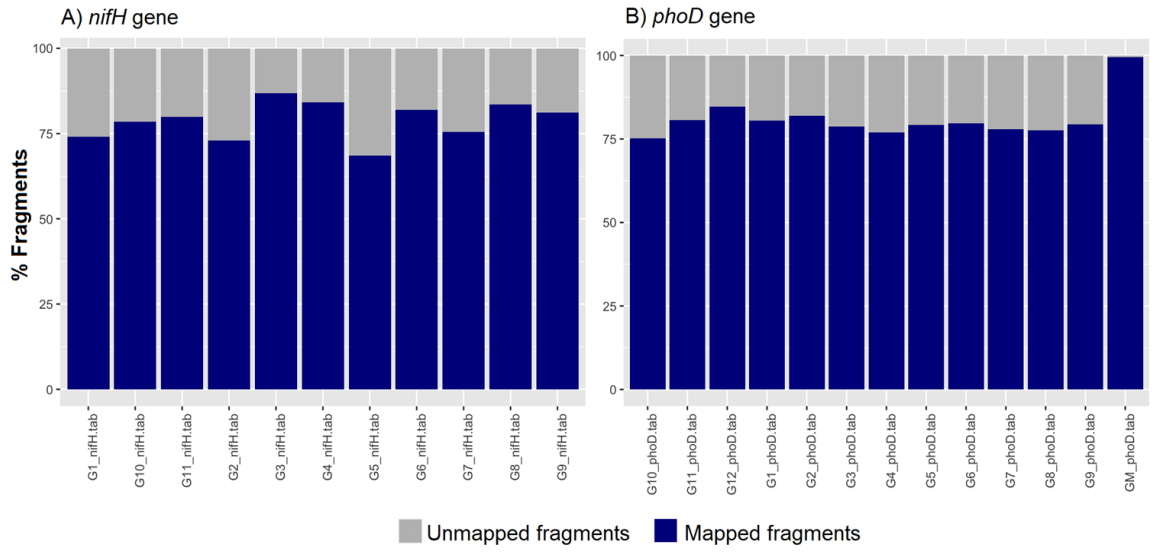



Figure 2.2.2. Percentage of mapped *nifH* and *phoD* fragments from samples of bulk soil and *A. lechuguilla* T. rhizosphere. In blue is the percentage of mapped fragments and in gray the unmapped. Samples 1-4 correspond to mature lechuguilla rhizosphere, 5-8 young plants, 9-12 bulk soil, and GM Mock community.

Finally, with a custom R script (R Core Team 2020), we processed the hisat2 results. With these results, we generated a data frame with the counts for all the samples.

Also, we generated the coverage for each assembled contig, and together with the counts, both parameters were used to filter out contigs.

For the analysis of the ASV sequences, the first quality control was a cut based on the length of the sequences. Each median value, as well as the lower and upper quartiles, were calculated. According to these, *nifH* gene sequences between 227 and 635 nucleotides were kept and 255 to 511 for *phoD*. The next filtering depended on the abundance, so sequences above 0.005% of relative abundance were kept, which excluded singletons and doubletons. This value is considered conservative and recommended when mock communities were not included for calibration (Bokulich *et al.* 2013). Lastly, the functional annotation of *nifH* and *phoD* sequences was revisited and the remaining ASV were those with functions related to nitrogenase and alkaline phosphomonoesterase enzymes.

The raw sequences obtained from Illumina sequencing of 16S rRNA, *nifH* and, *phoD* gene, are available in the GenBank Sequence Read Archive under the Bioproject accession PRJNA672793.

### 2.2.8 Statistical analyses

All statistical analyses described were performed with R (V 4.1.0.) in RStudio (V 1.4.1717) (R Core Team, 2020). Differences in the soil properties, the concentration of the amino acids, and relative abundance of predicted functions were assessed with a one-way analysis of variance (ANOVA), and groups were obtained with Tukey test. Furthermore, these differences were represented with a principal component analysis (PCA), with packages factoMineR V 2.4 and factoExtra V 1.0.7 (See supplementary material d).

The most abundant genera according to the 16S rRNA gene (>1%) were displayed in a heatmap (pheatmap V 1.0.12) (Supplementary material b). The differences in the bacterial composition were tested with an analysis of similarities (ANOSIM) (permutations = 999; vegan V 2.5-7) (Supplementary material c). To test the influence of the amino acid concentration and soil properties (independent variables) over the bacterial community composition, a redundancy analysis (RDA) was

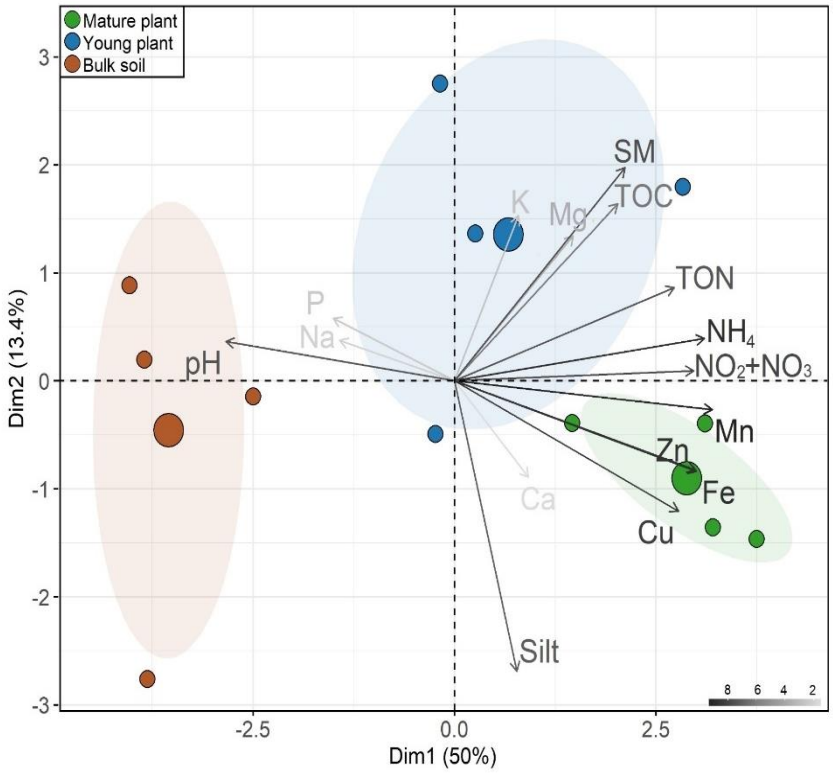
performed (Oksanen *et al.* 2020). First, a non-metric multidimensional scaling (NMDS) with environmental fit was performed for the selection of significant bacterial genera ($p<0.001$) and independent variables ($p<0.05$). The correlation between the genera and the variables was verified with the Pearson method ($p<0.001$). The RDA model was made with the selected variables and the Hellinger transformation of the relative abundance of the taxa (Paliy and Shankar 2016). The significance of the model was verified with a Montecarlo test (permutations = 999) (Supplementary material f). Lastly, the differences in predicted functions were analyzed with a one-way ANOVA and Tukey test.

The difference in the composition of NFB and OPMB by sample type was tested with an ANOSIM (>1% relative abundance), while, soil properties, amino acids, and taxa from *nifH* and *phoD* sequencing were used to obtain linear regression models to explain the variation in the abundance of *nifH* or *phoD* genes and nitrogenase and APM activity. Only those with $\rho>0.3$ were selected (Pearson method, $p<0.01$) and ASV above 1% of relative abundance to reduce the number of independent variables. With these data, we performed PCA for amino acid concentration and soil properties, and NMDS for the taxa; the orthogonal vectors were used in the models (Dimension reduction).

## 2.3 Results

### 2.3.1 The presence of lechuguilla increases nutrient availability in soil

To identify the differences in the soil properties between bulk soil and the rhizosphere of young and mature lechuguillas, one-way ANOVA and Tukey test were applied (Table 2.3-1). According to these tests, the presence of the plant is related to higher nitrogen concentration and EC than bulk soil. This pattern was reflected in the different nitrogen species (TON, $NH_4$, and $NO_2$+$NO_3$). For example, mature and young plants had 0.27 ± 0.02% TON and 0.24 ± 0.02% TON, respectively, while bulk soil had 0.17 ± 0.01% TON ($p<0.05$). As for EC, both mature and young agaves had about 0.22 ± 0.01 $dSm^{-1}$, while bulk soil had 0.17 ± 0.00 $dSm^{-1}$ ($p<0.01$). Furthermore, mature plants had 0.41 ± 0.04 g Cu $kg^{-1}$ and 3.08 ± 0.16 g Fe $kg^{-1}$, nearly twice the concentration in the samples from young lechuguilla and bulk soil ($p<0.01$).




| % Variance contribution | | |
|---|---|---|
| | Dim.1 | Dim.2 |
| TOC | 4.72 | 11.62 |
| TON | 8.55 | 3.20 |
| $NH_4$ | 11.07 | 0.66 |
| $NO_2$+$NO_3$ | 10.13 | 0.04 |
| P | 2.60 | 1.45 |
| Ca | 0.97 | 3.44 |
| K | 0.72 | 10.09 |
| Mg | 2.50 | 7.78 |
| Na | 2.34 | 0.60 |
| Cu | 8.88 | 6.28 |
| Fe | 10.36 | 3.03 |
| Mn | 11.76 | 0.31 |
| Zn | 10.29 | 2.95 |
| pH | 9.26 | 0.57 |
| SM | 5.15 | 16.81 |
| Silt | 0.69 | 31.16 |

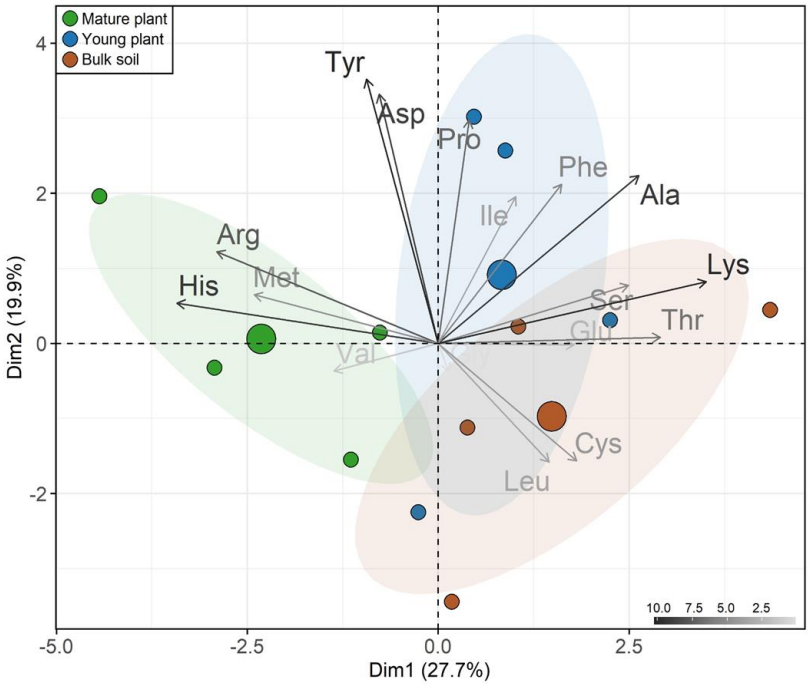
Figure 2.3.1. PCA of physicochemical soil properties. Samples belong to bulk soil and rhizosphere of mature and young lechuguilla, each group of samples shown their own centroid. On the right are displayed the contributions of the variables for each dimension. Explained variance = 63.4%. TOC=Total organic carbon, TON= Total organic nitrogen, and SM= Soil moisture.

Table 2.3-1. Physicochemical properties from the rhizosphere of *A. lechuguilla* T. Soil samples from lechuguilla rhizosphere and bulk soil from Los Amoles in Guadalcázar community, San Luís Potosí, Mexico. F value from one-way ANOVA and similar letters indicate no difference between groups according to Tukey test (* p<0.05, **p<0.01, ***p<0.001).

| Soil properties | Mature plant | Young plant | Bulk soil | F-value |
|---|---|---|---|---|
| **% Total C** | 9.05 ± 0.26 a | 9.32 ± 0.14 a | 8.51 ± 0.22 a | 3.73 |
| **% Total N** | 0.27 ± 0.02 a | 0.24 ± 0.02 a | 0.17 ± 0.01 b | 7.96* |
| **C:N ratio** | 34.11 ± 1.93 b | 40.07 ± 4.16 ab | 51.35 ± 2.64 a | 8.22** |
| **mg $NH_4$ $kg^{-1}$** | 26.85 ± 2.11 a | 23.25 ± 2.04 a | 7.50 ± 0.82 b | 34.15*** |
| **mg $NO_2$+$NO_3$ $kg^{-1}$** | 74.77 ± 5.46 a | 63.21 ± 3.30 a | 35.97 ± 2.67 b | 24.89*** |
| **mg P $kg^{-1}$** | 0.68 ± 0.04 a | 0.53 ± 0.04 a | 1.91 ± 1.15 a | 1.3 |
| **g Cu $kg^{-1}$** | 0.41 ± 0.04 a | 0.29 ± 0.02 b | 0.21 ± 0.01 b | 14.50** |
| **g Fe $kg^{-1}$** | 3.08 ± 0.16 a | 2.63 ± 0.12 ab | 2.13 ± 0.10 b | 12.71** |
| **g Mn $kg^{-1}$** | 14.77 ± 0.39 a | 11.75 ± 0.73 b | 8.13 ± 0.15 c | 46.54*** |
| **g Zn $kg^{-1}$** | 1.46 ± 0.04 a | 0.78 ± 0.07 b | 0.41 ± 0.01 c | 138.20*** |
| **g Ca $kg^{-1}$** | 0.86 ± 0.04 a | 0.84 ± 0.03 a | 0.85 ± 0.02 a | 0.13 |
| **g K $kg^{-1}$** | 0.62 ± 0.01 a | 0.63 ± 0.05 a | 0.62 ± 0.03 a | 0.02 |
| **g Mg $kg^{-1}$** | 0.11 ± 0.00 a | 0.11 ± 0.00 a | 0.10 ± 0.01 a | 1.11 |
| **g Na $kg^{-1}$** | 9.06 ± 0.13 a | 9.17 ± 0.22 a | 10.33 ± 1.22 a | 0.96 |
| **dS $m^{-1}$** | 0.22 ± 0.00 a | 0.22 ± 0.01 a | 0.17 ± 0.00 b | 12.27** |
| **pH** | 8.4 ± 0.03 c | 8.5 ± 0.01 b | 8.6 ± 0.01 a | 17.12*** |
| **% SM** | 27.30 ± 1.07 ab | 29.58 ± 0.89 a | 23.44 ± 1.25 b | 8.28** |
| **% Sand** | 30.10 ± 1.40 b | 33.50 ± 0.96 ab | 35.00 ± 1.29 a | 4.17 |
| **% Clay** | 21.20 ± 0.82 a | 24.60 ± 2.61 a | 18.80 ± 0.71 a | 3.2 |
| **% Silt** | 48.68 ± 0.94 a | 41.85 ± 2.26 b | 46.18 ± 1.26 ab | 4.71* |

Differences between plant-stage were only in pH and concentration of Mn and Zn. The pH of the soil samples decreased gradually from bulk soil (8.6 ± 0.01), young lechuguilla rhizosphere (8.5 ± 0.01), and has the lowest values in mature plants (8.4 ± 0.03; p<0.001). In contrast, Mn (14.77 ± 0.39 g $kg^{-1}$) and Zn (1.46 ± 0.04 g $kg^{-1}$) showed the highest concentration in the rhizosphere of mature lechuguillas than young plants or bulk soil (p<0.001). The differences in soil properties can be visualized in the PCA for which sand, clay, C:N ratio, and EC were omitted to avoid multicollinearity between variables. According to the clustering analysis, which explains about 63.4% of the variation, while lechuguilla samples were related to higher nitrogen concentration, bulk soil was characterized by higher pH. Between

plant stages, we found that mature plants had a higher concentration of micronutrients, and young plants had a higher TOC and SM (Figure 2.3.1).




| % Variance contribution | | |
|---|---|---|
| | Dim1 | Dim2 |
| Ala | 9.12 | 9.21 |
| Arg | 11.14 | 2.75 |
| Asp | 0.79 | 20.30 |
| Cys | 4.35 | 4.51 |
| Glu | 4.19 | 0.00 |
| Gly | 0.02 | 0.25 |
| His | 15.51 | 0.53 |
| Ile | 1.36 | 7.03 |
| Leu | 2.78 | 4.59 |
| Lys | 16.32 | 1.25 |
| Met | 7.69 | 0.77 |
| Phe | 3.46 | 8.27 |
| Pro | 0.22 | 16.33 |
| Ser | 8.21 | 1.10 |
| Thr | 11.20 | 0.01 |
| Tyr | 1.17 | 22.84 |
| Val | 2.47 | 0.25 |

Figure 2.3.2. PCA of amino acids concentration in soil. Samples belong to bulk soil and rhizosphere of mature and young lechuguilla, each group of samples shown their own centroid. On the right are displayed the contributions of the variables to each dimension. Explained variance = 47.6%.

### 2.3.2 Differences in amino acids prevail between mature lechuguilla and bulk soil

To discern the differences in the amino acid concentration between samples, ANOVA and Tukey *post hoc* were applied (Table 2.3-2). According to these analyses, leucine (Leu) and lysine (Lys) had lower concentrations in the rhizosphere than samples from bulk soil ($p<0.05$). For example, Leu concentration in the rhizosphere was about 0.05 mg $g^{-1}$ dry soil for both mature and young plants, which was half the concentration found in bulk soil, 0.10 ± 0.02 mg $g^{-1}$ dry soil. The 19.33 ± 5.81 mg Lys $g^{-1}$ dry soil in bulk soil had notably higher concentration than rhizosphere of mature plants where it only reached 0.19 ±0.06 mg Lys $g^{-1}$ dry soil ($p<0.05$). Lastly, the concentration of threonine (Thr), 0.17 ± 0.02 mg $g^{-1}$ dry soil, was higher in the rhizosphere of young lechuguillas than mature plants, which had 0.06 ± 0.01 mg $g^{-1}$ dry soil. Furthermore, to identify variations between the amino acids as a set, we

applied the ANOSIM test (999 permutations) and found difference between the rhizosphere of mature lechuguilla and bulk soil (R= 0.51, p<0.05). This contrast can be observed in the PCA (Explained variance= 47.6%), where samples from young lechuguilla were mixed with mature and bulk soil samples (Figure 2.3.2). Additionally, while bulk soil samples were clustered by a higher concentration of cysteine (Cys), Leu, and Lys, mature plants distinguish themselves by a higher concentration of arginine (Arg), histidine (His), and methionine (Met).

Table 2.3-2. Amino acids concentration from the rhizosphere of *A. lechuguilla* T. Rhizosphere and bulk soil samples from Los Amoles in Guadalcázar community, San Luís Potosí, Mexico. F value from one-way ANOVA and similar letters indicate no difference between groups according to Tukey test (* p<0.05, **p<0.01, ***p<0.001).

| Amino acid | Mature plant | Young plant | Bulk soil | F value |
|---|---|---|---|---|
| | mg $g^{-1}$ dry soil | | | |
| **Ala** | 12.14 ± 5.27 a | 38.35 ± 10.13 a | 23.33 ± 4.55 a | 3.43 |
| **Arg** | 0.36 ± 0.14 a | 0.14 ± 0.02 a | 0.12 ± 0.01 a | 1.05 |
| **Asp** | 0.26 ± 0.05 a | 0.24 ± 0.06 a | 0.17 ± 0.05 a | 0.74 |
| **Cys** | 8.28 ± 2.92 a | 5.19 ± 4.06 a | 13.90 ± 1.79 a | 2.07 |
| **Glu** | 0.28 ± 0.11 a | 0.50 ± 0.05 a | 0.36 ± 0.12 a | 1.26 |
| **Gly** | 0.16 ± 0.05 a | 0.04 ± 0.02 a | 0.12 ± 0.04 a | 2.76 |
| **His** | 0.27 ± 0.07 a | 0.13 ± 0.00 a | 0.12 ± 0.04 a | 3.01 |
| **Ile** | 0.22 ± 0.03 a | 0.17 ± 0.04 a | 0.23 ± 0.05 a | 0.56 |
| **Leu** | 0.05 ± 0.01 b | 0.05 ± 0.00 b | 0.10 ± 0.02 a | 4.82* |
| **Lys** | 0.19 ± 0.06 b | 11.43 ± 2.75 ab | 19.33 ± 5.81 a | 6.73* |
| **Met** | 0.12 ± 0.05 a | 0.05 ± 0.02 a | 0.06 ± 0.01 a | 1.04 |
| **Phe** | 0.20 ± 0.04 a | 0.18 ± 0.04 a | 0.23 ± 0.06 a | 0.34 |
| **Pro** | 0.19 ± 0.08 a | 1.09 ± 0.55 a | 0.14 ± 0.04 a | 2.73 |
| **Ser** | 0.06 ± 0.01 a | 0.07 ± 0.03 a | 0.09 ± 0.03 a | 0.25 |
| **Thr** | 0.06 ± 0.01 b | 0.17 ± 0.02 a | 0.12 ± 0.02 ab | 8.75** |
| **Tyr** | 0.24 ± 0.04 a | 0.25 ± 0.07 a | 0.13 ± 0.04 a | 1.67 |
| **Val** | 0.18 ± 0.06 a | 0.06 ± 0.01 a | 0.08 ± 0.02 a | 2.70 |

### 2.3.3 Proteobacteria and Bacteroidetes are the most abundant phyla in the rhizosphere of lechuguilla

The identification of the bacterial community was based on the 16S rRNA gene sequencing. Mock community and reactive-only samples were added as quality controls (Table 0-6), aided in discarding contamination due to manipulation, and set the data accuracy above 0.01% of the relative abundance.

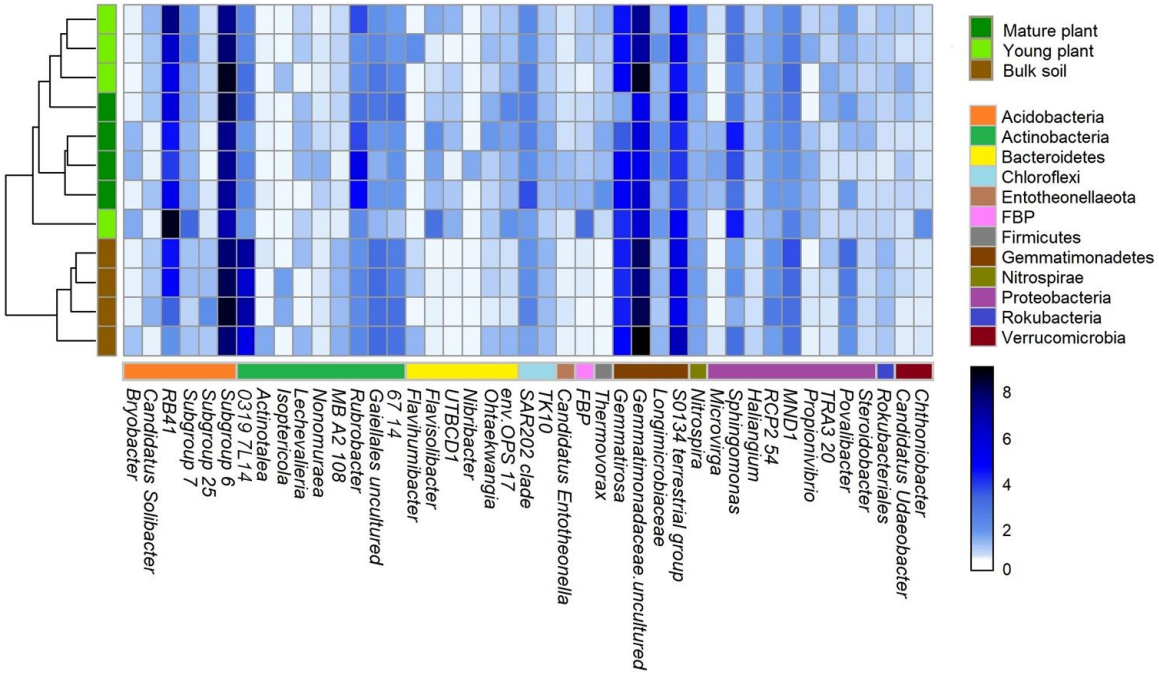
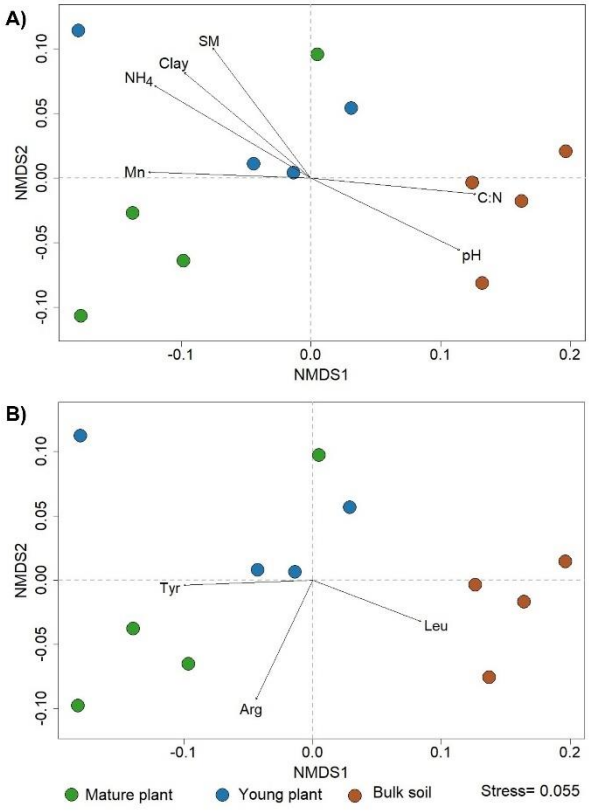



Figure 2.3.3. Relative abundance of bacterial genera in bulk soil and the rhizosphere of mature and young plants of *A. lechuguilla* T. The genera displayed had relative abundance above 1%. Tree on the left was built with Bray Curtis distance.

Among the most abundant phyla in the soil samples from Amoles were Proteobacteria, Actinobacteria, Acidobacteria, and Gemmatimonadetes (Figure 2.3.3). While the genera *0319 7L14, MB A2 108, Gaiellales uncultured*, and *67 14* (Actinobacteria) showed more relative abundance in bulk soil; the rhizosphere of lechuguilla seems to hold a higher abundance of *Flavisolibacter*, *UTBCD1* (Bacteroidetes), *Microvirga*, and *Propionivibrio* (Proteobacteria) (Figure 2.3.3). Clustering with Bray Curtis distance indicated that the main difference between groups corresponds to the presence of the plant, and this was corroborated with ANOSIM (999 permutations). The test showed that the bacterial communities of the

rhizosphere of mature (R=0.94, p<0.05) and young lechuguilla (R=0.82, p<0.05) were different from those in bulk soil samples; but there was no difference between plant stage (R=0.250, p>0.05).

Furthermore, to identify the soil properties related to the composition of the bacterial community, RDA was done with the relative abundance of the bacterial genera. Previously, the soil properties and amino acids selected for the model were those with significant envfit (p<0.05) in an NMDS (Figure 2.3.4A). Bacterial genera were selected in a similar manner (envfit p<0.001) (Figure 2.3.4).




| Phyla | Genera | R |
|---|---|---|
| **Acidobacteria** | *Acidobacteriaceae (Subgroup 1) uncultured* | 0.94 |
| | *Subgroup 25* | 0.80 |
| **Actinobacteria** | *0319 7L14* | 0.87 |
| | *Atopobium* | 0.92 |
| | *Cellulomonas* | 0.88 |
| | *Gaiellales uncultured* | 0.93 |
| | *MB A2 108* | 0.92 |
| | *Phytomonospora* | 0.92 |
| | *Quadrisphaera* | 0.83 |
| | *Ruania* | 0.86 |
| | *Saccharopolyspora* | 0.91 |
| | *Tenggerimyces* | 0.88 |
| | *Thermoleophilia uncultured* | 0.83 |
| **Bacteroidetes** | *Candidatus Paenicardinium* | 0.89 |
| | *Chitinophagaceae uncultured* | 0.85 |
| | *MSB 3C8* | 0.84 |
| **Chloroflexi** | *AKIW781* | 0.90 |
| | *Chloronema* | 0.85 |
| | *Kallotenue* | 0.93 |
| | *Kouleothrix* | 0.85 |
| | *Roseiflexus* | 0.81 |
| **Firmicutes** | *Fervidicola* | 0.80 |
| | *Paenibacillus* | 0.78 |
| **Gemmatimonadetes** | *Gemmatimonas* | 0.85 |
| **Proteobacteria** | *AT s16* | 0.84 |
| | *Beijerinckiaceae uncultured* | 0.80 |
| | *Candidatus Fokinia* | 0.85 |
| | *Caulobacteraceae* | 0.87 |
| | *Constrictibacter* | 0.87 |
| | *Craurococcus* | 0.81 |
| | *Crenalkalicoccus* | 0.64 |
| | *Dechlorosoma* | 0.83 |
| | *Hyphomicrobium* | 0.81 |
| | *Kushneria* | 0.69 |
| | *MBNT15* | 0.84 |
| | *Methylibium* | 0.81 |
| | *Methyloparacoccus* | 0.90 |
| | *Microvirga* | 0.92 |
| | *Rubellimicrobium* | 0.93 |
| | *Sphingoaurantiacus* | 0.71 |
| **Rokubacteria** | *Rokubacteriales* | 0.82 |
| **Verrucomicrobia** | *Roseimicrobium* | 0.80 |

Figure 2.3.4. NMDS of bacterial genera from bulk soil, mature and young lechuguilla rhizosphere from the Chihuahuan Desert. To explain the clustering of the samples an envfit was added. A) Soil physicochemical properties and B) amino acids (p<0.05). On the right, genera with significant envfit in NMDS (p<0.001).

RDA with soil properties indicated that SM, $NH_4$, and Mn were correlated with the composition of the rhizobacterial communities of lechuguilla (Explained variance= 73.06%). In contrast, a higher pH and C:N ratio were correlated with the bacterial

community in bulk soil (Figure 2.3.5A). In addition, according to RDA with amino acids (Explained variance = 59.72%), Leu was correlated with the composition of the bacterial communities in bulk soil. Finally, while Tyr was correlated with both rhizobacterial communities, Arg seemed to influence the microbiome composition in mature plants (Figure 2.3.5B).

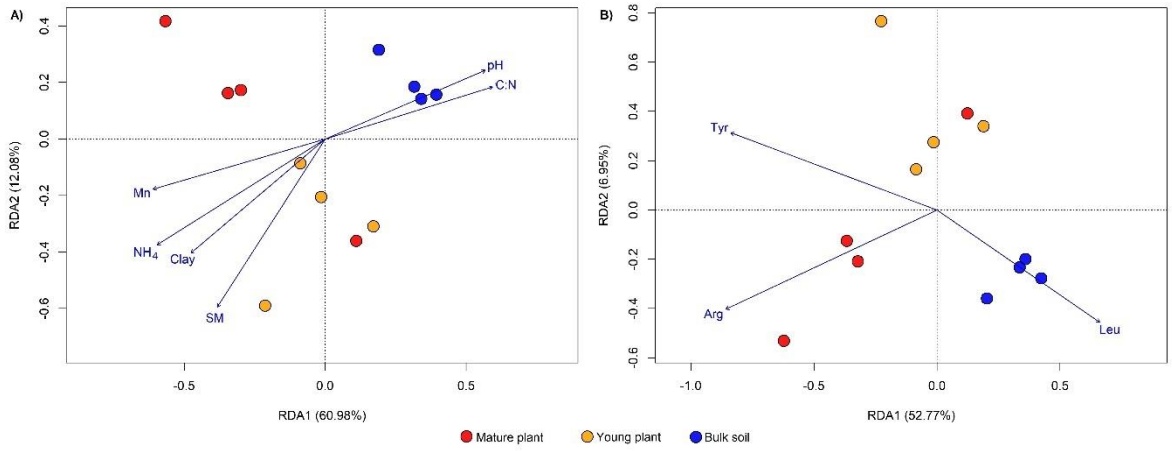



Figure 2.3.5. Redundancy analysis of the bacterial community associated to *A. lechuguilla* T. Bacterial diversity is explained by: A) Physicochemical soil properties (Exp. var. = 73.06%) and B) Amino acids (Exp. var. = 59.72%).

### 2.3.4 Theoretical functions in the rhizosphere of lechuguilla are centered on the metabolism of amino acids

The functional prediction of the microbial community was based on the 16S rRNA gene. OTU were assigned according to the SILVA database and transformed into functional traits with Tax4fun2. For our samples, this prediction included about 31.07% of the assigned OTU. The functions considered for further analysis included: Cellular processes, environmental information processing, genetic information processing, and metabolism. According to the ANOVA and Tukey test, the bacterial community in bulk soil showed an increase in the functions related to the metabolism of some amino acids (tryptophan - Trp, alanine - Ala, glutamine - Gln, and glutamate - Glu) (Figure 2.3.6), carbohydrates (butanoate, fructose, mannose, and propanoate) (Figure 2.3.7), lipids (fatty acids, alpha-linolenic acid, and steroids), the biosynthesis of terpenoids and xenobiotics degradation, such as, benzoate and caprolactam (Figure 2.3.8). Besides, photosynthesis and nucleotide excision repair were also

abundant functions in the root-free soil (Figure 2.3.8-9). In the rhizosphere of both stages of lechuguilla, the biosynthesis of valine (Val), Leu, isoleucine (Ile), and the metabolism of other amino acids had an increased abundance (Figure 2.3.6). In addition, carbon sources preferred by the bacteria in the rhizosphere included inositol phosphate, pentose, and glucuronate interconversions as the functions related to the metabolism of these carbohydrates were increased (Figure 2.3.7). There was also an augmented abundance of biosynthesis and degradation of secondary metabolites and xenobiotics, such as betalain, flavones, and flavonoids (Figure 2.3.8). Mature plants had higher relative abundance functions related to antibiotics and communications with *quorum sensing* (Figure 2.3.9), than younger lechuguilla or bulk soil bacterial communities.

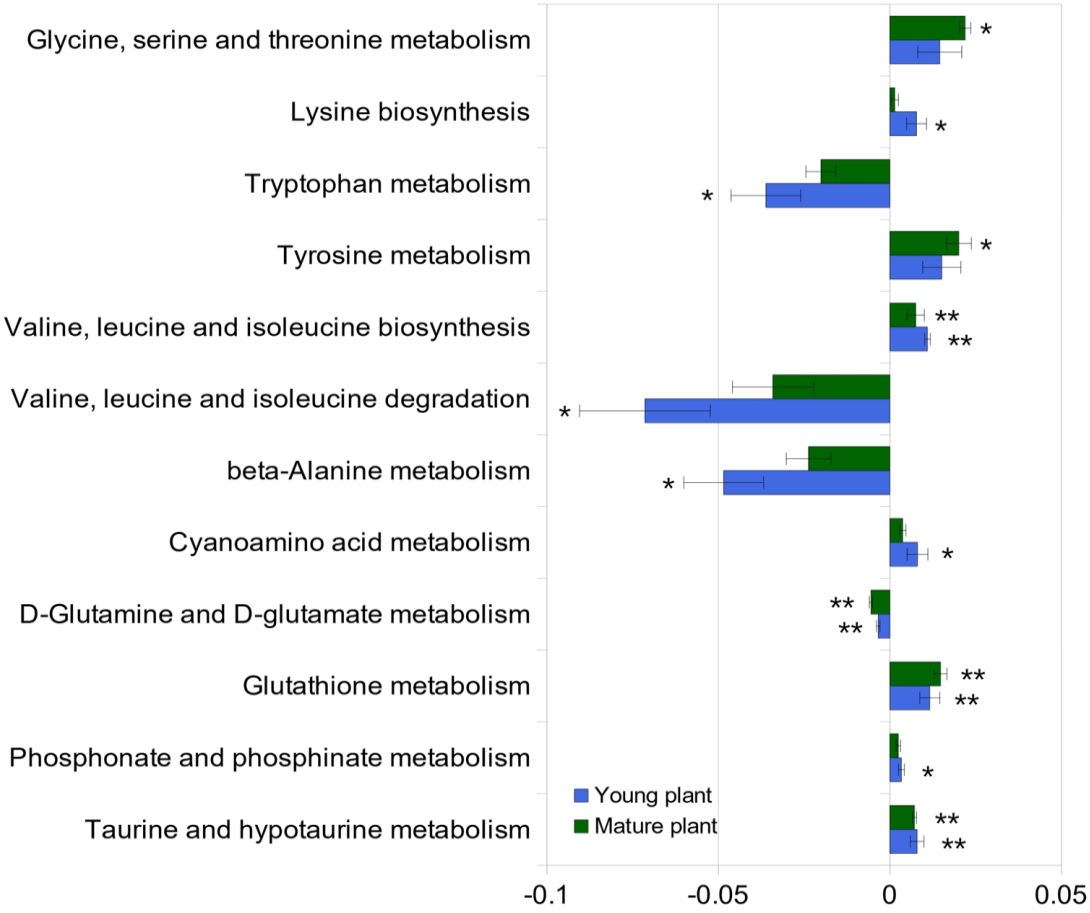



Figure 2.3.6. Predicted functions related to the metabolism of amino acids in the rhizobacterial communities of young and mature *A. lechuguilla* T. Values correspond to the difference between rhizosphere and bulk soil samples with standard error. Differences were tested with ANOVA and Tukey (* = p<00.5, ** = p<0.01, and *** p<0.001).

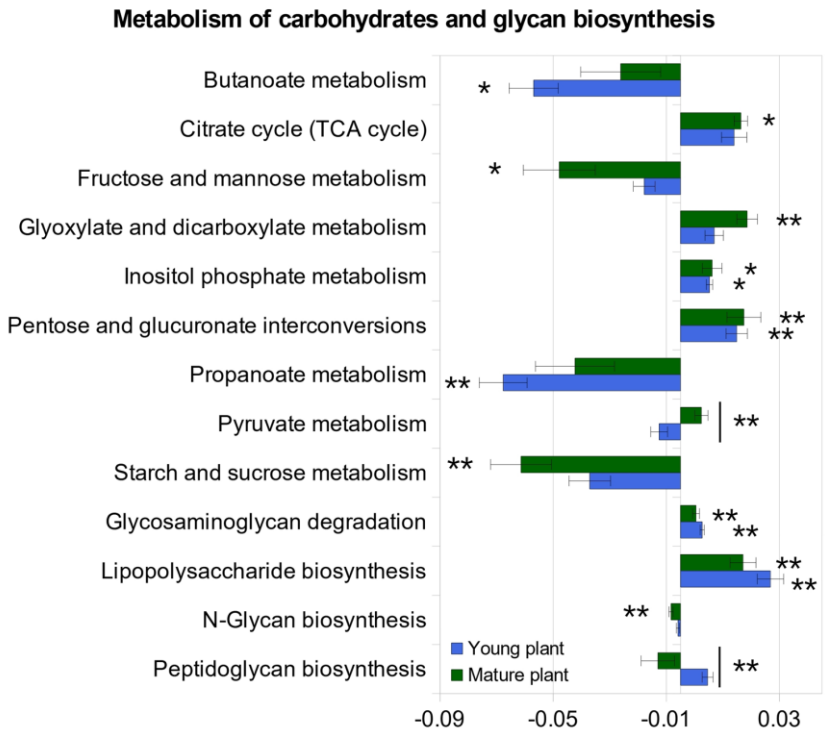



Figure 2.3.7. Predicted functions related to the metabolism of carbohydrates and glycan biosynthesis in the rhizobacterial communities of mature and young *A. lechuguilla* T. Values correspond to the difference between rhizosphere and bulk soil samples with standard error. Differences were tested with ANOVA and Tukey (* = p<00.5, ** = p<0.01, and *** p<0.001).

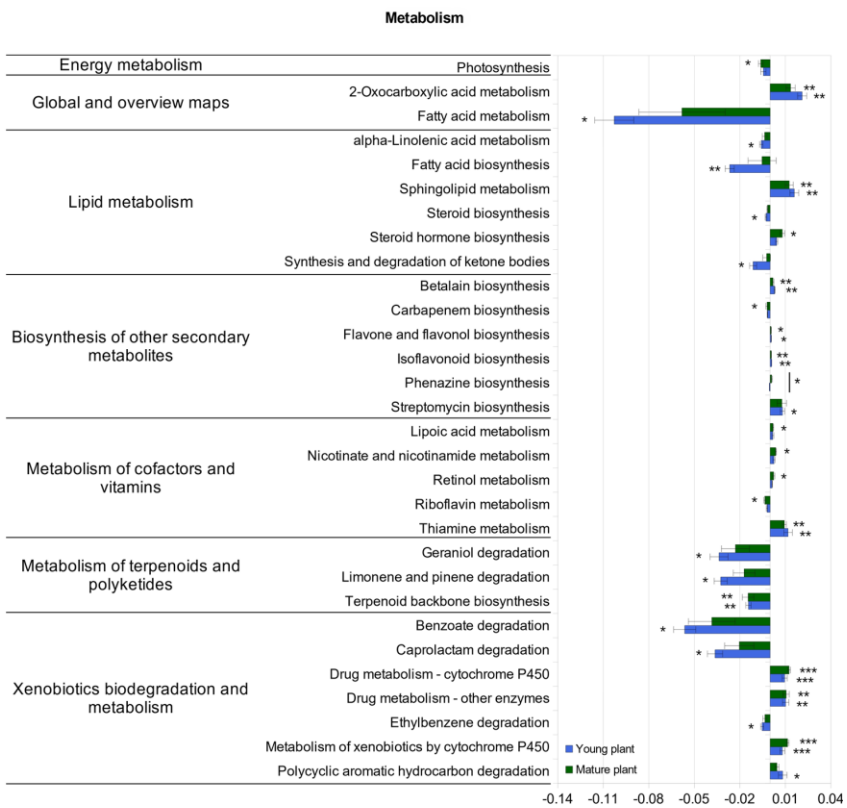



Figure 2.3.8. Predicted functions related to the metabolism of bacterial communities of young and mature *A. lechuguilla* T. Values correspond to the difference between rhizosphere and bulk soil samples with standard error. Differences were tested with ANOVA and Tukey (* = p<00.5, ** = p<0.01, and *** p<0.001).

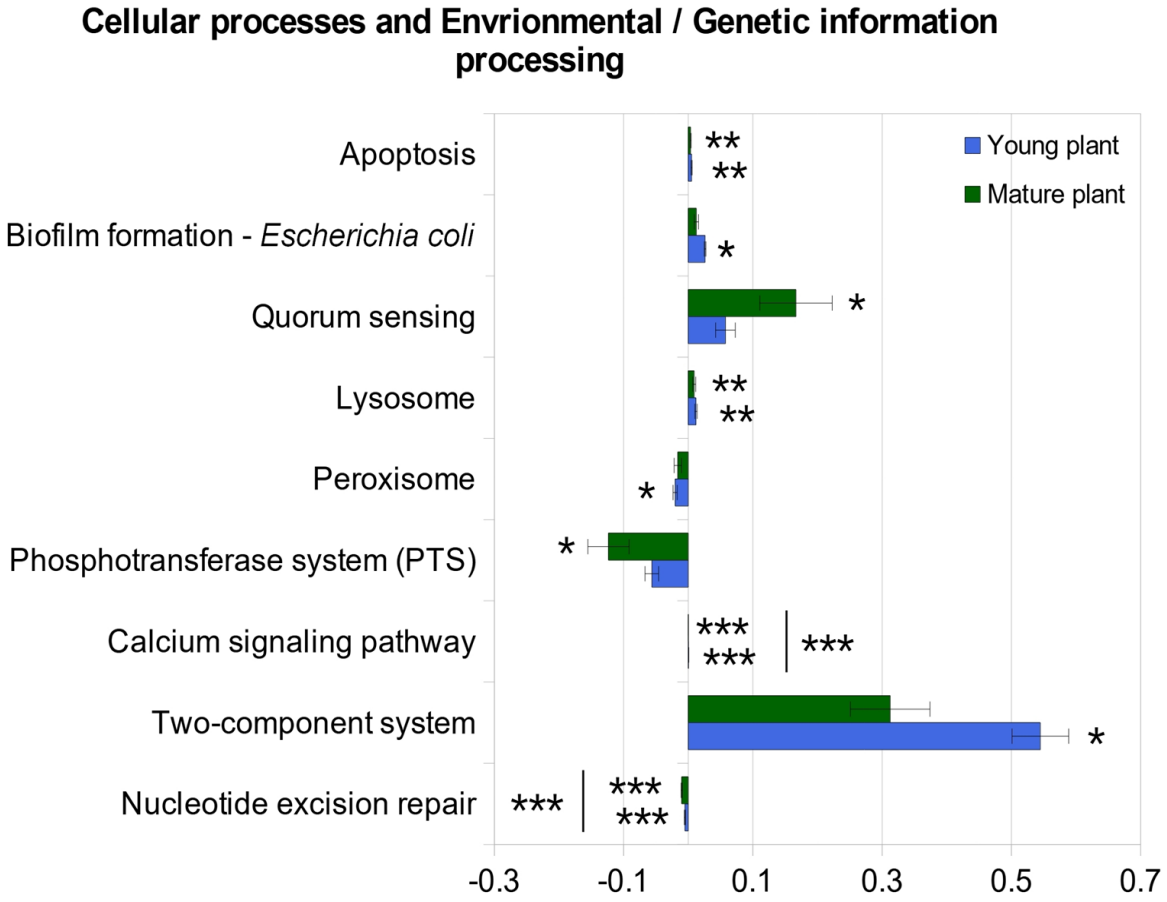



Figure 2.3.9. Predicted functions related to cellular processes and environmental / genetic information processing in the bacterial communities of young and mature *A. lechuguilla* T. Values correspond to the difference between rhizosphere and bulk soil samples with standard error. Differences were tested with ANOVA and Tukey (* = p<00.5, ** = p<0.01, and *** p<0.001).

### 2.3.5 Plant growth promoting related functions are enriched in the rhizosphere of lechuguilla

Bulk soil bacterial communities were enriched with microorganisms tolerant to environmental stress ($p<0.05$), associated with the production of biofilm and osmolytes (Figure 2.3.10). In addition, these communities had abundant microorganisms with the capacity of inorganic phosphate uptake, the transport of siderophores, and the production of hydrolytic enzymes (Figure 2.3.10). Instead, in the rhizosphere of lechuguilla, the acquisition of nutrients seemed to be favored by the fixation of nitrogen and the mineralization of organic phosphate ($p<0.05$).

Functions related to these pathways include the production of flavonoids, phosphate transporters, and denitrification (Figure 2.3.10), which were more abundant in the rhizosphere than bulk soil communities. Other abundant functions were those related to the metabolism of siderophores, osmolytes accumulation, ACC deaminase, indole pyruvic acid production, and protection against pathogens by means of the production of antibiotics, *quorum quenching* (Figure 2.3.10). The only difference between plant-stage was for indole-3-acetonitrile, which was higher in younger plants than mature ($p<0.05$).

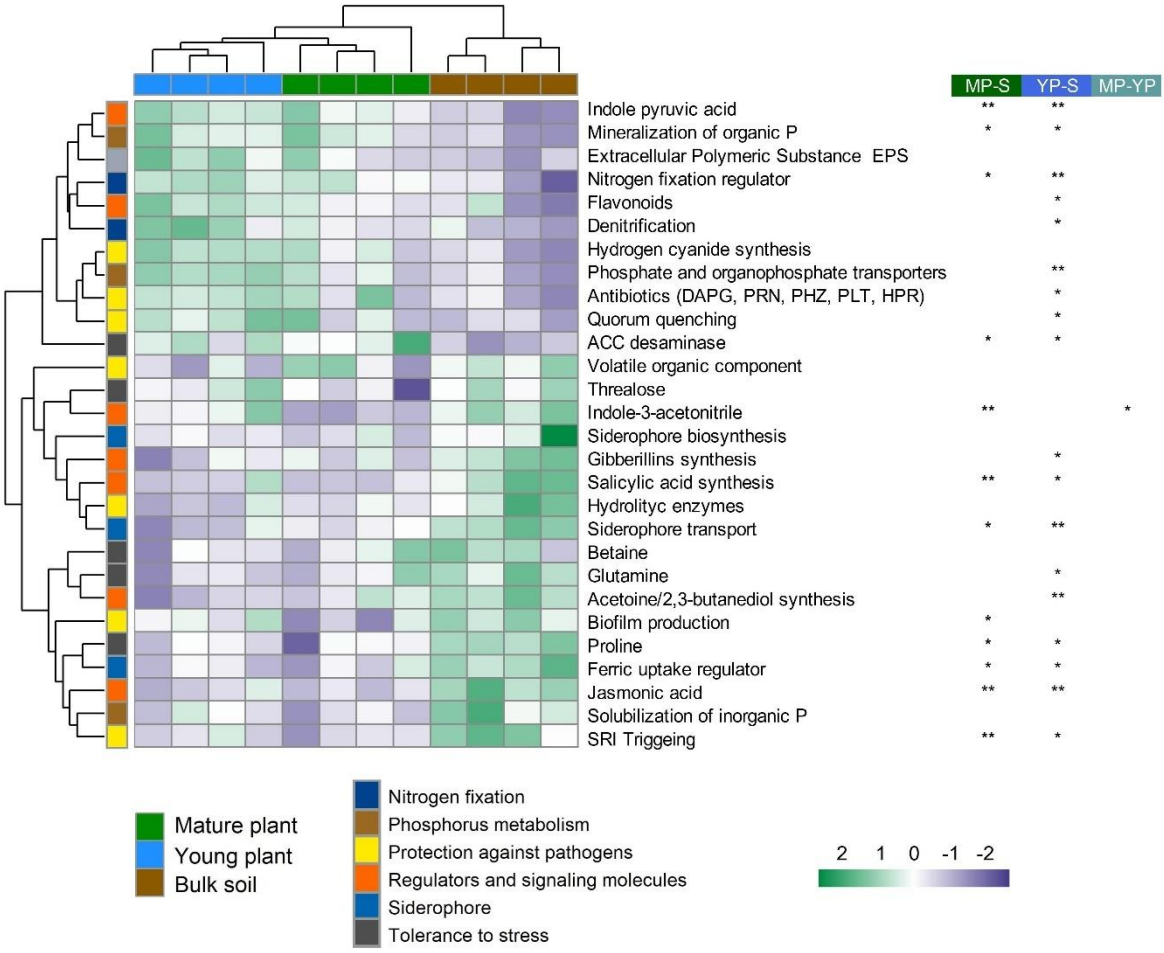



Figure 2.3.10. Plant growth-promoting related functions in the bacterial community of the rhizosphere of *A. lechuguilla* T. and bulk soil. Logarithm of the relative abundance of the functions. Difference in relative abundance was tested with ANOVA and Tukey (* = p<00.5, ** = p<0.01, and *** p<0.001) (Right). MP= Mature plant, YP= Young plant, and S= Bulks soil.

### 2.3.6 Firmicutes and Proteobacteria are the most abundant OPMB and NFB in the rhizosphere of lechuguilla

To identify the communities of NFB and OPMB associated with the rhizosphere of lechuguilla, *nifH* and *phoD* genes were sequenced (Table 0-7). In the rhizosphere *unclassified* of lechuguilla, we found 113 ASV from *nifH*, most of which belong to the phyla Proteobacteria (*Acidithiobacillus* -1 ASV-, *Sinorhizobium* -6 ASV-, *Gluconobacter* -1 ASV-, *Azospirillum* -1 ASV-, *Burkholderia* -2 ASV-, and *Pseudomonas* -21 ASV-) and Firmicutes (*Bacillus* -11 ASV-, *Geobacillus* -3 ASV-, and *Ruminiclostridium* -2 ASV-). These taxa diminish considerably in bulk soil, while Cyanobacteria (*Nostoc* -3 ASV-) and Unknown taxa (13 ASV), reach around 82.87% of the community (Figure 2.3.11B). However, the composition of the bacterial communities only differs between the rhizosphere of mature lechuguilla and bulk soil (ANOSIM, 999 permutations; R=0.222, p<0.049) and, even though there is no difference between mature and young lechuguilla, we found an increased abundance of the genera *Pseudomonas, Nostoc, Frankia* -1 ASV-, and *Unknown* taxa in the rhizosphere of young lechuguilla (Figure 2.3.11B).

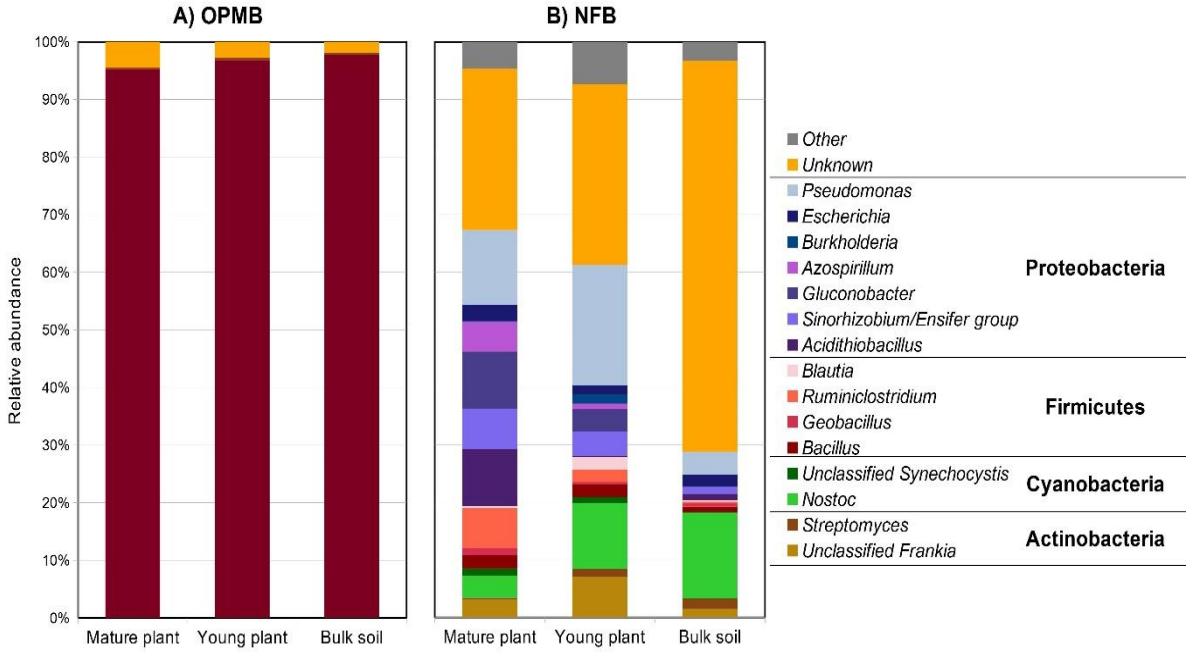
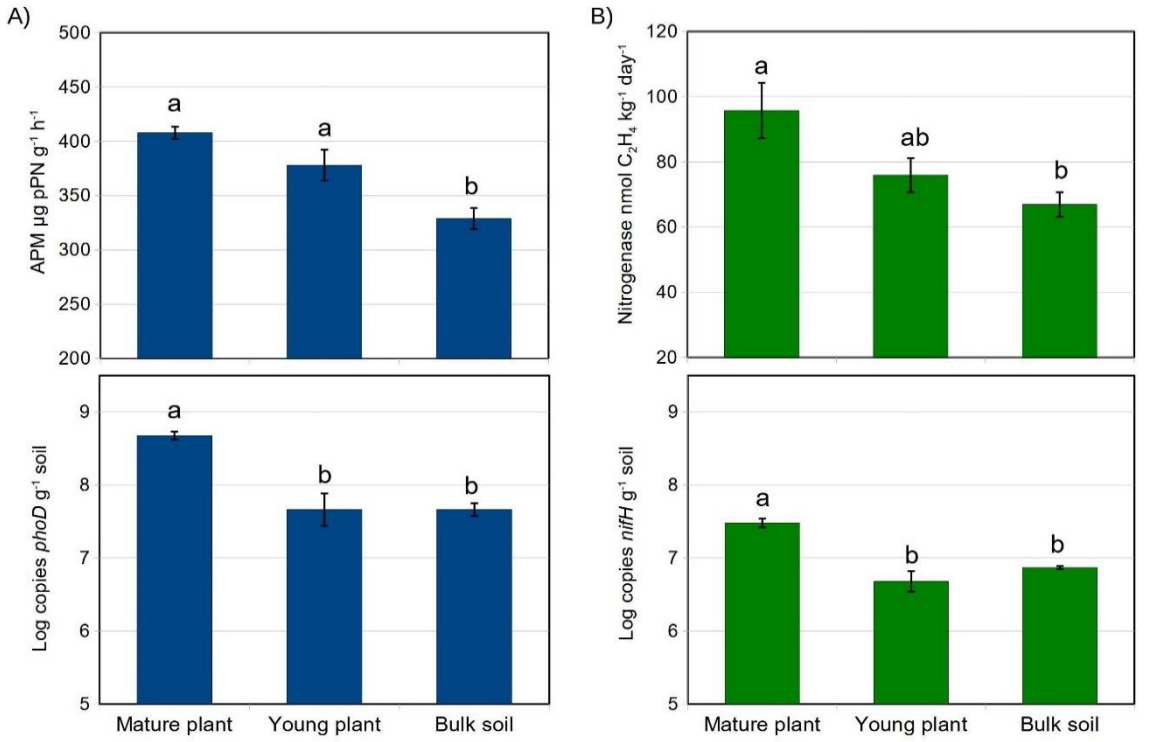



Figure 2.3.11. Bacterial composition in the rhizosphere of lechuguilla and bulk soil, obtained by *nifH* and *phoD* sequencing. A) Organic phosphorus mineralizing bacteria (OPMB) and B) Nitrogen-fixing bacteria (NFB). Genera displayed had relative abundance over 1%.

Further characterization of the nitrogen-fixing community included quantifying the nitrogenase activity and the abundance of the *nifH* gene. The activity of the nitrogenase in mature (95.72 ± 8.52 nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$) and young lechuguilla (75.89 ± 5.25 nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$) were higher than bulk soil (66.89 ± 3.77 nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$) (ANOVA, Tukey $p<0.05$). As for the *nifH* gene abundance, the samples from mature lechuguilla (7.48 ± 0.06 log copies *nifH* $g^{-1}$ $soil^{-1}$) had higher counts than young plants (6.68 ± 0.14 log copies *nifH* $g^{-1}$ $soil^{-1}$) or bulk soil (6.87 ± 0.02 log copies *nifH* $g^{-1}$ $soil^{-1}$; $p<0.001$) (Figure 2.3.12).



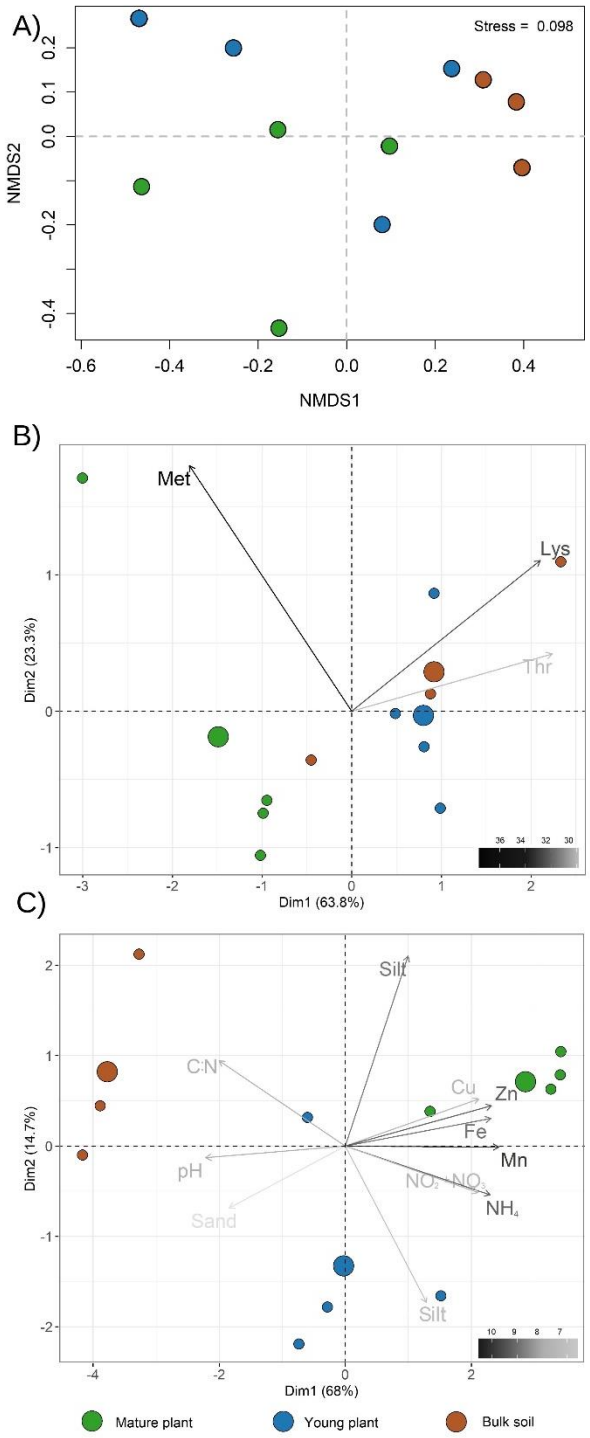

**Organic phosphorus mineralizing bacteria**

| | Variable | F value | Percentage contribution |
|---|---|---|---|
| **APM** | Soil Dim1 | 12.19** | 33.30% |
| | Aminoacids Dim1 | 12.52** | 36.16% |
| | Adj. $R^2$ = 0.6737, p <0.01 | | |
| **Log *phoD*** | Soil Dim1 | 5.89* | 30.51% |
| | Aminoacids Dim1 | 8.37* | 30.65% |
| | Adj. $R^2$ = 0.5274, p <0.05 | | |

**Nitrogen fixing bacteria**

| | Variable | F value | Percentage contribution |
|---|---|---|---|
| **Nitrogenase** | Soil Dim1 | 10.98* | 44.54% |
| | ASV Dim 2 | 5.67* | 22.99% |
| | Adj. R2 = 0.5942, p <0.05 | | |
| **Log *nifH*** | Aminoacids Dim 1 | 12.48** | 58.11% |
| | Adj. $R^2$ = 0.5345, p <0.01 | | |

Figure 2.3.12. Enzymatic activity and abundance of NFB and OPMB associated to the rhizosphere of *A. lechuguilla* T. A) Activity of nitrogenase and APM. B) Abundance of *nifH* and *phoD* genes. Different letters indicate significant difference (ANOVA, Tukey p<0.05). Capitals are for OPMB and low-case for NFB. Linear regression models were constructed with the loading dimensions of multivariate analysis from soil properties, amino acids, and ASV.

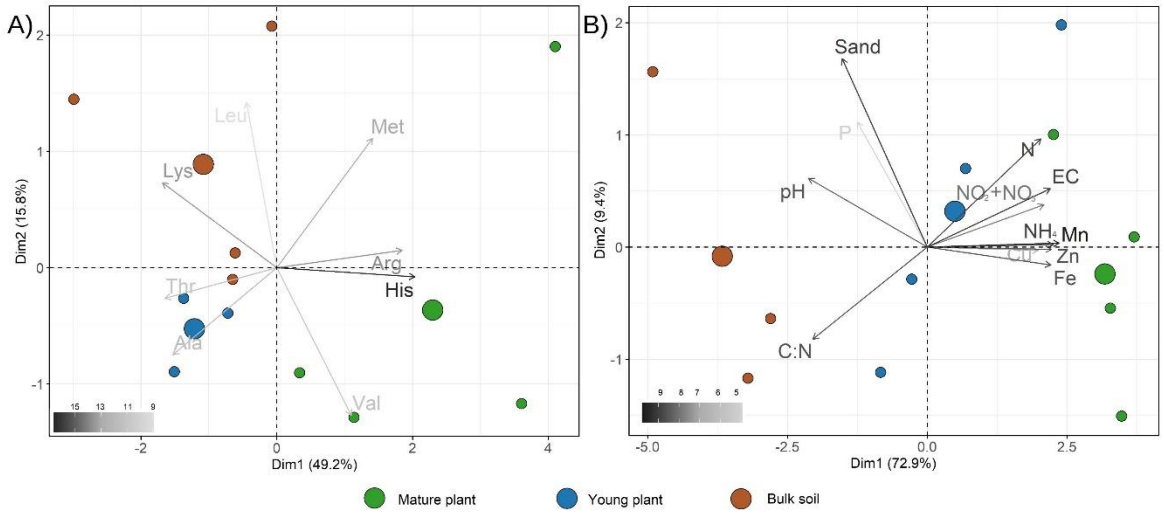



| *nifH* ASV >1% | NMDS1 | NMDS2 | $R^2$ | p value |
|---|---|---|---|---|
| *Acidithiobacillus_nifH109* | -0.14 | -0.99 | 0.40 | 0.11 |
| *Azospirillum_nifH056* | -0.72 | -0.69 | 0.14 | 0.56 |
| *Bacillus_nifH006* | -0.4 | -0.92 | 0.08 | 0.75 |
| *Bacillus_nifH010* | -0.43 | 0.9 | 0.17 | 0.48 |
| *Bacillus_nifH028* | -1 | -0.05 | 0.08 | 0.75 |
| *Bacillus_nifH035* | -0.61 | 0.8 | 0.47 | 0.1 |
| *Blautia_nifH075* | -0.46 | 0.89 | 0.12 | 0.67 |
| *Burkholderia_nifH019* | -0.59 | 0.81 | 0.46 | 0.11 |
| *Chloroherpeton_nifH068* | -0.73 | 0.68 | 0.09 | 0.7 |
| *Clostridium_nifH017* | -0.6 | 0.8 | 0.29 | 0.26 |
| *Colwellia_nifH009* | -0.59 | 0.81 | 0.51 | 0.07 |
| *Escherichia_nifH013* | -0.09 | -1 | 0.70 | 0.01 |
| *Escherichia_nifH021* | 0.25 | -0.97 | 0.00 | 1 |
| *Escherichia_nifH069* | 0.18 | 0.98 | 0.15 | 0.53 |
| *Geobacillus_nifH030* | -0.11 | -0.99 | 0.64 | 0.01 |
| *Geobacillus_nifH103* | 0.39 | 0.92 | 0.25 | 0.31 |
| *Gluconobacter_nifH066* | -0.99 | -0.12 | 0.27 | 0.28 |
| *Nitrosomonas_nifH023* | -0.19 | -0.98 | 0.00 | 0.99 |
| *Nostoc_ge_nifH043* | 0.92 | -0.39 | 0.10 | 0.69 |
| *Nostoc_nifH104* | 0.2 | -0.98 | 0.11 | 0.77 |
| *Nostoc_nifH110* | 0.79 | -0.61 | 0.35 | 0.18 |
| *Phormidium_nifH108* | 0.6 | 0.8 | 0.09 | 0.69 |
| *Pseudomonas_nifH033* | -0.7 | 0.71 | 0.40 | 0.16 |
| *Pseudomonas_nifH057* | -0.69 | 0.73 | 0.63 | 0.02 |
| *Pseudomonas_nifH061* | -0.14 | -0.99 | 0.54 | 0.03 |
| *Pseudomonas_nifH063* | 0.75 | 0.66 | 0.13 | 0.67 |
| *Pseudomonas_nifH073* | -0.48 | 0.88 | 0.43 | 0.12 |
| *Pseudomonas_nifH078* | -0.96 | -0.26 | 0.09 | 0.7 |
| *Pseudomonas_nifH090* | -0.86 | -0.51 | 0.09 | 0.7 |
| *Pseudomonas_nifH092* | 0.07 | -1 | 0.16 | 0.54 |
| *Pseudomonas_nifH100* | -0.58 | 0.82 | 0.67 | 0.02 |
| *Ruminiclostridium_nifH105* | -0.99 | -0.11 | 0.39 | 0.16 |
| *Salmonella_nifH102* | -0.09 | 1 | 0.25 | 0.3 |
| *Sinorhizobium/Ensifer group_nifH052* | -0.74 | -0.67 | 0.30 | 0.21 |
| *Sinorhizobium/Ensifer group_nifH101* | -1 | -0.03 | 0.12 | 0.61 |
| *Staphylococcus_nifH095* | 0.06 | -1 | 0.27 | 0.3 |
| *Streptomyces_nifH062* | -0.47 | 0.88 | 0.14 | 0.59 |
| *Streptomyces_nifH111* | 0.7 | 0.71 | 0.15 | 0.6 |
| *Teredinibacter_nifH001* | -0.77 | 0.64 | 0.07 | 0.79 |
| *Thermoanaerobacterium_nifH024* | -0.48 | 0.88 | 0.19 | 0.37 |
| *unclassified Frankia_nifH112* | -0.58 | 0.81 | 0.54 | 0.04 |
| *unclassified Synechocystis_nifH008* | -0.45 | 0.9 | 0.17 | 0.46 |
| *unclassified Synechocystis_nifH065* | -0.86 | -0.51 | 0.35 | 0.16 |
| *Unknown02_nifH038* | -0.64 | 0.77 | 0.42 | 0.09 |
| *Unknown06_nifH049* | -0.19 | -0.98 | 0.60 | 0.01 |
| *Unknown07_nifH058* | -0.53 | -0.84 | 0.67 | 0.02 |
| *Unknown08_nifH064* | -0.99 | 0.14 | 0.47 | 0.07 |
| *Unknown09_nifH070* | -0.15 | -0.99 | 0.61 | 0.01 |
| *Unknown10_nifH085* | -0.56 | 0.83 | 0.32 | 0.19 |
| *Unknown13_nifH113* | 0.89 | 0.46 | 0.87 | 0 |

| PCA Amino acids | | Correlation (p<0.05) | |
|---|---|---|---|
| Variable | % Contribution Dim 1 | Nitrogenase | Log *nifH* |
| Met | 25.84 | 0.54 | 0.45 |
| Thr | 39.39 | -0.53 | -0.85 |
| Lys | 34.77 | -0.45 | -0.48 |
| **PCA Soil properties** | | **Correlation (p<0.05)** | |
| Variable | % Contribution Dim 1 | Nitrogenase | Log *nifH* |
| C:N ratio | 8.76 | -0.27 | -0.41 |
| Cu | 9.82 | 0.63 | 0.64 |
| Fe | 11.64 | 0.69 | 0.43 |
| Mn | 12.95 | 0.65 | 0.50 |
| $NH_4$ | 11.52 | 0.60 | 0.30 |
| $NO_2$+$NO_3$ | 9.68 | 0.48 | 0.13 |
| pH | 10.72 | -0.78 | -0.57 |
| Sand | 7.38 | -0.76 | -0.59 |
| Silt | 2.17 | 0.36 | 0.44 |
| Zn | 11.73 | 0.66 | 0.71 |

Figure 2.3.13. Dimension reduction of variables for linear regression models of nitrogen-fixing bacteria. A) NMDS was applied for ASV above 1% of relative abundance, values of ASV envfit is displayed on the right. B) PCA of amino acids and C) Soil properties correlated with the abundance of *nifH* and nitrogenase activity (ρ>0.3, p<0.05, Pearson method). Contribution and correlation values are displayed.

Lastly, linear models were made to identify the variables that explained the variance of the nitrogenase activity and abundance of the *nifH* gene. These models were constructed with soil properties and amino acids correlated to the enzymatic activity and abundance of *nifH* gene (Pearson method ρ>0.3, p<0.05) and the most abundant ASV (>1%) (Figure 2.3.13). According to these models, while the abundance of the *nifH* gene is explained by the amino acids Lys, Met, and Thr (Explained variance= 58.11%, p<0.01); the nitrogenase activity depends on soil properties such as C:N ratio, micronutrients, nitrogen sources, SM (Explained variance= 44.54%, p<0.05), and abundant ASV (Explained variance= 22.99%, p<0.05) where the most common genera were *Bacillus, Escherichia, Nostoc*, *Pseudomonas*, and *Unknown* taxa (Figure 2.3.12).




| PCA Amino acids | | Correlation (p<0.05) | |
|---|---|---|---|
| Variable | % Contribution Dim 1 | APM | Log *phoD* |
| Ala | 12.41 | -0.18 | -0.31 |
| Arg | 18.1 | 0.42 | 0.61 |
| His | 22.01 | 0.53 | 0.64 |
| Leu | 1.07 | -0.54 | -0.29 |
| Lys | 15.06 | -0.76 | -0.61 |
| Met | 10.57 | 0.37 | 0.45 |
| Thr | 14.38 | -0.37 | -0.74 |
| Val | 6.4 | 0.36 | 0.64 |

| PCA Soil properties | | Correlation (p<0.05) | |
|---|---|---|---|
| Variable | % Contribution Dim 1 | APM | Log *phoD* |
| % Sand | 4.62 | -0.67 | -0.75 |
| C:N ratio | 8.35 | -0.57 | -0.44 |
| Cu | 7.8 | 0.81 | 0.72 |
| EC | 9.62 | 0.68 | 0.29 |
| Fe | 9.71 | 0.87 | 0.53 |
| Mn | 11.05 | 0.84 | 0.62 |
| N | 8.19 | 0.59 | 0.38 |
| $NH_4$ | 10.11 | 0.78 | 0.47 |
| $NO_2$+$NO_3$ | 8.65 | 0.72 | 0.23 |
| P | 3.11 | -0.59 | -0.31 |
| pH | 8.99 | -0.79 | -0.66 |
| Zn | 9.8 | 0.82 | 0.80 |

Figure 2.3.14. Dimension reduction of variables for linear regression models of organic phosphorus mineralizing bacteria. A) PCA was applied for amino acids and B) Soil properties correlated with the abundance of *phoD* and alkaline phosphomonoesterase activity (ρ>0.3, p<0.05, Pearson method). Contribution and correlation values are displayed.

OPMB community was composed mainly of the phyla Firmicutes (*Bacillus* -429 ASV- and *Clostridium* -1 ASV-*),* Actinobacteria (*Streptomyces* -11 ASV-*)*, and *Unknown* taxa (12 ASV). At genera level, we can appreciate a higher abundance of the *Unknown* taxa and lower from *Bacillus* in the rhizosphere of mature agaves, while *Bacillus* had up to 97.72% of relative abundance in bulk soil (Figure 2.3.11). However, if we compare the bacterial communities at ASV level (ANOSIM, 999 permutations), we found that the plant stage is significant ($R^2$=0.208, $p<0.05$), as well as the presence of mature ($R^2$=0.906, $p<0.05$) and young plants ($R^2$=0.615, $p<0.05$) in the composition of the bacterial communities. Additionally, we found that the enzymatic activity of APM is higher in both mature (407.72 ± 5.63 µg pNP $g^{-1}$ $h^{-1}$) and young lechuguillas (378.03 ± 14.15 µg pNP $g^{-1}$ $h^{-1}$) than bulk soil (328.90 ± 9.64 µg pNP $g^{-1}$ $h^{-1}$) ($p<0.01$). Moreover, like NFB communities, only the mature plants had a higher abundance of *phoD* gene (8.68 ± 0.05 log copies *phoD* $g^{-1}$ $soil^{-1}$), while young agaves and bulk soil had around 7 log copies *phoD* $g^{-1}$ $soil^{-1}$ ($p<0.01$, Figure 2.3.12).

Variables that explain the APM activity include the amino acids: Ala, Arg, His Leu, Lys, Met, Thr, and Val (Explained variance= 36.16%, $p<0.01$); and soil properties such as micronutrients, dissolved phosphorus, pH, EC, and sand (Explained variance= 33.30%, $p<0.01$) (Figure 2.3.14). Besides, the same variables explain about 61.16% of the variance in the abundance of *phoD* ($p<0.01$, Figure 2.3.12).

## 2.4 Discussion

The objective of this project was to understand the influence of the amino acids released by *A. lechuguilla* T. in the composition and the functional profiles of the rhizobacterial community, especially in the enzymatic activity and abundance of NFB and OPMB. We found that, through certain amino acids in the root exudates, the plants stimulated the accumulation of nutrients in the soil and enhanced the selection of bacteria with plant-growth promoting traits. Furthermore, the presence of some of these compounds in the soil may be due to the influence of rhizobacteria. These results help us further to understand interactions between plants and the rhizosphere microbiome.

### 2.4.1 *A. lechuguilla* T. accumulated nutrients in the soil and shaped the rhizobacterial community

The arid lands are known for their elevated evapotranspiration rates, high temperatures, poorly developed soils, and low nutrient availability (Gamalero *et al.* 2020). However, organisms that inhabit these ecosystems, such as lechuguilla, show remarkable tolerance against these conditions. For example, the results revealed an increase of SM, TOC, N forms, and micronutrient concentration in the rhizosphere of mature and young stages of lechuguilla, which could be influencing the composition of the rhizobacterial community (Figure 2.3.5A). This phenomenon is known as fertility island (Granados-Sánchez *et al.* 2011), and according to this, the plants in arid environments accumulate organic matter and nutrients, retain soil moisture, and provide shelter to other organisms (Bashan and De-Bashan, 2010). Moreover, the exudates from the root of lechuguilla and patches-like growth had been associated with the retention and nutrient cycling (Reyes, 1981; Barbosa-Biones *et al.* 2019). In addition, the organic acids in the root exudates of lechuguilla, had shown a reductive effect over the pH in the soil, increasing the availability of nutrients (Mapelli *et al.* 2018).

Under these conditions, Proteobacteria was the most abundant taxa in the rhizosphere of lechuguilla. As stated in previous works, lechuguilla from the

Chihuahuan desert favored a rhizosphere with plenty of beneficial taxa from Proteobacteria, such as *Bradyrhizobia*, *Microvirga*, *Rhizobia*, and bacteria from the order Sphingomonadales (López-Lozano *et al.* 2020). In addition, even though Bacteroidetes was not one of the most abundant taxa, its relative abundance in the rhizosphere was twice than in bulk soil (Figure 2.3.3), a characteristic found previously in the Meseta subregion of the Chihuahuan desert (Medina-de la Rosa *et al.* 2021). The phylum Bacteroidetes includes beneficial bacteria such as *Flavisolibacter*, a genus able to subdue heavy metal oxidation and accelerate nutrient dynamics in the rhizosphere, thus promoting plant-growth (Liu *et al.* 2020). The phylum Bacteroidetes, along with Acidobacteria, Chloroflexi, and Firmicutes, had been identified as the community that replace the dominant Proteobacteria in the late stages of leguminous shrubs (Na *et al.* 2018). In lechuguilla these phyla along with *Entotheonellaeota* and *FBP*, showed an increment in their relative abundance in the rhizosphere. In contrast, we found a higher abundance of Actinobacteria in bulk soil communities. This phylum is common and dominant in soils (Buée *et al.* 2009), and some of its taxa are known to have mechanisms for the tolerance against alkaline soils, salinity, high temperatures, and formation of spores to overcome the scarce availability of nutrients (Ventura *et al.* 2007; Barka *et al.* 2016), conditions that were exacerbated in the bulk soil samples as RDA showed, since samples seemed to be related to higher C:N ratio and pH (Figure 2.3.5A).

As we can observe, the composition of the bacterial communities was greatly determined by the soil properties, influence which is duplicated since soil properties are also significant for the concentration of the root exudates (Kamilova *et al.* 2006), crucial factors for the selection of the rhizobacterial communities (De-la-Peña, Badri and Loyola-Vargas 2012).

### 2.4.2 The amino acids served as nutrient sources and aid against environmental stress.

Root exudates are essential elements in the conformation of the rhizobacterial communities (Huang *et al.* 2014), and among its components, the amino acids are some of the most critical by their multiple uses as metabolic intermediaries, signals,

structural components, and sources of C and N for plants and microorganisms (Moe 2013).

However, the composition of root exudates changes according to different variables, including abiotic stress, water availability, plant species or cultivar, growth stage, and interactions (Rengel 2002; Bertin, Yang and Weston 2003; Chaparro, Badri and Vivanco 2014; Bobille *et al.* 2019; Balasubramanian *et al.* 2020). In lechuguilla, we found differences in the amino acids profile according to the growth stage that influence the bacterial community. For example, Tyr was related with the composition of the rhizobacterial community related to young and mature lechuguillas, while Arg was mostly related with the mature plants. It has been proposed that Arg, Asp, and Gln are used as nitrogen storage and the transport of compounds inside the plants (Hildebrandt *et al.* 2015). Arginine, however, is the only source of urea, and its catabolism is essential for the mobilization of N through the plant (Witte, 2011). The concentration of this amino acid was associated with the late growth stage in underground tissues of plants of *Dioscorea opposite* and *Rumex acetosa* (Bausenwein *et al.* 2001; Ninomiya *et al.* 2004), and as nutrient storage, it is degraded during the senescence (Witte, 2011). Also, Arg, Met, and Ser increased in the root exudates of rice in the reproductive stage (Jackson and Ilamurugu, 2014). Though the accumulation of Arg in mature plants and reproductive stages has been registered, little is known about its influence over the bacterial community. Nonetheless, there are reports about chemotaxis effect over strains of *Bacillus* and *Ralstonia* or the inhibition of *Stenotrophomonas* (Elhalag *et al.* 2015; Feng *et al.* 2018), so further research is needed to clarify how the rhizosphere of lechuguilla seemed to enhance the presence of genera from Proteobacteria and Gemmatimonadetes.

Stress in plants has been related to the production of amino acids (Naylor and Coleman-Derr 2018). For example, high concentrations of Tyr in plants had been associated with stressful conditions or late growth stages (Schenck and Maeda 2018), which could explain its abundance in both stages of lechuguilla and its importance in the bacteria recruitment since, in the rhizobacterial communities, the

functions related to its metabolism were more abundant than in bulk soil. Similarly, Lys and Thr, which were at higher concentration in the rhizosphere of young plants, could be the response of a younger stages of lechuguilla exposed to stressful conditions (Obata and Fernie 2012). It has been proposed that plants increment the production of Lys, Thr, and Met for its use as osmolytes or alternative donors for the mitochondrial electron transport chain and the production of adenosine triphosphate (ATP) when C sources are scarce (Ishizaki *et al.* 2006; Joshi *et al.* 2010; Obata and Fernie 2012). Though the physiological response of lechuguilla to stressful conditions may need more investigation. Therefore, the production of Lys and Thr in young plants may help them tolerate arid conditions.

Nevertheless, we could associate the presence of some amino acids in the soil to the influence of the microorganisms and the abiotic conditions (Phillips *et al.* 2004). For example, it has been proved that the pace at which the amino acids are extracted depends significantly on the soil moisture (Ma *et al*. 2017), so the low availability of water in the arid lands may explain, at some level, the accumulation of these molecules in the arid ecosystems. Furthermore, even though there was no difference between samples in the concentration of Ala, it was one of the most abundant amino acids. To explain this behavior, we could rely on the study of other plants, since it has been reported that along with Val, Pro, Asp, and Glu, Ala is released by barley under Zn deficiency and aids to the selection of bacteria able to enhance its availability (Rengel 2015). Zinc is a yield-limiting micronutrient readily adsorbed into insoluble forms in calcareous soils such as those from the Chihuahuan desert (Weldua, Haileb and Habtegebrieb 2012). It is possible that the concentration of some of these amino acids in the rhizosphere of younger lechuguilla may be related to the lower availability of Zn, than in mature plants. Moreover, some of the strains related to the Zn solubilization are among the beneficial genera *Azospirillum*, *Bacillus*, *Rhizobium*, and *Pseudomonas* (Kamran *et al*. 2017; Eshaghi *et al.* 2019), which had higher relative abundance in the rhizosphere of lechuguilla.

Furthermore, another osmolyte associated with low SM and salinity tolerance is Pro which aids the plant to keep the cell turgor, prevent electrolyte leakage, and even

maintain reactive oxygen species at standard rates (Hayat *et al.* 2012). In the rhizosphere of lechuguilla, according to the PCA, higher concentration of Pro was related to young agaves (Figure 2.3.2). However, this mechanism may also be implemented by bacteria exposed to the harsher conditions of the arid land (Bobille *et al.* 2019), like those in bulk soil, since they do not have the shelter from the plant. As such, it may explain the higher abundance of functions related to the production of Pro in bulk soil (Figure 2.3.10). On the same note, other metabolites used to increase their tolerance against water deficiency are betaine, glutamine, trehalose, and exopolysaccharides (Paul and Nair 2008; Hayat *et al*. 2012; Vurukonda *et al*. 2016), all of which were enhanced in the communities from bulk soil where could be related with the increased abundance of Actinobacteria, a phylum which includes several mechanisms for drought tolerance (Niu *et al.* 2021).

Lastly, the lower concentration of amino acids in the rhizosphere of mature lechuguilla than young plants or bulk soil could be attributed to a different strategy of the plant stage. It has been proposed that younger plants release higher amounts of amino acids than in late stages. Pereira *et al*. (2020) proposed that as younger plants require high amounts of nutrients for biomass production, so the nutrient transport patterns may favor the release of root exudates (Pereira *et al.* 2020). For example, in soils from successional vegetation of fast growth, Ala, Pro, and Glu were found in higher concentrations than mature ecosystems (Henry and Jefferies 2002), whose higher concentrations were related to the rhizosphere of young lechuguilla samples.

As we can see, the production of some amino acids may respond to abiotic stimuli; however, further studies may be needed to understand the physiological changes in the plants that permit the release of specific components in the root exudates, as well as its influence on the bacterial communities.

### 2.4.3 The rhizobacterial community of lechuguilla heightened plant-growth-promoting functions and the tolerance functions of the communities from bulk soil.

The bulk soil communities are exposed to environmental conditions which may be harsher without the aid of a nursing plant (Bashan and De-Bashan 2010). In addition to osmolytes, we found that the potential functions of the bacterial community in bulk soil include other mechanisms to tolerate arid conditions. For example, the intense UV radiation in the arid ecosystems is harmful to bacteria, as they can die by exposure damage in its structures (Vélez *et al.* 2018). However, bacteria have functions to preserve the integrity of their DNA, such as nucleotide excision repair (Rouws *et al.* 2013), which we found to be in higher abundance in bulk soil samples than in young plant rhizosphere and lower still in mature communities (Figure 2.3.9). This difference could evidence the tolerance of the bacterial communities in bulk soil and how the host protection to the rhizospheric microbiota grows with the plant stage.

Another mechanism of protection in the bacteria is the formation of biofilm which provides tolerance to bacteria in the soil against abiotic stress like pH and osmotic changes, evaporation, and UV radiation (Kokare *et al.* 2009). In our samples, bulk soil bacteria showed a higher abundance of biofilm-related functions than rhizobacterial communities (Figure 2.3.10), which could be explained by the high abundance of Actinobacteria in said samples. The phylum is widely spread on some of the most extreme ecosystems, like the Atacama Desert (Bull *et al.* 2016), and includes bacteria with multiple tolerance mechanisms to abiotic stress and competence against other bacteria, such as biofilm production, siderophores, and an ample variety of antibiotics (Rateb, Ebel and Jaspars 2018; El Othmany *et al.* 2021), functions that were more abundant in bulk soil than the rhizosphere of lechuguilla. Along with Actinobacteria, another abundant phylum in bulk soil was Chloroflexi, which had been reported to have an essential role in biofilm production (Kindaichi *et al.* 2012). Consequently, both of them could explain the abundance of biofilm related functions and it use as a strategy in bacterial communities exposed to arid conditions.

In the young lechuguilla rhizosphere, we found an increased potential biosynthesis of peptidoglycans, which supports the previous statement about the stress to which rhizobacteria of young lechuguillas are exposed. Glycans, exopolysaccharides, and lipopolysaccharides are structures attached to the surface of bacteria and biofilm components that work as an energy reserve, are involved in plant-bacteria communication, and thus, are related to the first stages of microbial infection of pathogens (Oldroyd 2013; Sánchez-Cañizares *et al.* 2017; Arnold *et al.* 2018; Chaliha *et al.* 2018) and beneficial bacteria alike (Wanke *et al.* 2021). For example, several strains of biofilm-producing *Pseudomonas* and *Bacillus* had shown mechanisms for nutrient solubilization, nitrogen fixation, production of phytohormones and siderophores, and enhanced the salinity tolerance of the host (Kasim *et al.* 2016; Haque *et al.* 2020). For rhizobia, glycans aid in the suppression of the immune response of plants and the establishment of bacteria, nodulation, and survival of the bacteroids (Gully *et al.* 2016). Also, in mature plants, the enhanced *quorum quenching*-related functions could be associated to the suppression of biofilm production and bacteria colonization (Azman *et al.* 2019), which suggests an extra resource of competition between the microbiome and biocontrol protection for the host.

On the same note, antibiotic-related functions and stimulants of immune response had higher relative abundance in the rhizosphere of lechuguilla than bulk soil. Antibiotics are competence mechanisms between microorganisms that could be used in the bacterial community to eliminate the contestants for nutrients or space (Crépin *et al.* 2012). However, the production of these compounds varies between species, abiotic conditions, and C sources (Coulthurst, Barnard and Salmond 2005). Some of the functions related to the production of antibiotics found in the rhizosphere of lechuguilla were phenazine (Mavrodi, Blankenfeldt and Thomashow. 2006), streptomycin (Ohnishi *et al.* 2008), and betalain, which is not an antibiotic but it does stimulate the immune defense system of the host (Moola and Diana 2019). Betalain favor the production of secondary metabolites and has similar Tyr precursor of EPS

(Steiner *et al.* 1999; Contreras-Llano *et al.* 2019), which could explain its influence over the rhizobacterial community of both lechuguilla stages.

Protection and communication mechanisms, as well as nutrition and transport were abundant functions in the bacterial communities of lechuguilla and bulk soil, however, there was a clear difference between the use of C sources in the rhizosphere and the scarce resources from bulk soil. For example, Chloroflexi could be partially responsible for the increased functions concerning the autotrophic metabolism to acquire nutrients. Photosynthesis is a mechanism of autotrophic organisms to convert light into chemical energy to produce organic carbon compounds (Blankenship 2010). Different types of bacteria perform this process, and taxa from Chloroflexi are some examples (Hoff and Deisenhofer 1997; Garrity *et al.* 2001). Which could explain the higher abundance of these functions in the bulk soil, a space where bacteria are exposed to UV radiation, but organic matter and root exudates are restricted.

Meanwhile, in the rhizosphere of lechuguilla, with the root exudates, bacteria had an ample source of C compounds in the form of sugars, organic acids, amino acids, and other metabolites that serve as chemo attractants and substrates (Zhalnina *et al*. 2018; Vives-Peris *et al.* 2020). This influence could be demonstrated in the rhizospheric samples, as functions related to the metabolism of carbon sources may evidence heterotrophic behavior. For example, propanoate and butanoate, which are associated to autotrophic processes and uptake of diverse carbon sources (Flores-Nuñez *et al.* 2020), showed a lower abundance of related functions to their metabolism in the rhizosphere of lechuguilla than bulk soil. Instead, the host-related communities, had a high relative abundance of functions for the metabolism of carbon sources from the citrate cycle, glyoxylate, and carboxylate (Figure 2.3.7), thus suggesting that this microbiome utilizes C sources from root exudates. Other carbon sources drawn on by rhizobacteria could be related to phosphorus compounds, which implies the importance of this element and its source; however, lower biosynthesis of fatty acids could be related to lower P availability (Geiger *et al.* 1999). The enhanced metabolism of sphingolipids in rhizobacterial communities

could also evidence the recycling of organic forms to obtain phosphorus. Sphingolipids are essential components of eukaryotic cells, and the ability to metabolize them had been related to the first stages of the infection of bacteria such as *Legionella*, *Pseudomonas*, *Neisseria*, *Staphylococci*, *Mycobacteria*, *Helicobacter*, or *Clostridia* (Rolando and Buchrieser 2019; Stankeviciute *et al.* 2019). Accordingly, our data suggest that rhizobacterial communities tend to use organic sources to obtain phosphorus, increasing the functions of mineralization and transporters of P, while in the bulk soil community, inorganic P solubilization, was favored (Figure 2.3.10). Furthermore, the same preference for organic P sources, over inorganic P, has been identified in *Bacillus* dominated soil in Pozas Azules and Churince in Cuatro Ciénegas, Mexico (Tapia-Torres *et al.* 2016).

Mineralization also had an essential role in the nourishment of the microbiome from the rhizobacterial communities. Denitrification is the respiration process in bacteria by which the soluble nitrogen forms ($NO_2$ and $NO_3$) are reduced to gas like nitric oxide (NO), nitrous oxide, and atmospheric nitrogen ($N_2$) (Giri *et al.* 2005). The enhancement of the abundance of taxa related to denitrification in the rhizosphere of lechuguilla, and plants in general, can be associated with the availability of organic matter and the consequent high rates of mineralization and denitrification (Christensen *et al.* 1990), a process facilitated by the $O_2$ from the water-logged roots and root exudates (Arth and Frenzel 2000), which agree with the higher SM and denitrification function in the rhizosphere of young lechuguilla than bulk soil.

Other PGPR traits enhanced in the rhizosphere of lechuguilla were related to the tolerance against abiotic stress and plant-growth promotion. The first corresponds to 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase, an enzyme able to degrade ethylene precursor ACC, a compound that can inhibit the elongation of the plants and promotes its senescence (Saravanakumar and Samiyappan 2007). The second trait, indole-3-pyruvic acid, is an auxin that stimulates the growth of biomass in the plants and has been found in *Acinetobacter, Azospirillum, Bacillus, Klebsiella, Pseudomonas,* and *Rhizobium* (Meena 2018); some of which were found in the rhizosphere of lechuguilla (Figure 2.3.11).

However, even though functional prediction is a valuable method to characterize the bacterial communities, it is essential to view these results cautiously since they are mainly based on theoretical assumptions (Kavamura *et al.* 2018).

### 2.4.4 Amino acids in the root exudates of lechuguilla influence enzymatic activity and abundance of NFB and OPMB

Other mechanisms enhanced in the rhizosphere of lechuguilla to obtain nutrients were those related to the nitrogen fixation and the mineralization of organic phosphorus, which are helpful for plant development and could be improved by the amino acids contained in the root exudates. Nitrogen fixation is an essential process for nutrient recycling and the sustenance of the plants, as it is one of the few forms in which nitrogen is incorporated into the soil (Gaby and Bucley 2014). As part of this enhanced process in the rhizosphere, the flavonoid-related functions were also increased. Flavonoids are compounds that stimulate the expression of Nod genes in rhizobia to start the symbiosis process with host plants, and even in no-nodulating plants are employed for chemotaxis of diazotrophic bacteria like *Azospirillum* (Coskun *et al.* 2020).

While a significant part of the NFB found were unknown taxa in the bulk soil and young lechuguilla communities, *Nostoc* was one of the most abundant genera. Cyanobacteria are organisms widely spread in arid lands, aided perhaps by its photoautotrophic metabolism and the formation of heterocysts to isolate the nitrogenase since these enzymes are oxygen-sensitive (Herrero, Muro-Pastor and Flores 2001). These traits could be highly beneficial under the oligotrophic root-free soil and the earlier stages of the lechuguilla, where the plant still has yet to develop its full size and is recruiting its microbiome.

The rhizobacterial community of mature lechuguilla had a broad diversity of diazotrophic bacteria, and according to *nifH* gene sequencing, the most abundant taxa were unknown bacteria, followed by genera from Proteobacteria like *Acidithiobacillus, Azospirillum, Burkholderia, Escherichia, Gluconobacter, Pseudomonas,* and *Sinorhizobium*, which had been reported to possess more plant-growth promotion traits, such as the production of phytohormones and osmolytes

(Mandon *et al.* 2003; Spaepen *et al.* 2008). For example, during the colonization of *Azospirillum*, the bacteria can reduce nitrate to produce NO, and this compound is associated with auxins pathway, which controls the root growth (Vacheron *et al.* 2013). Similar results were found in cacti from the Sonoran Desert, where *Azospirillum*, *Pseudomonas,* and *Rhizobium* were widely distributed (López-Lozano, Carcaño-Montiel, and Bashan 2016). However, most of the sequences were not related with the GenBank NCBI database.

Furthermore, according to the linear models, we found that the taxa recruited in the rhizosphere of lechuguilla and its soil physicochemical properties were related to the enhanced activity of NFB, explaining up to 67.53% of the variation (Figure 2.3.12). Among the most critical soil properties included in the model were C:N, Mn, Fe, Zn, and N inorganic forms. Nitrogen availability is probably easy to explain since the $N_2$ reduction is energy demanding and utilizes up to 16 ATP molecules per each molecule of $N_2$ reduced (Bahuguna and Pal 2011), so the bacteria need carbon sources balanced with the availability of nitrogen, which is reflected in the C:N ratio. The metallocluster in each nitrogenase component and its use as catalytic core, explains the influence of Fe, Mn, and the simultaneous increase of the activity with their concentrations (Byer *et al.* 2015). However, it has been found that concentrations above 0.2% of these elements can suppress the nitrogenase activity of diazotrophic bacteria (Gulten, Gonul and Sedat 2007; Karimi *et al.* 2021). As for the influence of the composition of the bacteria, it may hint at the importance in the selection of specific microbiota in the rhizosphere of the lechuguilla, where the most relevant ASV, were *Bacillus*, *Escherichia*, *Nostoc*, *Pseudomonas*, and *Unknown* taxa (Figure 2.3.13).

On the same note, even though a relation had been found between the nitrogenase activity and the amino acid concentration (Yousef *et al.* 2012), the results suggested that amino acid concentration was related to the abundance of the *nifH* gene. The amino acids included were Met, Lys, and Thr, with a matching contribution to the clustering of mature plant samples on one side and young lechuguilla rhizosphere and bulk soil on the other (Figure 2.3.13). Tough little is known about the effect of

bacteria in the production of amino acids by plants, it has been found that the inoculation of strains of *Agrobacterium, Bacillus, Pseudomonas,* and *Rhizobium* increased the concentration of amino acids like His, Arg, Asp, and Tyr in the roots and leaves of *Datura stramonium* (Rahmoune *et al.* 2019). In the case of *Pseudomonas*, it has been found that it could increase the exudation of amino acids in maize (Phillips *et al.* 2004). Likewise, the improvement of amino acids in the plants could be related to the diazotrophic activity of the microbiome, since the bacteria increase the availability of nitrogen for the host, the plants tend to storage C and N sources (Moe *et al.* 2013), as amino acids pools (Hildebrant *et al.* 2015). For example, Hutapea *et al*. (2018) observed that inoculation with rhizobia increased the concentration of proteins and amino acids like Met and His in leguminous plants. Thus, we could consider a direct influence of the bacteria with the amino acids in the rhizosphere since it had been observed that NFB release amino acids into the soil (González-López *et al.* 2005). This trait coincides with the high concentration of Arg, Asp, His, and Tyr with the pattern of the relative abundance of *nifH*, which seemed to have the high concentration/abundance in mature plants.

Up to this point, we have observed that the activity of enzymes is related to the concentration of nutrients available in the organic matter for the microbiome (Tapia-Torres *et al.* 2015), a pattern that may be just as necessary for the OPMB. The sequencing of the *phoD* gene revealed that most of the OPMB in our samples were from the phylum Firmicutes, and the most abundant by far was *Bacillus*. Firmicutes is a phylum known for its tolerance to extreme conditions (Galperin *et al.* 2013), and *Bacillus*, the most abundant genera found in our samples, is also a generalist widely spread in multiple environments and easily adaptable to different rough conditions (Torres *et al.*2018). One of its tolerance mechanisms is the ability to form spores, in which it can stay in dormancy for years till the conditions are manageable (Khanna, Lopez-Garrido and Pogliano 2020). Moreover, in the Chihuahuan Desert, *Bacillus* has been identified as some of the most abundant taxa, and to supply its high demand of phosphorus for their cellular wall, they can produce phosphatases (Tapia-Torres *et al.* 2016). In addition, even under the limiting availability of phosphorus,

*Bacillus* can use different forms of P oxidation, an advantage over other taxa in an oligotrophic environment like our study site (Tapia-Torres and Olmedo, 2018). This versatility, suggest a great importance of *Bacillus* in the availability of P in the Chihuahuan Desert.

Beyond the composition of the OPMB, other unknown factor may be responsible of the difference in the enzymatic activity and abundance of OPMB. In this regard, we found that APM activity was higher in the rhizosphere of both stages of lechuguilla, while the copy number of the *phoD* gene only was increased in the mature stage of the plant, and according to the linear model, soil physicochemical properties and amino acids composition are significant. Like nitrogenase, micronutrient concentration influenced APM activity and could relate to their role in the enzyme as a component of the catalytic centers (Zappa *et al.* 2001). Moreover, APM activity seemed to be restricted by a high C:N ratio, which implies unbalanced carbon and nitrogen concentration and thus a halt in the metabolism of the bacteria (Su *et al.* 2015); a concentration that may be related to the availability of amino acids (Ma *et al.* 2017). For example, it had been suggested that the microorganisms can metabolize Ala more easily (Ma *et al.* 2017), which could lead to the increase of the biomass and thus explain, its influence on the abundance of *phoD*. Besides, the more complex amino acids like Arg or Met could have favored the APM activity since they offer more nutrients than those with short C chains. As mentioned before, Arg is an essential nutrient store, and its versatile metabolism can produce several compounds like NO, urea, ATP, ornithine, and other amino acids (Morris 2004; Witte 2011); which had been found to accumulate in vegetal tissues as a response to P deficient conditions and inhibition of the plant growth (Rabe and Lovatt, 1986). In turn, Met is involved in the initiation and stability of proteins; and provides them with resistance to oxidation (Si *et al.* 2020). Even an increased activity of the APM had been reported with Met as substrate; however, the mechanisms are not precise (Spohn, Ermark and Kuzyakov 2013).

Likewise, the negative relationship that suggest the amino acids with the OPMB in the components of the model (Figure 2.3.14), had been reported in previous works,

though the mechanisms remain unclear. For example, Tavakoli *et al*. (2019) found that Leu, Ala, and Val had an inhibitory effect on alkaline phosphatases. Bodansky (1948) proposed that while low concentrations of amino acids like Gly increase the APM activity, higher doses may decrease it and suggested that the inhibition could be related to the presence of the free carboxyl and amino groups of the biomolecule since the esterification of the carboxyl groups reverted the inhibition of the phosphatases (Bodansky 1948). Others suggest that amino acids may prevent water access to the enzyme which is essential for the hydrolysis of organic phosphorus (Fernley and Walker 1970). Still, further research is needed to test the influence of amino acids over the enzymatic activity and abundance of PGPR.

## 2.5 Conclusions

We studied the effect of amino acids over bacterial composition in the rhizosphere of *A. lechuguilla* T. from the Meseta subregion in the Chihuahuan Desert and the nutrient-improvement bacteria activity. To understand the mechanisms that support the bacterial community and its interaction with its host, we studied the theoretical functions of the community and the diversity, enzymatic activity, and abundance of NFB and OPMB. In this regard, we found that lechuguilla plants improve the nutrient availability in the soil. As seen in different plants, root exudates contain a significant percentage of the assimilated carbon and nitrogen, thus shaping the conditions that recruit the rhizobacterial community.

In addition, the structure and functions of the bacterial community could have been influenced by the amino acids, which reflected the growth stage and abiotic conditions at which lechuguilla was exposed. Among these metabolites, we found that bulk soil communities seemed to be influenced by Leu and were potentially equipped to tolerate abiotic stress and oligotrophic conditions. At the same time, the amino acids Arg and Tyr were related to the rhizobacterial community of lechuguilla, where the most abundant phyla were Proteobacteria and Bacteroidetes, with several plant growth-promoting traits in the community. Likewise, the most common mechanisms of the rhizobacteria were the mineralization of organic phosphorus, nitrogen-fixing activity, antibiotic production, phytohormones, and *quorum quenching*, while bulk soil community seem to favor functions related to the tolerance against abiotic stress.

Lastly, we found that the enzymatic activity and abundance of nutrient-improvement bacteria like NFB and OPMB were increased in the rhizosphere of lechuguilla, mostly the mature stage, and that this response was related to soil properties, amino acids composition, and the bacterial community composition. OPMB was mainly composed of *Bacillus*, a tolerant taxon with several mechanisms to overcome the harsh conditions of the Chihuahuan Desert, while NFB included diazotrophic organisms from Proteobacteria, Cyanobacteria, and Firmicutes; however, a large

part of the diazotrophic bacteria could not be identified, especially in the bulk soil samples; which demonstrate the bast unexplored diversity. Both functional groups showed influence by nutrient composition, especially nitrogen forms and micronutrients. As for the amino acids, Arg and Met were some of the most relevant for the enhancement of the enzymatic activity of NFB and OPMB; and even though we can hint at their importance as structural components the mechanisms for the increase of the enzymatic activity are not completely clear, thus the need to further research to clarify their role in the bacterial communities. Besides, since the rhizosphere is subjected to the depositions of plants, it is possible that some of the amino acids in the soil came from the biosynthesis process from bacteria. This work provides valuable information since it tries to conjugate the influence of the abiotic variables over the root exudates released by the plant and the resulting recruited bacteria and the benefits that offer the microbiome to the host so it can tolerate the arid conditions and improve the nutrient availability (Figure 2.5.1).

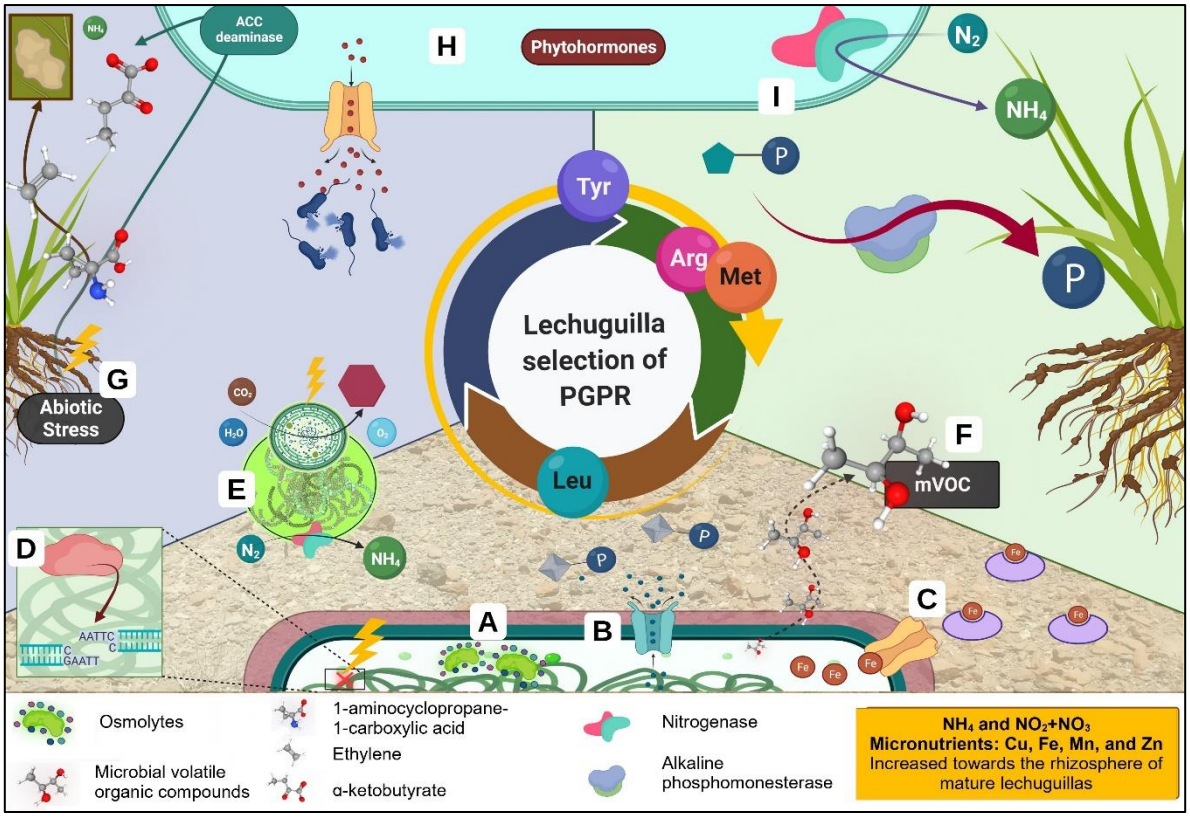



Figure 2.5.1. Selection of PGPR in the rhizosphere of *Agave lechuguilla* T. by amino acids from its root exudates. Exposed to less water and nutrient availability, the bacterial communities seemed to be influenced by the presence of Leucine (Leu) and have tolerance mechanisms like the production of osmolytes, biofilm production (A), phosphorus (P) solubilization (B) and siderophore transport for iron (Fe) obtention (C). The communities associated with young lechuguillas share some similarities with those in bulk soil such as nucleotide excision repair mechanism (**D**) and the increased abundance of *Nostoc* a Cyanobacteria able to perform photosynthesis and nitrogen fixation (**E**). In addition, the production of volatile organic compounds (mVOC) of bulk soil communities is also present in the bacterial community in the rhizosphere of mature lechuguillas (**F**). The release of amino acids like Tyrosine (Tyr) and Arginine (Arg) from lechuguilla may improve the selection of PGPR with mechanisms that favor the production of ACC deaminase to prevent ethylene production as response to abiotic stress (G); or the production of antibiotics and phytohormones (H), especially in the more susceptible stages of young lechuguilla. In mature lechuguilla, the high abundance and activity of nitrogenase and alkaline phosphomonoesterase (I) seemed to be related to Arg, Methionine (Met); and the concentration of micronutrients and nitrogen sources available in the rhizosphere, which increased with the plant stage. Furthermore, among OPMB, *Bacillus* was a dominant genus that gradually decreased with the growth stages of lechuguilla. Nitrogen fixing bacteria associated to mature lechuguilla were among Proteobacteria like *Pseudomonas*, *Azospirillum*, *Gluconobacter*, *Sinorhizobium*, and *Acidithiobacillus*.

## 2.6 References


Abdel-Lateif K, Bogusz D, Hocher V. The role of flavonoids in the establishment of plant roots endosymbioses with arbuscular mycorrhiza fungi, rhizobia and *Frankia* bacteria. Plant Signal Behav 2012;7:636–641.

Abu-Zaitoon Yousef. Ecological studies on nitrogen fixing bacteria from leguminous plants at the north of Jordan. African J Microbiol Res 2012;6:3656–61.

Altschul S, Gish W, Miller W *et al*. Basic local alignment search tool. J Mol Biol 1990;215:403–10.

Arnold MFF, Penterman J, Shabab M *et al*. Important Late-Stage Symbiotic Role of the *Sinorhizobium meliloti* Exopolysaccharide Succinoglycan. Brun Y V. (ed.). J Bacteriol 2018;200:1–17.

Arth I, Frenzel P. Nitrification and denitrification in the rhizosphere of rice: The detection of processes by a new multi-channel electrode. Biol Fertil Soils 2000;31:427–35.

Ashburner M, Ball CA, Blake JA *et al*. Gene Ontology: tool for the unification of biology. Nat Genet 2000;25:25–9.

Ayangbenro AS, Babalola OO. Reclamation of arid and semi-arid soils: The role of plant growth-promoting archaea and bacteria. Curr Plant Biol 2021;25:100173.

Azman AS, Mawang CI, Khairat JE *et al*. Actinobacteria—a promising natural source of anti-biofilm agents. Int Microbiol 2019;22:403–9.

Badri D V., Vivanco JM. Regulation and function of root exudates. Plant Cell Environ 2009;32:666–81.

Bahuguna RN, Pal M. Physiology of nitrogen fixation in legumes under elevated $CO_2$. In: Singh MP, Khetarpal S, Pandey R, *et al*. (eds.). Climate Change: Impacts and Adaptations in Crop Plants. Vol 2. New Delhi: Toay and Tomorrow, 2011, 213–28.

Balasubramanian V, Sur A, Nayak KK *et al*. Plant Root Exudates as Determinant of Rhizomicrobiome. In: Sharma SK, Singh UB, Sahu PK, *et al*. (eds.). Rhizosphere Microbes, Microorganisms for Sustainability. Singapore: Springer, Singapore, 2020, 105–26.

Barbosa-Briones E, Cardona-Benavides A, Reyes-Hernández H *et al*. Ecohydrological function of vegetation patches in semi-arid shrublands of central Mexico. J Arid Environ 2019;168:36–45.

Barka EA, Vatsa P, Sanchez L *et al*. Taxonomy, Physiology, and Natural Products of Actinobacteria. Microbiol Mol Biol Rev 2016;80:1–43.

Barret M, Morrissey JP, O'Gara F. Functional genomics analysis of plant growth-promoting rhizobacterial traits involved in rhizosphere competence. Biol Fertil Soils 2011;47:729–43.

Bashan Y, De-Bashan LE. Microbial Populations of Arid Lands and their Potential for Restoration of Deserts. Soil Biology and Agriculture in the Tropic. Berlin, Heidelberg.: Springer, Berlin, Heidelberg, 2010, 109–37.

Bausenwein U, Millard P, Thornton B *et al*. Seasonal nitrogen storage and remobilization in the forb Rumex acetosa. Funct Ecol 2001;15:370–7.

Baziramakenga R, Simard RR, Leroux GD. Determination of organic acids in soil extracts by ion chromatography. Soil Biol Biochem 1995;27:349–56.

Behnke R, Mortimore M eds. The End of Desertification? Berlin, Heidelberg: Springer Berlin Heidelberg, 2016.

Bertin C, Yang X, Weston LA. The role of root exudates and allelochemicals in the rhizosphere. Plant Soil 2003;256:67–83.

Blankenship RE. Early Evolution of Photosynthesis. Plant Physiol 2010;154:434–8.

Bobille H, Fustec J, Robins RJ *et al*. Effect of water availability on changes in root amino acids and associated rhizosphere on root exudation of amino acids in *Pisum sativum* L. Phytochemistry 2019;161:75–85.

Bodansky O. The inhibitory effects of DL-Alanine, L-Glutamic acid, L-Lysine, and L-Histidine on the activity of intestinal, bone and kidney phosphatases. J Biol Chem 1948;174:465–76.

Bokulich NA, Subramanian S, Faith JJ *et al*. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nat Methods 2013;10:57–9.

Bray RH, Kurtz LT. Determination of total, organic, and available forms of phosphorus in soils. Soil Sci 1945;59:39–46.

Buée M, De Boer W, Martin F *et al*. The rhizosphere zoo: An overview of plant-associated communities of microorganisms, including phages, bacteria, archaea, and fungi, and of some of their structuring factors. Plant Soil 2009;321:189–212.

Bull AT, Asenjo JA, Goodfellow M *et al*. The Atacama Desert: Technical Resources and the Growing Importance of Novel Microbial Diversity. Annu Rev Microbiol 2016;70:215–34.

Byer AS, Shepard EM, Peters JW *et al*. Radical S-Adenosyl-l-methionine Chemistry in the Synthesis of Hydrogenase and Nitrogenase Metal Cofactors. J Biol Chem 2015;290:3987–94.

Carvalhais LC, Dennis PG, Badri D V. *et al*. Activation of the Jasmonic Acid Plant Defence Pathway Alters the Composition of Rhizosphere Bacterial Communities. PLoS One 2013a;8:1–5.

Carvalhais LC, Dennis PG, Fan B *et al*. Linking Plant Nutritional Status to Plant-Microbe Interactions. Baxter I (ed.). PLoS One 2013b;8:e68555.

Chaliha C, Rugen MD, Field RA *et al*. Glycans as Modulators of Plant Defense Against Filamentous Pathogens. Front Plant Sci 2018;9, DOI: 10.3389/fpls.2018.00928.

Chaparro JM, Badri D V, Vivanco JM. Rhizosphere microbiome assemblage is affected by plant development. ISME J 2014;8:790–803.

Chen X, Jiang N, Chen Z *et al*. Response of soil phoD phosphatase gene to long-term combined applications of chemical fertilizers and organic materials. Appl Soil Ecol 2017;119:197–204.

Christensen S, Groffman P, Mosier A *et al*. Rhizosphere Denitrification; A Minor Process but Indicator of Decomposition Activity. In: Revsbech NP, Sørensen J (eds.). Denitrification in Soil and Sediment. Boston, MA: Springer US, 1990, 199–211.

Clark JS, Campbell JH, Grizzle H *et al*. Soil Microbial Community Response to Drought and Precipitation Variability in the Chihuahuan Desert. Microb Ecol 2009;57:248–60.

Contreras-Llano LE, Guerrero-Rubio MA, Lozada-Ramírez JD *et al*. First Betalain-Producing Bacteria Break the Exclusive Presence of the Pigments in the Plant Kingdom. Søgaard-Andersen L (ed.). MBio 2019;10, DOI: 10.1128/mBio.00345-19.

Coskun D, Britto DT, Shi W *et al*. How Plant Root Exudates Shape the Nitrogen Cycle. Trends Plant Sci 2017;22:661–73.

Coulthurst SJ, Barnard AML, Salmond GPC. Regulation and biosynthesis of carbapenem antibiotics in bacteria. Nat Rev Microbiol 2005 34 2005;3:295–306.

Crépin A, Barbey C, Cirou A *et al*. Biological control of pathogen communication in the rhizosphere: A novel approach applied to potato soft rot due to *Pectobacterium atrosepticum*. Plant Soil 2012;358:27–37.

De-la-Peña C, Badri D V, Loyola-Vargas VM. Plant Root Secretions and Their Interactions with Neighbors. In: Vivanco JM, Baluška F (eds.). Secretions and Exudates in Biological Systems. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, 1–26.

Elhalag KM, Emara HM, Messiha NAS *et al*. The Relation of Different Crop Roots Exudates to the Survival and Suppressive Effect of *Stenotrophomonas maltophilia* (PD4560), Biocontrol Agent of Bacterial Wilt of Potato. J Phytopathol 2015;163:829–40.

Eshaghi E, Nosrati R, Owlia P *et al*. Zinc solubilization characteristics of efficient siderophore-producing soil bacteria. Iran J Microbiol 2019;11:419–30.

Farrar J, Hawes M, Jones D *et al*. How roots control the flux of carbon to the rhizosphere. Ecology 2003;84:827–37.

Feng H, Zhang N, Du W *et al*. Identification of Chemotaxis Compounds in Root Exudates and Their Sensing Chemoreceptors in Plant-Growth-Promoting Rhizobacteria *Bacillus amyloliquefaciens* SQR9. Mol Plant-Microbe Interact 2018;31:995–1005.

Fernley HN, Walker PG. Inhibition of alkaline phosphatase by L-phenylalanine. Biochem J 1970;116:543–4.

Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 2011;39:W29.

Flores-Núñez VM, Fonseca-García C, Desgarennes D *et al*. Functional Signatures of the Epiphytic Prokaryotic Microbiome of Agaves and Cacti. Front Microbiol 2020;10:3044.

Freeman CE, Reid WH. Aspects of the reproductive biology of *Agave lechuguilla* Torr. Desert Plants 1985;7:75–80.

Gaby JC, Buckley DH. A comprehensive aligned *nifH* gene database: A multipurpose tool for studies of nitrogen-fixing bacteria. Database 2014;2014:1–8.

Galperin MY. Genome Diversity of Spore-Forming Firmicutes. Microbiol Spectr 2013;1:TBS-0015-2012.

Gamalero E, Bona E, Todeschini V *et al*. Saline and Arid Soils: Impact on Bacteria, Plants, and Their Interaction. Biology (Basel) 2020;9:116.

García-Oliva F, Merino A, Fonturbel MT *et al*. Severe wildfire hinders renewal of soil P pools by thermal mineralization of organic P in forest soil: Analysis by sequential extraction and 31P NMR spectroscopy. Geoderma 2018;309:32–40.

García-Robledo E, Corzo A, Papaspyrou S. A fast and direct spectrophotometric method for the sequential determination of nitrate and nitrite at low concentrations in small volumes. Mar Chem 2014;162:30–6.

Garrity GM, Holt JG, Castenholz RW *et al*. Phylum BVI. Chloroflexi phy. nov. Bergey's Manual® of Systematic Bacteriology. New York, NY: Springer New York, 2001, 427–46.

Geiger O, Rohrs V, Weissenmayer B *et al*. The regulator gene *phoB* mediates phosphate stress-controlled synthesis of the membrane lipid diacylglyceryl-N,N,N-trimethylhomoserine in *Rhizobium* (*Sinorhizobium*) *meliloti*. Mol Microbiol 1999;32:63–73.

Giri B, Giang PH, Kumari R *et al*. Microbial Diversity in Soils. In: Buscot F, Varma A (eds.). Microorganisms in Soils: Roles in Genesis and Functions. 1st ed. Germany: Springer, 2005, 19–58.

GNU P. Free Software Foundation. 2007.

González-López J, Rodelas B, Pozo C *et al*. Liberation of amino acids by heterotrophic nitrogen fixing bacteria. Amino Acids 2005;28:363–7.

Granados-Sánchez D, Sánchez-González A, Granados-Victorino RL *et al*. Ecología de la vegetación del Desierto Chihuahuense. Rev Chapingo Ser Ciencias For y del Ambient 2011;XVII:111–30.

Gransee A, Wittenmayer L. Qualitative and quantitative analysis of water-soluble root exudates in relation to plant species and development. J Plant Nutr Soil Sci 2000;163:381–5.

Gully D, Gargani D, Bonaldi K *et al*. A Peptidoglycan-remodeling enzyme is critical for bacteroid differentiation in *Bradyrhizobium* spp. during legume symbiosis. Mol Plant-Microbe Interact 2016;29:447–57.

Gulten OK, Gonul D, Sedat D. Influence of osmotic and metal stresses on nitrogenase activity of cyanobacteria isolated from paddy fields. African J Biotechnol 2007;6:1828–32.

Haque MM, Mosharaf MK, Khatun M *et al*. Biofilm Producing Rhizobacteria With Multiple Plant Growth-Promoting Traits Promote Growth of Tomato Under Water-Deficit Stress. Front Microbiol 2020;11, DOI: 10.3389/fmicb.2020.542053.

Hayat S, Hayat Q, Alyemeni MN *et al*. Role of proline under changing environments. Plant Signal Behav 2012;7:1456–66.

Hennion N, Durand M, Vriet C *et al*. Sugars en route to the roots. Transport, metabolism and storage within plant roots and towards microorganisms of the rhizosphere. Physiol Plant 2019;165:44–57.

Henry HAL, Jefferies RL. Free amino acid, ammonium and nitrate concentrations in soil solutions of a grazed coastal marsh in relation to plant growth. 2002;25:665–75.

Herrero A, Muro-Pastor AM, Flores E. Nitrogen Control in Cyanobacteria. J Bacteriol 2001;183:411–25.

Hildebrandt TM, Nunes Nesi A, Araújo WL *et al*. Amino Acid Catabolism in Plants. Mol Plant 2015;8:1563–79.

Hoff AJ, Deisenhofer J. Photophysics of photosynthesis. Structure and spectroscopy of reaction centers of purple bacteria. Phys Rep 1997;287:1–247.

Houri A, Machaka-Houri N. *Agave Lechuguilla* as a potential biomass source in arid areas. J Sustain Dev Energy, Water Environ Syst 2016;4:89–93.

Huang X-F, Chaparro JM, Reardon KF *et al*. Rhizosphere interactions: root exudates, microbes, and microbial communities. Botany 2014;92:267–75.

Huo C, Luo Y, Cheng W. Rhizosphere priming effect: A meta-analysis. Soil Biol Biochem 2017;111:78–84.

Hutapea PS, Abdullah L, Karti PDMH *et al*. Improvement of *Indigofera zollingeriana* Production and Methionine Content through Inoculation of Nitrogen-Fixing Bacteria. Trop Anim Sci J 2018;41:37–45.

Ishizaki K, Schauer N, Larson TR *et al*. The mitochondrial electron transfer flavoprotein complex is essential for survival of Arabidopsis in extended darkness. Plant J 2006;47:751–60.

Jackson KM, Ilamurugu K. Metabolic profiling of rice root exudates and its impact on rhizosphere microbial dynamics under aerobic conditions. Reswarch J Agric Sci 2014;5:777–81.

Jones DL, Darrah PR. Amino-acid influx at the soil-root interface of *Zea mays* L. and its implications in the rhizosphere. Plant Soil 1994;163:1–12.

Jones DL, Healey JR, Willett VB *et al*. Dissolved organic nitrogen uptake by plants - An important N uptake pathway? Soil Biol Biochem 2005;37:413–23.

Joshi V, Joung JG, Fei Z *et al*. Interdependence of threonine, methionine and isoleucine metabolism in plants: accumulation and transcriptional regulation under abiotic stress. Amino Acids 2010;39:933–47.

Kamilova F, Kravchenko L V., Shaposhnikov AI *et al*. Organic Acids, Sugars, and l -Tryptophane in Exudates of Vegetables Growing on Stonewool and Their Effects on Activities of Rhizosphere Bacteria. Mol Plant-Microbe Interact 2006;19:250–6.

Kamran S, Shahid I, Baig DN *et al*. Contribution of Zinc Solubilizing Bacteria in Growth Promotion and Zinc Content of Wheat. Front Microbiol 2017;8, DOI: 10.3389/fmicb.2017.02593.

Kanehisa M, Goto S, Sato Y *et al*. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res 2012;40:D109–14.

Karimi N, Goltapeh EM, Amini J *et al*. Effect of *Azospirillum zeae* and Seed Priming with Zinc, Manganese and Auxin on Growth and Yield Parameters of Wheat, under Dryland Farming. Agric Res 2021;10:44–55.

Kasim WA, Gaafar RM, Abou-Ali RM *et al*. Effect of biofilm forming plant growth promoting rhizobacteria on salinity tolerance in barley. Ann Agric Sci 2016;61:217–27.

Kavamura VN, Hayat R, Clark IM *et al*. Inorganic Nitrogen Application Affects Both Taxonomical and Predicted Functional Structure of Wheat Rhizosphere Bacterial Communities. Front Microbiol 2018;9:1074.

Khanna K, Lopez-Garrido J, Pogliano K. Shaping an Endospore: Architectural Transformations During *Bacillus subtilis* Sporulation. Annu Rev Microbiol 2020;74:361–86.

Kim D, Paggi JM, Park C *et al*. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 2019 378 2019;37:907–15.

Kindaichi T, Yuri S, Ozaki N *et al*. Ecophysiological role and function of uncultured Chloroflexi in an anammox reactor. Water Sci Technol 2012;66:2556–61.

Kokare CR, Chakraborty S, Khopade AN *et al*. Biofilm: Importance and applications. Indian J Biotechnol 2009;8:159–68.

Krogh A, Larsson B, von Heijne G *et al*. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes11Edited by F. Cohen. J Mol Biol 2001;305:567–80.

Lagesen K, Hallin P, Rødland EA *et al*. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007;35:3100–8.

Li B, Li Y-Y, Wu H-M *et al*. Root exudates drive interspecific facilitation by enhancing nodulation and N 2 fixation. Proc Natl Acad Sci 2016;113, DOI: 10.1073/pnas.1523580113.

Li T, Tao Q, Liang C *et al*. Complexation with dissolved organic matter and mobility control of heavy metals in the rhizosphere of hyperaccumulator *Sedum alfredii*. Environ Pollut 2013;182:248–55.

Lindsay WL, Norvell WA. Development of a DTPA Soil Test for Zinc, Iron, Manganese, and Copper. Soil Sci Soc Am J 1978;42:421–8.

Liu C, Lin H, Li B *et al*. Responses of microbial communities and metabolic activities in the rhizosphere during phytoremediation of Cd-contaminated soil. Ecotoxicol Environ Saf 2020;202:110958.

Liu Y, Tao J, Yan Y *et al*. Biocontrol efficiency of *Bacillus subtilis* SL-13 and characterization of an antifungal chitinase. Chinese J Chem Eng 2011;19:128–34.

López-Lozano NE, Carcaño-Montiel MG, Bashan Y. Using native trees and cacti to improve soil potential nitrogen fixation during long-term restoration of arid lands. Plant Soil 2016;403:317–29.

López-Lozano NE, Echeverría Molinar A, Ortiz Durán EA *et al*. Bacterial Diversity and Interaction Networks of *Agave lechuguilla* Rhizosphere Differ Significantly From Bulk Soil in the Oligotrophic Basin of Cuatro Cienegas. Front Plant Sci 2020;11:1028.

Ma H, Pei G, Gao R *et al*. Mineralization of amino acids and its signs in nitrogen cycling of forest soil. Acta Ecol Sin 2017;37:60–3.

Mandon K, Østerås M, Boncompagni E *et al*. The *Sinorhizobium meliloti* Glycine Betaine Biosynthetic Genes ( betICBA ) Are Induced by Choline and Highly Expressed in Bacteroids. Mol Plant-Microbe Interact 2003;16:709–19.

Mapelli F, Marasco R, Fusi M *et al*. The stage of soil development modulates rhizosphere effect along a High Arctic desert chronosequence. ISME J 2018;12:1188–98.

Mavrodi D V., Blankenfeldt W, Thomashow LS. Phenazine Compounds in Fluorescent *Pseudomonas* Spp. Biosynthesis and Regulation. Annu Rev Phytopathol 2006;44:417–45.

Medina-de la Rosa G, García-Oliva F, Alpuche-Solís ÁG *et al*. The nutrient-improvement bacteria selected by *Agave lechuguilla* T. and their role in the rhizosphere community. FEMS Microbiol Ecol 2021;97:1–14.

Meena VS. Role of Rhizospheric Microbes in Soil. Meena VS (ed.). Singapore: Springer Singapore, 2018.

Miltner A, Kindler R, Knicker H *et al*. Fate of microbial biomass-derived amino acids in soil and their contribution to soil organic matter. Org Geochem 2009;40:978–85.

Moe LA. Amino acids in the rhizosphere: From plants to microbes. Am J Bot 2013;100:1692–705.

Moola AK, Diana RKB. Elicitation as a means for enhanced plant secondary metabolites through hairy root system. J Adv Sci Res 2019;10:6–12.

Morris SM. Enzymes of Arginine Metabolism. J Nutr 2004;134:2743S-2747S.

Na X, Li X, Zhang Z *et al*. Bacterial community dynamics in the rhizosphere of a long-lived, leguminous shrub across a 40-year age sequence. J Soils Sediments 2018;18:76–84.

Naher UA, Othman R, Shamsuddin ZHJ *et al*. Growth Enhancement and Root Colonization of Rice Seedlings by *Rhizobium* and *Corynebacterium* spp. Int J Agric Biol 2009;11:586–90.

Naylor D, Coleman-Derr D. Drought Stress and Root-Associated Bacterial Communities. Front Plant Sci 2018;8:2223.

Nelson DW. Determination of ammonium in KCl extracts of soils by the salicylate method. Commun Soil Sci Plant Anal 1983;14:1051–62.

Ninomiya A, Murata Y, Tada M *et al*. Changes in Allantoin and Arginine Contents in *Dioscorea opposita* “Tsukuneimo” during the Growth. Engei Gakkai zasshi 2004;73:546–51.

Niu S, Gao Y, Zi H *et al*. The osmolyte-producing endophyte *Streptomyces albidoflavus* OsiLf-2 induces drought and salt tolerance in rice via a multi-level mechanism. Crop J 2021, DOI: 10.1016/j.cj.2021.06.008.

Nurk S, Meleshko D, Korobeynikov A *et al*. metaSPAdes: a new versatile metagenomic assembler. Genome Res 2017;27:824–34.

Obata T, Fernie AR. The use of metabolomics to dissect plant responses to abiotic stresses. Cell Mol Life Sci 2012;69:3225–43.

Ohnishi Y, Ishikawa J, Hara H *et al*. Genome Sequence of the Streptomycin-Producing Microorganism *Streptomyces griseus* IFO 13350. J Bacteriol 2008;190:4050–60.

Oksanen J, Blanchet FG, Friendly M *et al*. vegan: Community Ecology Package. 2020.

Oku S, Komatsu A, Tajima T *et al*. Identification of Chemotaxis Sensory Proteins for Amino Acids in *Pseudomonas fluorescens* Pf0-1 and Their Involvement in Chemotaxis to Tomato Root Exudate and Root Colonization. Microbes Environ 2012;27:462–9.

Oldroyd GED. Speak, friend, and enter: Signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol 2013;11:252–63.

El Othmany R, Zahir H, Ellouali M *et al*. Current Understanding on Adhesion and Biofilm Development in Actinobacteria. Falkinham J (ed.). Int J Microbiol 2021;2021:1–11.

Owen A., Jones D. Competition for amino acids between wheat roots and rhizosphere microorganisms and the role of amino acids in plant N acquisition. Soil Biol Biochem 2001;33:651–7.

Paliy O, Shankar V. Application of multivariate statistical techniques in microbial ecology. Mol Ecol 2016;25:1032–57.

Paul D, Nair S. Stress adaptations in a Plant Growth Promoting *Rhizobacterium* (PGPR) with increasing salinity in the coastal agricultural soils. J Basic Microbiol 2008;48:378–84.

Pereira LC, Bertuzzi Pereira C, Correia LV *et al*. Corn Responsiveness to *Azospirillum*: Accessing the Effect of Root Exudates on the Bacterial Growth and Its Ability to Fix Nitrogen. Plants 2020;9:923.

Perroni Y, García-Oliva F, Tapia-Torres Y *et al*. Relationship between soil P fractions and microbial biomass in an oligotrophic grassland-desert scrub system. Ecol Res 2014;29:463–72.

Petersen TN, Brunak S, von Heijne G *et al*. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 2011 810 2011;8:785–6.

Phillips DA, Fox TC, King MD *et al*. Microbial Products Trigger Amino Acid Exudation from Plant Roots. Plant Physiol 2004;136:2887–94.

Poly F, Monrozier LJ, Bally R. Improvement in the RFLP procedure for studying the diversity of nifH genes in communities of nitrogen fixers in soil. Res Microbiol 2001;152:95–103.

Powell S, Szklarczyk D, Trachana K *et al*. eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. Nucleic Acids Res 2012;40:D284–9.

Punta M, Coggill PC, Eberhardt RY *et al*. The Pfam protein families database. Nucleic Acids Res 2012;40:D290–301.

Qiao Q, Wang F, Zhang J *et al*. The Variation in the Rhizosphere Microbiome of Cotton with Soil Type, Genotype and Developmental Stage. Sci Rep 2017;7:3940.

Rabe E, Lovatt CJ. Increased Arginine Biosynthesis during Phosphorus Deficiency. Plant Physiol 1986;81:774–9.

Ragot SA, Kertesz MA, Bünemann EK. phoD Alkaline Phosphatase Gene Diversity in Soil. Voordouw G (ed.). Appl Environ Microbiol 2015;81:7281–9.

Rahmoune B, Zerrouk IZ, Bouzaa S *et al*. Amino acids profiling in *Datura stramonium* and study of their variations after inoculation with plant growth promoting Rhizobacteria. Biologia (Bratisl) 2019;74:1373–83.

Rateb ME, Ebel R, Jaspars M. Natural product diversity of actinobacteria in the Atacama Desert. Antonie Van Leeuwenhoek 2018;111:1467–77.

Ray S, Mishra S, Bisen K *et al*. Modulation in phenolic root exudate profile of Abelmoschus esculentus expressing activation of defense pathway. Microbiol Res 2018;207:100–7.

R Core Team. R: A language and environment for statistical computing. 2020.

Rees DC, Howard JB. Nitrogenase: standing at the crossroads. Curr Opin Chem Biol 2000;4:559–66.

Rengel Z. Genetic control of root exudation. Plant Soil 2002 2451 2002;245:59–70.

Rengel Z. Availability of Mn, Zn and Fe in the rhizosphere. J soil Sci plant Nutr 2015;15:0–0.

Reyes R. La Lechuguilla (*Agave Lechuguilla*) Como Importante Factor En La Formación Del Suelo de Las Zonas Áridas. México, D.F., 1981.

Reynolds JF, Smith DMS, Lambin EF *et al*. Global Desertification: Building a Science for Dryland Development. Science (80- ) 2007;316:847–51.

Rolando M, Buchrieser C. A Comprehensive Review on the Manipulation of the Sphingolipid Pathway by Pathogenic Bacteria. Front Cell Dev Biol 2019;7:168.

Rouws LFM, Meneses CHSG, Hemerly AS *et al*. Nucleotide excision repair and photoreactivation in sugarcane endophyte *Gluconacetobacter diazotrophicus* strain PAL5. African J Microbiol Res 2013;7:5284–7.

Rudi K, Skulberg OM, Larsen F *et al*. Strain characterization and classification of oxyphotobacteria in clone cultures on the basis of 16S rRNA sequences from the variable regions V6, V7, and V8. Appl Environ Microbiol 1997;63:2593–9.

Sánchez-Cañizares C, Jorrín B, Poole PS *et al*. Understanding the holobiont: the interdependence of plants and their microbiome. Curr Opin Microbiol 2017;38:188–96.

Saravanakumar D, Samiyappan R. ACC deaminase from *Pseudomonas fluorescens* mediated saline resistance in groundnut (*Arachis hypogea*) plants. J Appl Microbiol 2007;102:1283–92.

Schenck CA, Maeda HA. Tyrosine biosynthesis, metabolism, and catabolism in plants. Phytochemistry 2018;149:82–102.

Schmidt JE, Kent AD, Brisson VL *et al*. Agricultural management and plant selection interactively affect rhizosphere microbial community structure and nitrogen cycling. Microbiome 2019;7:146.

Schubert CJ, Nielsen B. Effects of decarbonation treatments on δ13C values in marine sediments. Mar Chem 2000;72:55–9.

SEMARNAT. Norma Oficial Mexicana Nom-021-Semarnat-2000 Que establece las especificaciones de fertilidad, salinidad y clasificación de suelos, estudio, muestreo y análisis. Diario Oficial de la Federación 2002:85.

Si Z, Guan N, Zhou Y *et al*. A Methionine Sulfoxide Reductase B Is Required for the Establishment of *Astragalus sinicus–Mesorhizobium* Symbiosis. Plant Cell Physiol 2020;61:1631–45.

Simonet P, Grosjean MC, Misra AK *et al*. *Frankia* genus-specific characterization by polymerase chain reaction. Appl Environ Microbiol 1991;57:3278–86.

Spaepen S, Dobbelaere S, Croonenborghs A *et al*. Effects of *Azospirillum brasilense* indole-3-acetic acid production on inoculated wheat plants. Plant Soil 2008;312:15–23.

Spohn M, Ermak A, Kuzyakov Y. Microbial gross organic phosphorus mineralization can be stimulated by root exudates – A 33P isotopic dilution study. Soil Biol Biochem 2013;65:254–63.

Stankeviciute G, Guan Z, Goldfine H *et al*. *Caulobacter crescentus* adapts to phosphate starvation by synthesizing anionic glycoglycerolipids and a novel glycosphingolipid. MBio 2019;10, DOI: 10.1128/MBIO.00107-19/SUPPL_FILE/MBIO.00107-19-S0001.DOCX.

Steiner U, Schliemann W, Böhm H *et al*. Tyrosinase involved in betalain biosynthesis of higher plants. Planta 1999 2081 1999;208:114–24.

Su J-Q, Ding L-J, Xue K *et al*. Long-term balanced fertilization increases the soil microbial functional diversity in a phosphorus-limited paddy soil. Mol Ecol 2015;24:136–50.

Tabatabai MA, Bremner JM. Use of p-nitrophenyl phosphate for assay of soil phosphatase activity. Soil Biol Biochem 1969;1:301–7.

Tao Q, Hou D, Yang X *et al*. Oxalate secretion from the root apex of *Sedum alfredii* contributes to hyperaccumulation of Cd. Plant Soil 2016;398:139–52.

Tapia-Torres Y, Elser JJ, Souza V *et al*. Ecoenzymatic stoichiometry at the extremes: How microbes cope in an ultra-oligotrophic desert soil. Soil Biol Biochem 2015;87:34–42.

Tapia-Torres Y, Olmedo-Álvarez G. Life on Phosphite: A Metagenomics Tale. Trends Microbiol 2018;26:170–2.

Tapia-Torres Y, Rodríguez-Torres MD, Elser JJ *et al*. How to live with phosphorus scarcity in soil and sediment: Lessons from bacteria. Appl Environ Microbiol 2016;82:4652–62.

Tavakoli R, Rastegar S, Khaledi M *et al*. Investigating effect of Amino Acids Leucine, Valine and Alanine on Alkaline Phosphatase Activity of purified Acetone fractions of Sweet Lemon, Garlic and Onion. Int J Ayurvedic Med 2019;10:34–8.

Torres MDR, Alcaraz LD, Souza V *et al*. Single Genus Approach to Understanding Bacterial Diversity, Niche, Distribution, and Genomics: The *Bacillus* in Cuatro Ciénegas. In: Souza V, Eguiarte LE, Olmedo-Álvarez G (eds.). Cuatro Ciénegas Ecology, Natural History and Microbiology. Switzerland: Springer, 2018, 103–12.

Trame A-M, Coddington AJ, Paige KN. Field and genetic studies testing optimal outcrossing in *Agave schottii*, a long-lived clonal plant. Oecologia 1995;104:93–100.

Turner S, Pryer KM, Miao VPW *et al*. Investigating Deep Phylogenetic Relationships among Cyanobacteria and Plastids by Small Subunit rRNA Sequence Analysis. J Eukaryot Microbiol 1999;46:327–38.

Ueda T, Suga Y, Yahiro N *et al*. Remarkable $N_2$-fixing bacterial diversity detected in rice roots by molecular evolutionary analysis of nifH gene sequences. J Bacteriol 1995;177:1414–7.

Vacheron J, Desbrosses G, Bouffaud M-L *et al*. Plant growth-promoting rhizobacteria and root system functioning. Front Plant Sci 2013;4:1–19.

Vélez P, Espinosa-Asuar L, Travisano M *et al*. The Niche at the Edge of Life or the Microbial Ecology (Including Microfungi) of Cuatro Ciénegas: Mutualisms with Locals, Antagonisms Against Foreigners. In: Souza V, Eguiarte LE, Olmedo-Álvarez G (eds.). Cuatro Ciénegas Ecology, Natural History and Microbiology. Switzerland: Springer, 2018, 73–82.

Ventura M, Canchaya C, Tauch A *et al*. Genomics of Actinobacteria: Tracing the Evolutionary History of an Ancient Phylum. Microbiol Mol Biol Rev 2007;71:495–548.

Vives-Peris V, de Ollas C, Gómez-Cadenas A *et al*. Root exudates: from plant to rhizosphere and beyond. Plant Cell Rep 2020;39:3–17.

Vurukonda SSKP, Vardharajula S, Shrivastava M *et al*. Enhancement of drought stress tolerance in crops by plant growth promoting rhizobacteria. Microbiol Res 2016;184:13–24.

Wanke A, Malisic M, Wawra S *et al*. Unraveling the sugar code: the role of microbial extracellular glycans in plant–microbe interactions. Weber A (ed.). J Exp Bot 2021;72:15–35.

Weldua Y, Haileb M, Habtegebrielb K. Effect of zinc and phosphorus fertilizers application on yield and yield components of faba bean (*Vicia faba* L.) grown in calcaric cambisol of semi-arid northern. J Soil Sci Environ Manag 2012;3:320–6.

Wemheuer F, Taylor JA, Daniel R *et al*. Tax4Fun2: prediction of habitat-specific functional profiles and functional redundancy based on 16S rRNA gene sequences. Environ Microbiome 2020;15:11.

Witte CP. Urea metabolism in plants. Plant Sci 2011;180:431–8.

Yoneyama T, Terakado-Tonooka J, Minamisawa K. Exploration of bacterial $N_2$-fixation systems in association with soil-grown sugarcane, sweet potato, and paddy rice: a review and synthesis. Soil Sci Plant Nutr 2017;63:578–90.

Yuan J, Zhao J, Wen T *et al*. Root exudates drive the soil-borne legacy of aboveground pathogen infection. Microbiome 2018;6:12.

Zappa S, Rolland JL, Flament D *et al*. Characterization of a Highly Thermostable Alkaline Phosphatase from the *Euryarchaeon Pyrococcus abyssi*. Appl Environ Microbiol 2001;67:4504–11.

Zavala-Hurtado JA, Jiménez M. Diversity and Uniqueness at Its Best: Vegetation of the Chihuahuan Desert. Plant Di. Springer, Cham, 2020, 1–17.

Zehr JP, Jenkins BD, Short SM *et al*. Nitrogenase gene diversity and microbial community structure: a cross-system comparison. Env Microbiol 2003;5:539–54.

Zehr JP, Turner PJ. Nitrogen fixation: Nitrogenase genes and gene expression. Methods Microbiol 2001;30:271–86.

Zhalnina K, Louie KB, Hao Z *et al*. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. Nat Microbiol 2018;3:470–80.

Zheng L, Dean DR. Catalytic formation of a nitrogenase iron-sulfur cluster. J Biol Chem 1994;269:18723–6.

# Chapter 3. Arginine and methionine increase the enzymatic activity of nutrient-improvement bacteria and the mineralization of organic nitrogen sources in arid soil from the Chihuahuan Desert.

## 3.1 Introduction

The rhizosphere is a nutrient-rich space around the roots that receive up to 20% of photosynthetic carbon from the plant (Richardson and Simpson 2011). The mix in the root exudates is mainly composed of sugars, organic acids, amino acids, and other low molecular weight organic substances (Stintzing and Carle 2005; Tao *et al.* 2020), which increase the microbial biomass and preferentially recruits plant-growth-promoting rhizobacteria (PGPR) (Farrar *et al.* 2003; Huang *et al.* 2014). As such, the presence of these beneficial bacteria in the rhizosphere aids the plant through different mechanisms like nutrient availability enhancement, biological control against pathogens, phytohormones production, and favor the tolerance of plants to abiotic stress (Selvakumar *et al.* 2014; Olanrewaju *et al.* 2017; Jeyanthi and Kanimoshi, 2018). These traits are even more valuable in arid environments where the plants are subjected to low water levels, high UV radiation, salinity, alkaline soils, and nutrient deficiency (Gamalero *et al.* 2020).

Nitrogen-fixing bacteria (NFB), for example, include a extensive diversity of taxa distributed in all ecosystems (Dos Santos *et al.* 2012; Pereira and Silva, 2013), where they are closely related to the nitrogen cycle (Jones and Darrah *et al.* 1994; Phillips *et al.* 2004). It is estimated that up to 70% of the nitrogen needed in crops is contributed by diazotrophic organisms (Yoneyama *et al.* 2017), either as endophytes or free-living organisms (Ma *et al.* 2003; Bahulikar *et al.* 2014). In turn, organic phosphorus mineralizing bacteria (OPMB) liberate phosphorus (P) from organic compounds by the action of phosphatases (Eder *et al.* 1996). Several phosphatases intervene in P mineralization, including phosphomonoesterases, phosphodiesterases, phytases, and phosphonatases (Rodríguez *et al.* 2006; Stout *et al.* 2014). Since organic P is probably the primary P source in the soil, OPMB are essential for the nutrient availability (García-Oliva *et al*. 2018).

The importance of NFB and OPMB among other PGPR for the improvement of the nourishment of plants has been reported (Coleman-Derr *et al.* 2016; Santos *et al.* 2019; Flores-Núñez *et al.* 2020), and their recruitment in the rhizosphere has been associated with the release of root exudates, that through their high concentration of carbon, enhance its availability (Dechassa and Schenk 2004), and in the case of organic acids, contribute to the mobilization of the organic matter in the soil (Shi *et al.* 2011). For example, organic acids such as tartaric acid produced by the crassulacean acid metabolism (CAM) plant *Sedum alfredii* can solubilize nutrients from the soil and aid in their mobilization (Tao *et al.* 2020). Oxalic acid is also a known solubilizing agent of mineral phosphates in soils (Jiang *et al.* 2020), and it had been related with a 'priming effect' in the microbial activity in amended soils, that may affect a specialized group of microorganisms in contrast to the less specific use of sugars (Landi *et al.* 2006).

Sugars, exert chemoattraction towards microorganisms and thus have an essential role in microbial selection (Feng *et al.* 2018). Even though the composition of root exudates depends on the plant species and age (Gransee and Wittenmayer 2000), the most common sugars identified in the root exudates include glucose, fructose, sucrose, ribose, xylose, and arabinose (Giri *et al.* 2005; Badri *et al.* 2013; Meier *et al.* 2017; Zhang *et al.* 2020). In general, these compounds had been related to the presence of PGPR (Vives-Peris *et al.* 2020). For example, sucrose metabolism has been associated to the symbiotic N fixation in the nodules (Hennion *et al.* 2019), *Rhizobia* in the nodules of legumes receive sugars in exchange of fixed nitrogen (Bezrutczyk *et al.* 2018), and the aggregation of *Azospirillum brasilense* to the root surface increases with the addition of plant-derived sugars (Marini *et al.* 1995). While other PGPR from the *Bacillus* genera, have shown a better response to amino acids than sugars or organic acids (Feng *et al.* 2021).

Amino acids are the most abundant in the root exudates (Paungfoo-Lonhienne *et al.* 2008; Ma *et al.* 2017), and offer structural components for protein synthesis or essential elements after being metabolized (Jones and Darrah 1994). Previous studies in arid lands from the Chihuahuan Desert, showed that the rhizobacteria

associated to *Agave lechuguilla* Torrey, favored the recruitment of PGPR thorough the plant-growth (Medina-de la Rosa *et al.* 2021); and that Methionine (Met) and Arginine (Arg) may increase the enzymatic activity and abundance of NFB and OPMB (Chapter 2). However, the mechanisms for this increment are not completely clear and despite their importance, there are few studies about the influence of specific components of the plant root exudates over the enzymatic activity and abundance of the bacterial community, and even less regarding the root exudates of CAM plants from arid lands. Thus, to understand the influence of the root exudates over nutrient-improvement bacteria, we proposed to identify and test different metabolites from the model plant, *A. lechuguilla* T., in the heterogeneous bacterial community of bulk soil. We hypothesized that the metabolites identified in lechuguilla, would increase the abundance and enzymatic activity of NFB and OPMB, and thus the availability of nutrients in bulk soil samples. The specific objectives of this work were: 1) Evaluate the variation in the abundance and enzymatic activity of NFB and OPMB as a result of the treatment with lactose, maltose, oxalic acid, tartaric acid, methionine, and arginine identified in the rhizosphere of lechuguilla in bulk soil samples to avoid the influence of other root exudates. The activity of the enzymes to quantify were nitrogenase and two types of alkaline phosphatases (mono- and diesterase types). 2) Determine the change in the dynamic of nitrogen and phosphorus concentration in the soil samples after the metabolite treatment. This work is another step to solve the complex mechanisms that maintain plant-bacteria interactions. Furthermore, this information may help to improve the establishment of PGPR as an aid to plants in restoration programs and crops since these bacteria are already tolerant to harsh conditions of arid lands.

## 3.2 Methods

### 3.2.1 Sampling site

The soil samples for the experiment were collected from the previously studied site in the Chihuahuan Desert at Los Amoles in San Luís Potosí, Mexico. The site is located at 1,443 m. The climate is hot semi-arid (Bsh), with annual precipitation of 667.4 mm and temperatures from 17.6° C to 24.6°C. The vegetation is dominated by desert shrubs like cacti, agaves, and *Yucca* woodlands, with large patches of lechuguilla known as “lechuguillales” (Zavala-Hurtado and Jiménez 2020).

We selected four pairs of mature and young agaves, at least 2 m apart, to guarantee the sampling from different specimens (Trame *et al.* 1995), and collected about 1 kg of the closest soil to the roots in four points, within the circumference of the plant and 10 cm below the surface. Similarly, we collected four samples of root-free soil and 2 m afar from any plant, following the same depth parameters as in the rhizosphere samples. The subsamples were homogenized, oven-dried till constant weight at 60°C, sieved, and kept at 4°C till the identification of the metabolites.

### 3.2.2 Metabolite selection

The amino acids were identified in soil samples with the AcQTag TM Ultra Derivatization Kit (Waters) and read with high-performance liquid chromatography with a fluorescence detector. The soil was degreased with hexane and hydrolyzed in a dry bath at 115°C with hydrochloric acid. After that, the samples were reconstituted, derivatized, and read. Methionine (Met) and Arginine (Arg) were selected according to the results from previous studies, as these may be related with the structure of the rhizobacterial community, especially of the nutrient-improvement bacteria from the rhizosphere of *A. lechuguilla* T. (See Chapter 2). The concentration of amino acids in the rhizosphere of mature plants was 0.36 mg $g^{-1}$ dry soil for Arg and 0.12 mg $g^{-1}$ for Met (Table 2.3-2).

The sugars and organic acids were identified and semi-quantified with thin layer chromatography (TLC). First, the metabolites were extracted from 40 g of sieved

rhizospheric soil (33 mm) with a solution of methanol (MeOH), distilled water, and formic acid (50:49.95:0.05) (Pétriacq *et al.* 2017). The mix was shaken for one hour at 120 revolutions per minute (rpm) and centrifuged at 4500 rpm for 10 min. The supernatant was collected and filtered (Whatman®, Grade 42). Then, the MeOH was eliminated from the samples at 50°C (40 rpm) in a rotavapor Büchi. Finally, the samples were frozen dried and kept at -80°C for storage.

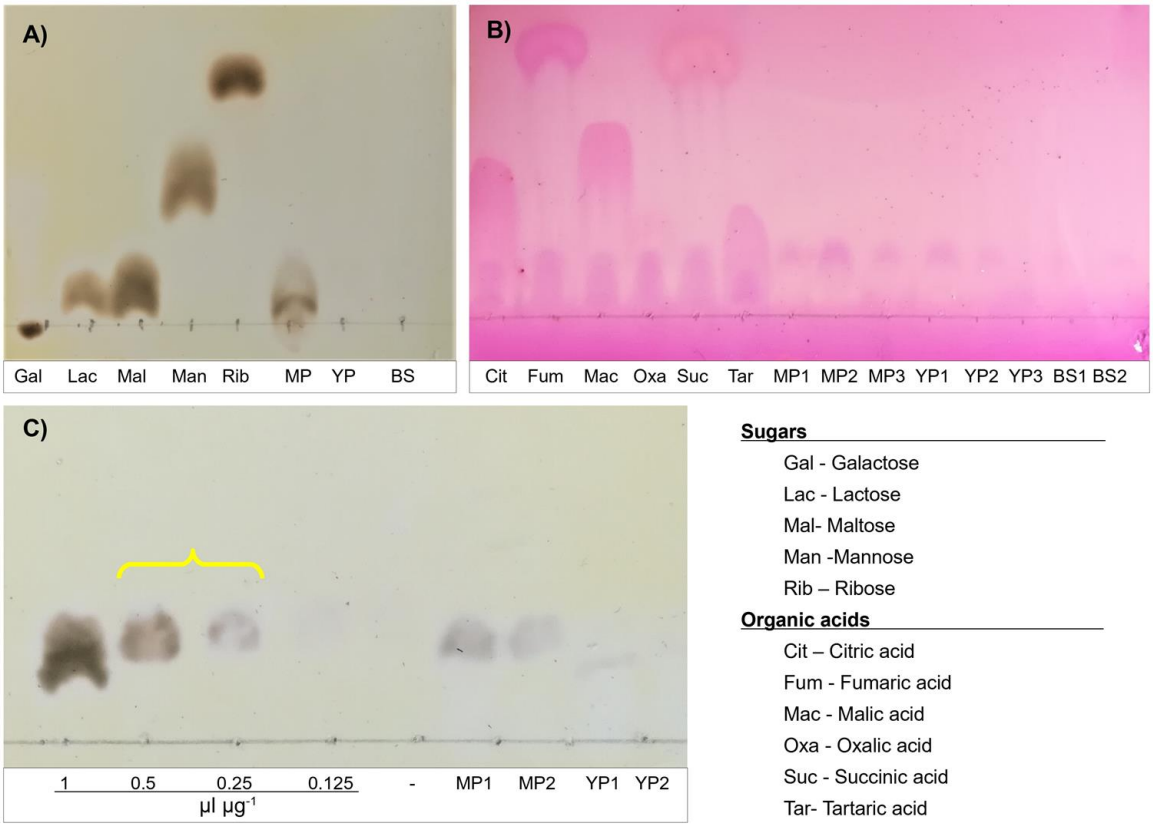



Figure 3.2.1.Identification and semi quantification of sugars and organic acids in the rhizosphere of *A. lechuguilla* T. A) Sugars identified in the rhizosphere of mature lechuguilla (MP) were designed as lactose and maltose. B) In the rhizosphere of mature and young lechuguilla (YP), the organic acids were similar to oxalic and tartaric acids. C) The semi quantification was performed with different concentrations of maltose as reference (1 to 0.125 µl $\mu g^{-1}$). The bands in the rhizosphere samples were similar to the spots in 0.5 and 0.25 µl $\mu g^{-1}$.

The assay with the metabolites was carried on aluminum TLC plates precoated with silica gel and fluorescent indicator F254 (Macherey-Nagel). About 2.5 mg of each extract was dissolved in 500 µl MeOH and run alongside standards. The mobile phase for sugars was a modification to the method of Han and Robyt (1998), with acetonitrile and water (90:10) (Han and Robyt 1998); the plates were sprayed with

a solution of 0.1 N ceric sulfate and 2 N sulfuric acid (1:2), and heated for final development (Fell 1990). Meanwhile, the eluent for organic acids was a mix of water, butanol, and formic acid (50:40:10); the spots were developed with methyl red and bromophenol blue (1:1) in 70% ethanol (Lee *et al.* 2001).

The sugars lactose and maltose and the organic acids oxalic and tartaric were identified in the root exudates of the rhizosphere of lechuguilla at an estimated concentration between 0.5 and 1.5 μg $g^{-1}$ dry soil (Figure 3.2.1).

### 3.2.3 Soil conditioning

A second sampling was made to collect soil to test the metabolites immediately before the experiment setup. Bulk soil samples were taken from interspaces without plant influence, 2 m of distance from any plant, and within the first 10 cm below the surface. About 2 kg of soil were collected and transported to the laboratory, where the soil was sieved (2 mm) and kept in a growth chamber at 27°C to 28°C, with 60% humidity. After that period, the soil was weighed and separated into four bags with about 2.5 kg each. The samples were conserved in the growth chamber till the experiment setup (60 h).

### 3.2.4 Treatments and soil incubation

The treatments consisted of six metabolites: lactose, maltose, oxalic acid, tartaric acid, arginine, and methionine. Each metabolite was tested in three different concentrations and the four repetitions were set up in blocks, each 24 h to facilitate the processing of the samples.

The conditioned soil was distributed in separate bags, each with 90 g of soil moisturized with 20 ml of a sterilized solution of the corresponding metabolite (Figure 3.2.2). The mix was homogenized and incubated for 7 days in a growth chamber (27°C to 28°C, 60% humidity, and 12 h of light). Each repetition followed the same procedure. At the end of the incubation period, all the samples were stored at 4°C till their processing.

| Control | Organic acids | | | | | | Sugars | | | | | | Amino acids | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Water | Oxalic | | | Tartaric | | | Maltose | | | Lactose | | | Arginine | | | Methionine | | |
| T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 | T10 | T11 | T12 | T13 | T14 | T15 | T16 | T17 | T18 | T19 |
| | 5 µg | 30 µg | 60 µg | 5 µg | 30 µg | 60 µg | 5 µg | 30 µg | 60 µg | 5 µg | 30 µg | 60 µg | 100 µg | 600 µg | 1200 µg | 50 µg | 300 µg | 600 µg |
| R1 | | | | | | | | | | | | | | | | | | |
| R2 | | | | | | | | | | | | | | | | | | |
| R3 | | | | | | | | | | | | | | | | | | |
| R4 | | | | | | | | | | | | | | | | | | |

Figure 3.2.2. Experimental design for individual test of metabolites identified in the root exudates from lechuguilla. Each metabolite was tested at three different concentrations, with four repetitions set up as blocks. The control treatment consisted of sterilized distilled water.

### 3.2.5 Enzymatic activity quantification

As the nitrogenase activity was quantified with the acetylene reduction method, the treated soils were deposited in serologic flasks and sealed with rubber stoppers and aluminum caps. About 20% of the headspace air was replaced with acetylene and incubated alongside the rest of the treatments. The nitrogenase activity was quantified through the ethylene production in a gas chromatograph Agilent 6890 GC N.05.04 with a flame ionization detector. The calculations were made according to a standard curve of ethylene (0.3-2.1 nmol) (López-Lozano, Carcaño-Montiel and Bashan 2016). Results were listed as nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$.

Both alkaline phosphomonoesterase (APM) and phosphodiesterase (APD) were determined with the colorimetric method of Tabatabai and Bremner (1969). Briefly, 0.5 g of soil was mixed with 2.375 buffer Tris 0.5 M (pH 11) and 0.125 ml of the substrate; p-Nitrophenol phosphate (Sigma-Aldrich P7998) for APM and Bis(4-nitrophenyl) phosphate (Sigma-Aldrich 123943) for APD. The mixture was incubated at 37°C while stirring at 150 rpm for 30 min. The reaction was terminated with 2 ml of 0.5 M NaOH and 2 ml of 0.5 M $CaCl_2$. The blend was centrifuged at 2,500 rpm for 5 min, and the absorbance of the supernatant was read at 405 nm (Tabatabai and Bremner 1969). Blank samples without the substrate and a standard curve of p-nitrophenol (0 to 180 pNP) were run alongside the samples. The results are reported as µg pNP $g^{-1}$ $h^{-1}$.

### 3.2.6 *nifH* and *phoD* genes abundance

First, the DNA extraction was performed with the DNeasy® PowerSoil® Kit (Qiagen®), following the manufacturer instructions. The quantification of the *nifH* and *phoD* genes present in NFB and OPMB, respectively, was conducted with the real-time polymerase chain reaction (qPCR) in the PikoReal 96 Real-Time PCR system (Thermo Fisher Scientific Inc., TCR0096). Previously, it was verified that the DNA concentration employed did not hinder the reaction.

The reaction for the *nifH* gene had a final volume of 10 µl, which included: 5 µl of 2X SYBR Green, 0.3 mM *PolR* (5'-ATSGCCATCATYTCRCCGGA-3') (Poly, Monrozier and Bally 2001), 0.15 mM *FPGH19* (5'-TACGGCAARGGTGGNATHG-3') (Simonet *et al.* 1991), and 1 µl DNA (1:20). The reaction was carried on with an initial denaturation of 10 min (95°C), 40 cycles: denaturation (94°C), alignment (57°C), and extension (72°C), 1 min each; the final step was an extension of 5 min at 72°C (Simonet *et al.* 1991).

For the OPMB, the abundance of the *phoD* gene was determined with a 10 µl reaction constituted of 5 µl 2x SYBR Green, 0.4 mM of the primers *F733* (5'-TGGGAYGATCAYGARGT-3') and *R1083* (5'-CTGSGCSAKSACRTTCCA-3'); and 1 µl DNA (1:10). The reaction consisted of 10 min of the initial denaturation (95°C) followed by 40 cycles of: 15 s of denaturation (95°C), 1 min alignment (57°C) and 1 min extension (72°C); with final extension of 72°C for 5 min (Ragot, Kertsez and Bünemann 2015). The quality of the reactions was assessed with a melting step up to 95°C. Lastly, we included standard curves (100 to $10^8$ copies) with clones of *Bacillus subtilis* (*phoD*) and *Geobacter sulfurreducens* (*nifH*), obtained with the pGEM-T Easy Vector system (PROMEGA®).

### 3.2.7 Nutrient quantification in treated soils

To assess the variations in nitrogen and phosphorus, we quantified extractable phosphorus (P ext.) with the method of Bray-Kurtz (Bray and Kurtz 1945). The quantification of nitrite and nitrate ($NO_2$+$NO_3$) was performed with the vanadium reduction (García-Robledo *et al.* 2014), and a variation to the Berthelot method

(Nelson 1983) for $NH_4$. The dissolved organic phosphorus (DOP) and nitrogen (DON) were determined at the Instituto de Investigaciones en Ecosistemas y Sustentabilidad, UNAM, Mexico. Both were calculated as the difference between the digested and the inorganic soluble concentrations (Joergensen and Muller 1996). Briefly, the total nitrogen and phosphorus dissolved concentration were obtained through the digestion of the samples with sulfuric acid and hydroxide peroxide; in their mineralized forms ammonium and orthophosphate, respectively. The soluble forms were determined from fresh soil samples (2 g) stirred in deionized water (1 h), and filtered (Chávez-Vergara *et al.* 2018). In addition, with these results, we obtained the dissolved inorganic nitrogen (DIN) and DON ratio, dissolved inorganic phosphorus (DIP) and DOP ratio, and DON:DOP ratios. All of these relations, give us information about the dynamics of the nutrients, as these relations change according to the season, locations or agricultural activities, among other factors (Islam *et al.* 2019). Moreover, the ratios of nutrients let us know about their availability for plants or microorganisms, the mineralization process, and even if the nutrient concentration may promote the competition in the soil (Mattsson *et al.* 2009).

### 3.2.8 Statistical analysis

All statistical analyses were performed with the software R (V 4.1.0.) in RStudio (V 1.4.1717) and base packages (R Core Team, 2020). The difference between the effects of the treatments and control was tested with Kruskal-Wallis and Dunn methods ($p<0.05$). Moreover, we tested the influence of the metabolite type and concentration with a two-way analysis of variance (ANOVA). Finally, canonical correspondence analysis (CCA) was performed with vegan (V 2.5-7), to determinate the variance of nutrient concentration in the soil, explained by the enzymatic activity and abundance of the nutrient-improvement bacteria. Before the analysis, the variables were selected according to linear regression models, and the significance of the CCA model and individual axis was tested with a permutation test (Permutations = 999).

## 3.3 Results

### 3.3.1 Amino acids treatments enhanced the enzymatic activity of the nutrient-improvement bacteria

Metabolite type and concentration significantly affected most of the parameters evaluated in the treated soils ($p<0.05$). However, amino acids seemed to be the most influencing metabolites to the enzymatic activity of NFB (Figure 3.3.1A). For example, the soil nitrogenase activity similarly increased with the medium concentration of tartaric acid up to 599.80 ± 8.99 nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$, the low dose of lactose was 595.64 ± 6.14 nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$, and most of the treatments with Arg, which ranged from 605.07±3.75 to 604.34± 9.22 nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$; however, nitrogenase activity with Met was the highest, from 620.91 up to 659.23 nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$ (Kruskal-Wallis and Dun test, $p<0.05$, Figure 3.3.1).

Moreover, ammonium concentration increased under the Arg treatments ($p<0.05$, Table 3.3-1), where it reached about 79.12 ± 2.74 mg $NH_4$ $kg^{-1}$ (Figure 3.3.2A), while the DON concentration, considerably decreased with low lactose to 0.30 ± 0.05 µg $g^{-1}$. In amino acids treatments the DON also remained lower than the control but at a different pace; for example, DON concentration increased from 0.33 ± 0.06 µg $g^{-1}$ at low Arg ($p<0.01$) to 0.53 ± 0.12 µg $g^{-1}$ ($p<0.05$) in the high dose. In contrast, the DON concentration in Met dosed soil decreased from 1.30 ± 0.32 µg $g^{-1}$, to the lowest value of all samples, 0.28 ± 0.04 µg DON $g^{-1}$ ($p<0.001$, Figure 3.3.1). Lastly, as the DIN:DON ratio suggested, most of the nitrogen in the soil samples treated with Arg (Figure 11A), was in the inorganic forms and showed the highest values among the other samples ($p<0.001$, Table 3.3-1). The number of *nifH* gene copies (Figure 3.3.1B) and $NO_2$+$NO_3$ concentration showed no variation compared to the control treatment ($p>0.05$, Figure 3.3.2).

As for OPMB, the activity of alkaline phosphodiesterase increased mainly with sugars and amino acids (Table 3.3-2). Both sugars seemed to enhance gradually the activity of APD ($p<0.05$); and the highest activity in these treatments was reported for the high dose of lactose, 261.41 ± 20.30 µg pNP $g^{-1}$ $h^{-1}$, and maltose with 272.66

± 22.84 µg pNP $g^{-1}$ $h^{-1}$ (Table 3.3-2). However, APD activity remained at similar levels, around 256 µg pNP $g^{-1}$ $h^{-1}$, with the low and high concentrations of Met (Figure 3.3.1D). The low and high treatments with Arg, also showed similar levels of APD activity, 266.80 ± 18.94 µg pNP $g^{-1}$ $h^{-1}$ and 277.44 ± 18.30 µg pNP $g^{-1}$ $h^{-1}$, respectively; while the highest value, 284.03 ± 20.85 µg pNP $g^{-1}$ $h^{-1}$, was recorded for the medium dose of Arg (Table 3.3-2). Furthermore, the abundance of the *phoD* gene increased only with the medium dose of oxalic acid to 7.40 ± 0.01 log copies *phoD* $g^{-1}$ soil ($p<0.05$, Figure 3.3.1E).

Table 3.3-1. Nitrogenase activity and *nifH* gene abundance in soil samples after their inoculation with metabolites identified in *A. lechuguilla* T. Nitrogen forms include: $NH_4$= Ammonium, $NO_2$+$NO_3$= Nitrite + nitrate, DIN= Dissolved inorganic nitrogen, and DON= dissolved organic nitrogen. Values correspond to mean and standard error (n = 4). Two-way ANOVA of interaction between metabolite and concentration[1] (*** p <0.001, ** p <0.01, * p <0.05). Met= Metabolite type. [1]Interaction between metabolite type and concentration.

| Metabolite | µg $g^{-1}$ | Nitrogenase nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$ | Log copies *nifH* $g^{-1}$ soil | $NH_4$ mg $kg^{-1}$ | $NO_2$+$NO_3$ mg $kg^{-1}$ | DON µg $g^{-1}$ | DIN:DON |
|---|---|---|---|---|---|---|---|
| Control | 0 | 563.04 ± 10.59 | 5.51 ± 0.06 | 3.38 ± 0.38 | 51.77 ± 16.38 | 1.67 ± 0.08 | 32.99 ± 9.47 |
| Oxalic | 5 | 577.21 ± 11.13 | 5.46 ± 0.19 | 2.87 ± 0.46 | 31.96 ± 6.17 | 1.65 ± 0.12 | 21.70 ± 4.19 |
| | 30 | 586.34 ± 12.52 | 5.83 ± 0.20 | 2.85 ± 0.35 | 47.32 ± 15.81 | 1.17 ± 0.41 | 83.24 ± 55.31 |
| | 60 | 589.00 ± 8.92 | 5.70 ± 0.06 | 3.02 ± 0.68 | 28.35 ± 1.96 | 1.53 ± 0.17 | 21.70 ± 3.90 |
| Tartaric | 5 | 582.44 ± 7.85 | 5.64 ± 0.15 | 3.65 ± 0.83 | 27.12 ± 6.13 | 1.66 ± 0.12 | 18.11 ± 2.87 |
| | 30 | 599.80 ± 8.99 | 5.61 ± 0.22 | 1.87 ± 0.20 | 35.32 ± 5.93 | 1.40 ± 0.26 | 27.31 ± 1.45 |
| | 60 | 576.07 ± 1.34 | 5.71 ± 0.24 | 2.72 ± 0.43 | 26.10 ± 7.37 | 1.65 ± 0.12 | 16.85 ± 3.59 |
| Lactose | 5 | 595.64 ± 6.14 | 5.51 ± 0.02 | 2.81 ± 0.19 | 26.18 ± 5.06 | 0.30 ± 0.05 | 95.40 ± 9.56 |
| | 30 | 585.06 ± 7.08 | 5.70 ± 0.20 | 2.98 ± 0.27 | 31.66 ± 8.77 | 1.40 ± 0.11 | 25.10 ± 6.28 |
| | 60 | 588.87 ± 10.12 | 5.55 ± 0.07 | 2.78 ± 0.74 | 21.79 ± 3.77 | 1.64 ± 0.11 | 15.14 ± 2.48 |
| Maltose | 5 | 576.08 ± 9.17 | 5.58 ± 0.14 | 3.87 ± 0.43 | 32.91 ± 7.07 | 1.53 ± 0.11 | 24.15 ± 4.71 |
| | 30 | 577.31 ± 2.10 | 5.54 ± 0.15 | 2.90 ± 0.28 | 41.00 ± 19.75 | 1.57 ± 0.11 | 28.66 ± 13.67 |
| | 60 | 596.58 ± 9.67 | 5.33 ± 0.14 | 3.42 ± 0.40 | 35.38 ± 6.19 | 1.59 ± 0.10 | 24.73 ± 4.40 |
| Arginine | 100 | 576.04 ± 10.51 | 5.34 ± 0.21 | 23.82 ± 1.39 | 80.53 ± 15.81 | 0.33 ± 0.06 | 327.03 ± 38.28 |
| | 600 | 605.07 ± 3.75 | 5.63 ± 0.15 | 76.33 ± 2.51 | 496.67 ± 120.31 | 0.40 ± 0.11 | 1700.30 ± 460.78 |
| | 1200 | 604.34 ± 9.22 | 5.43 ± 0.07 | 79.12 ± 2.74 | 640.54 ± 99.51 | 0.53 ± 0.12 | 1758.44 ± 603.14 |
| Methionine | 50 | 620.91 ± 9.69 | 5.37 ± 0.09 | 6.57 ± 0.35 | 28.21 ± 3.62 | 1.30 ± 0.32 | 35.26 ± 11.50 |
| | 300 | 633.39 ± 15.14 | 5.52 ± 0.14 | 21.96 ± 1.18 | 19.85 ± 1.50 | 0.65 ± 0.27 | 83.41 ± 28.17 |
| | 600 | 659.23 ± 5.46 | 5.54 ± 0.10 | 17.79 ± 1.43 | 43.73 ± 12.15 | 0.28 ± 0.04 | 253.05 ± 72.19 |
| Two-way ANOVA | Met. | 0.000 | 1.069 | 0.000 | 0.000 | 22.016 | 85.000 |
| | Con. | 0.000 | 0.673 | 0.000 | 0.000 | 0.547 | 3.057 |
| | M*C[1] | 1.679 | 0.689 | 0.000 | 0.000 | 0.000 | 0.000 |

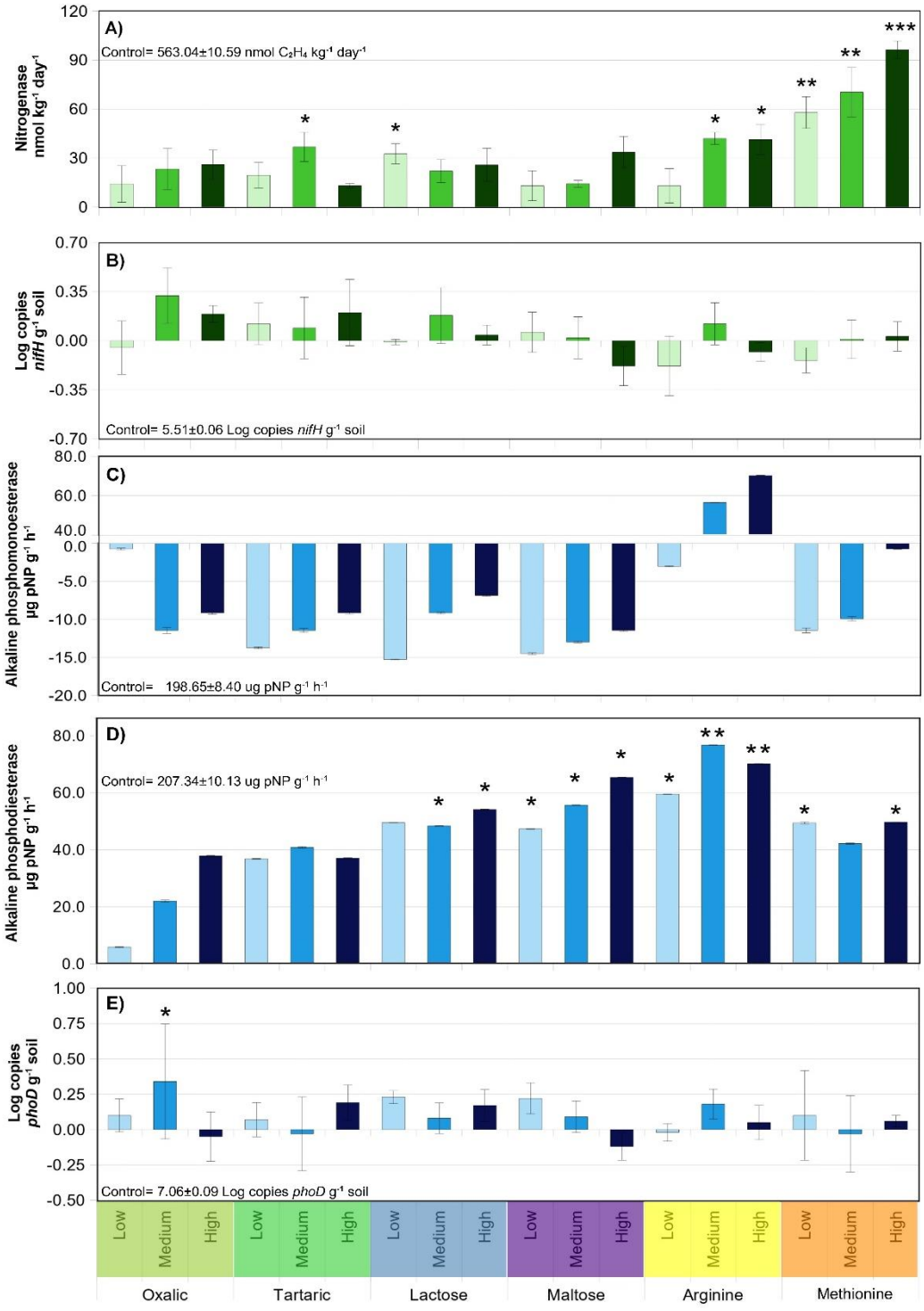



Figure 3.3.1. Enzymatic activity and abundance of NFB and OPMB in dosed soil with different metabolites. NFB characterization includes nitrogenase activity (A) and *nifH* gene abundance (B); as for OPMB, two types of phosphatases were quantified: Alkaline phosphomonoesterase (C) and phosphodiesterase (D), plus the abundance of *phoD* gene (E). Values depict difference against control treatment and standard error. Kruskal-Wallis and Dunn tests were applied (* = $p<0.05$, ** = $p<0.01$, *** = $p<0.05$). Control means are displayed.

Lastly, the metabolites and its concentration showed no influence over the organic or inorganic P forms, nor their ratio (Table 3.3-2); except for the lower concentration of DOP, 1.49 ± 0.36 μg $g^{-1}$, with the medium dose of tartaric acid ($P<0.05$, Figure 3.3.3E). The ratio between DON and DOP was influenced by both metabolite type and concentration ($p<0.001$) and decreased mainly with the addition of amino acids to similar levels as with the low concentration of lactose (0.17 ± 0.02; $p<0.01$) (Figure 3.3.3C).

Table 3.3-2. Phosphatases activity and *phoD* gene abundance in soil samples after their inoculation with metabolites identified in *A. lechuguilla* T. Phosphorus inorganic and organic forms are also displayed: DIP= Dissolved inorganic phosphorus, DOP= Dissolved organic phosphorus, and DON:DOP= Ratio between dissolved organic forms of nitrogen and phosphorus. Values correspond to mean and standard error (n = 4). At the bottom of the table, a two-way ANOVA is shown (*** p <0.001, ** p <0.01, * p <0.05). Met= Metabolite type. [1]Interaction between metabolite type and concentration.

| Metabolite | μg $g^{-1}$ | APM ug pNP $g^{-1}$ $h^{-1}$ | APD ug pNP $g^{-1}$ $h^{-1}$ | log copies *phoD* $g^{-1}$ soil | DIP mg $kg^{-1}$ | DOP μg $g^{-1}$ | DIP:DOP | DON:DOP |
|---|---|---|---|---|---|---|---|---|
| Control | 0 | 198.65 ± 8.40 | 207.34 ± 10.13 | 7.06 ± 0.09 | 4.20 ± 0.68 | 1.85±0.01 | 2.26±0.36 | 0.91±0.04 |
| Oxalic | 5 | 197.89 ± 7.32 | 213.16 ± 13.32 | 7.16 ± 0.20 | 4.14 ± 0.73 | 1.84±0.01 | 2.26±0.39 | 0.90±0.06 |
| | 30 | 187.21 ± 8.36 | 229.27 ± 15.64 | 7.40 ± 0.01 | 4.00 ± 0.75 | 1.49±0.36 | 3.88±1.70 | 0.80±0.17 |
| | 60 | 189.49 ± 10.38 | 245.12 ± 9.27 | 7.01 ± 0.11 | 3.98 ± 0.48 | 1.84±0.01 | 2.16±0.26 | 0.84±0.10 |
| Tartaric | 5 | 184.92 ± 7.62 | 244.11 ± 28.77 | 7.13 ± 0.09 | 4.11 ± 0.49 | 1.84±0.01 | 2.24±0.27 | 0.90±0.06 |
| | 30 | 187.20 ± 9.74 | 248.15 ± 20.72 | 7.03 ± 0.13 | 3.60 ± 0.19 | 1.83±0.01 | 1.97±0.11 | 0.76±0.14 |
| | 60 | 189.50 ± 10.23 | 244.36 ± 23.18 | 7.25 ± 0.13 | 3.31 ± 0.43 | 1.84±0.01 | 1.80±0.23 | 0.90±0.06 |
| Lactose | 5 | 183.39 ± 10.46 | 256.80 ± 20.80 | 7.29 ± 0.08 | 3.59 ± 0.20 | 1.85±0.01 | 1.95±0.12 | 0.17±0.02 |
| | 30 | 189.50 ± 10.46 | 255.66 ± 11.81 | 7.15 ± 0.15 | 3.77 ± 0.40 | 1.87±0.01 | 2.02±0.22 | 0.75±0.05 |
| | 60 | 191.79 ± 9.03 | 261.41 ± 20.30 | 7.23 ± 0.02 | 3.44 ± 0.18 | 1.85±0.01 | 1.86±0.09 | 0.88±0.06 |
| Maltose | 5 | 184.15 ± 11.08 | 254.62 ± 13.23 | 7.28 ± 0.18 | 3.09 ± 0.28 | 1.84±0.00 | 1.68±0.15 | 0.83±0.06 |
| | 30 | 185.68 ± 8.68 | 262.94 ± 19.79 | 7.15 ± 0.02 | 3.84 ± 0.15 | 1.86±0.00 | 2.07±0.08 | 0.85±0.06 |
| | 60 | 187.21 ± 12.28 | 272.66 ± 22.84 | 6.94 ± 0.16 | 3.55 ± 0.44 | 1.84±0.01 | 1.94±0.23 | 0.87±0.05 |
| Arginine | 100 | 195.61 ± 9.93 | 266.80 ± 18.94 | 7.04 ± 0.24 | 3.74 ± 0.44 | 1.84±0.01 | 2.02±0.23 | 0.18±0.04 |
| | 600 | 255.14 ± 16.22 | 284.03 ± 20.85 | 7.24 ± 0.08 | 3.86 ± 0.54 | 1.84±0.01 | 2.09±0.28 | 0.22±0.06 |
| | 1200 | 268.88 ± 17.56 | 277.44 ± 18.30 | 7.11 ± 0.24 | 3.61 ± 0.60 | 1.84±0.00 | 1.97±0.32 | 0.29±0.07 |
| Methionine | 50 | 187.21 ± 4.83 | 256.73 ± 11.06 | 7.16 ± 0.09 | 4.01 ± 0.69 | 1.62±0.23 | 2.66±0.56 | 0.75±0.13 |
| | 300 | 188.73 ± 4.75 | 249.48 ± 4.55 | 7.03 ± 0.18 | 3.70 ± 0.40 | 1.83±0.01 | 1.90±0.25 | 0.36±0.15 |
| | 600 | 197.89 ± 4.75 | 256.98 ± 4.57 | 7.12 ± 0.09 | 4.89 ± 0.82 | 1.84±0.01 | 2.66±0.45 | 0.15±0.02 |
| Two-way ANOVA | Met. | **0.000** | **0.000** | 0.303 | 0.895 | 0.853 | 1.324 | **0.000** |
| | Con. | **0.000** | 0.576 | 0.401 | 0.002 | 0.495 | 0.330 | 0.379 |
| | M*C[1] | **0.000** | 0.168 | 0.950 | 0.599 | 0.976 | 0.646 | **0.000** |

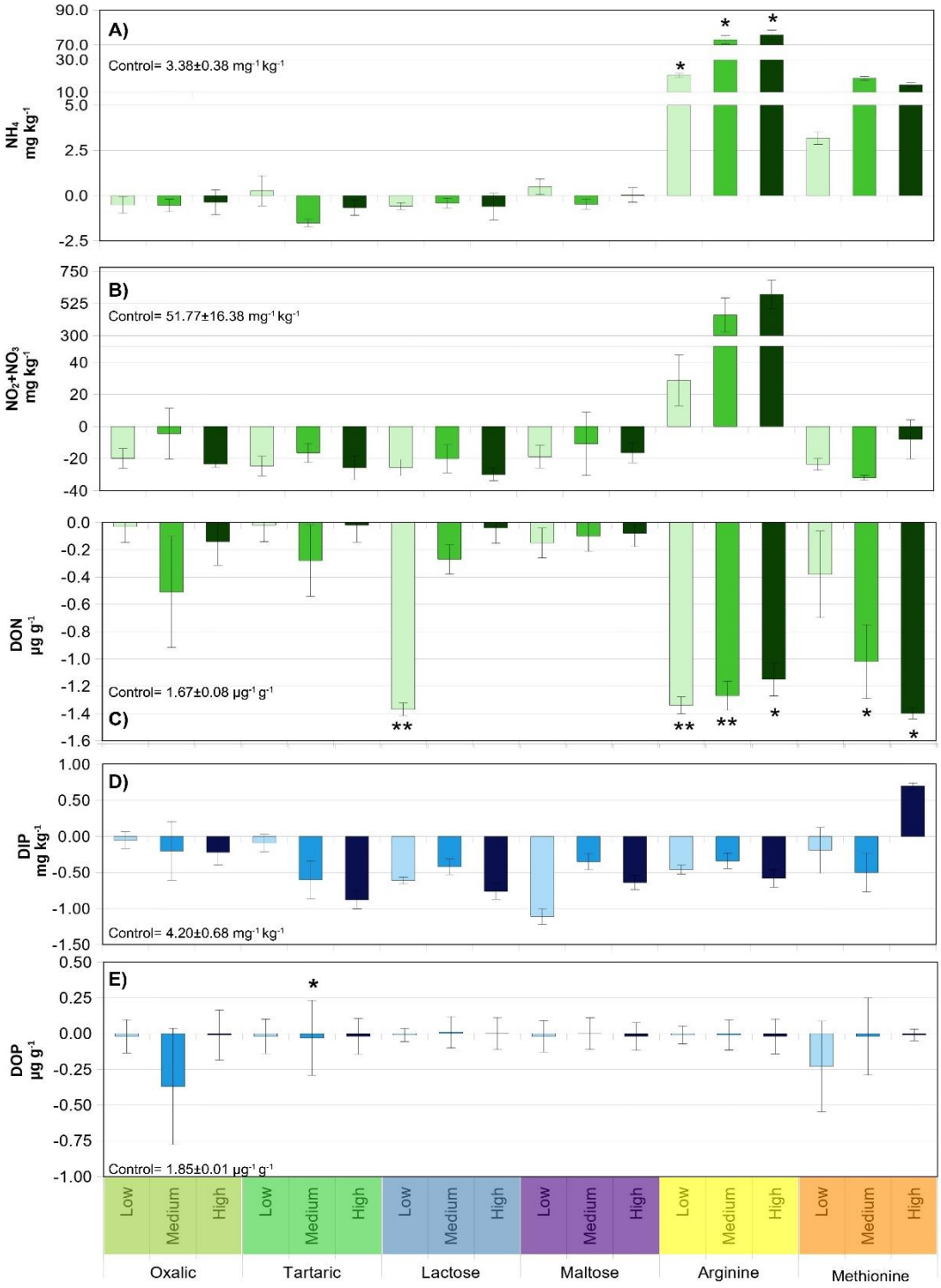



Figure 3.3.2. Nitrogen and phosphorus concentration in soil samples dosed with metabolites. The quantified forms of nitrogen in the soil samples included: Ammonium (A), Nitrite + Nitrate (B), and dissolved organic nitrogen (C); as for phosphorus, the Dissolved inorganic phosphorus (D) and dissolved organic phosphorus (E). The depicted values correspond to the difference against the control and the standard error. Kruskal-Wallis and Dunn test were applied (* = p<0.05, ** = p<0.01, *** = p<0.05). Control means are displayed.

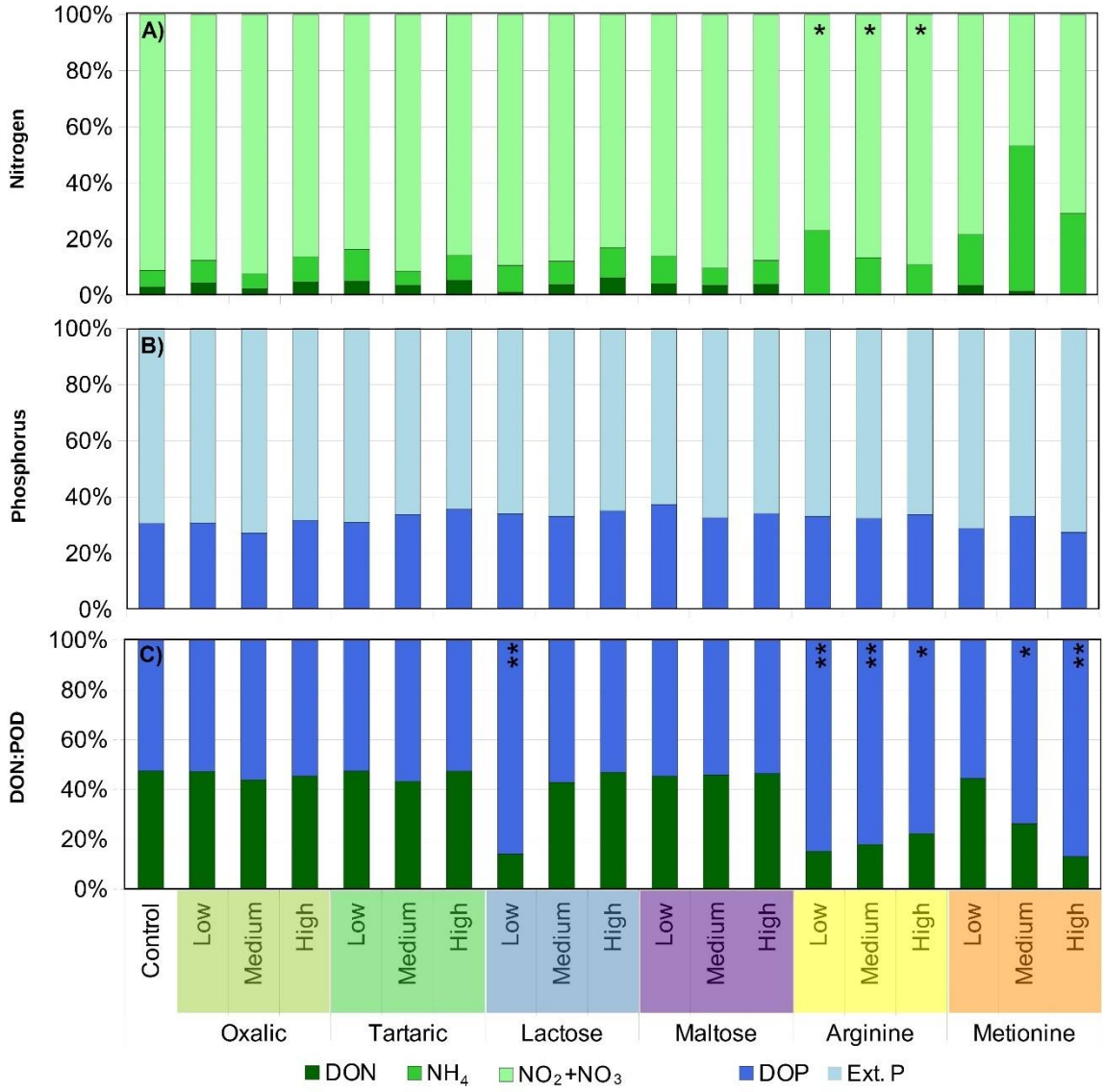



Figure 3.3.3.Nutrient coefficients in amended soil samples with different metabolite. A) Relation between ammonium ($NH_4$), nitrite + nitrate ($NO_2$+$NO_3$), and dissolved organic nitrogen (DON); B) Proportion between dissolved inorganic and organic phosphorus; C) Relation between DON and DOP. Differences in the coefficients of treatments against the control were tested with Kruskal-Wallis and Dunn method (* = $p<0.05$, ** = $p<0.01$, *** = $p<0.05$).

### 3.3.2 Enzymatic activities were related to organic and inorganic nitrogen forms

To understand the influence of the enzymatic activity and abundance of NFB and OPMB over the nutrient concentration in our experimental units, we performed a CCA. Previous to the analysis, we explored the linear relations between all the

variables, and selected $NH_4$, $NO_2$+$NO_3$, DON:DOP, and DIP, as these had most of the significant relations with the bacterial response and avoided the collinearity between the nutrients ($p<0.05$; Table 3.3-3).

Table 3.3-3. Linear relation between nutrient concentration and enzymatic activity and abundance of nutrient-improvement bacteria was tested with linear regression models. In bold are significant F and p values (* = $p<0.05$, ** = $p<0.01$, *** = $p<0.05$).

| Nutrient | Nitrogenase | Log *nifH* | APM | APD | Log *phoD* |
|---|---|---|---|---|---|
| **$NH_4$** | **4.57*** | 0.49 | **111.80***** | **7.01**** | 0.00 |
| **$NO_2$+$NO_3$** | 0.82 | 0.30 | **121.00***** | **5.72*** | 0.41 |
| **DIN:DON** | 2.24 | 0.10 | **60.73***** | **7.64**** | 0.09 |
| **DON** | **27.09***** | 0.19 | **5.25*** | **13.63***** | 0.71 |
| **DIP** | 0.03 | **4.38*** | 0.28 | **11.38**** | **3.98*** |
| **DOP** | **4.60*** | 1.32 | 0.07 | 0.43 | 0.72 |
| **DIP:DOP** | 2.20 | 0.17 | 0.09 | 1.47 | 0.03 |
| **DON:DOP** | **21.94***** | 0 | **5.47*** | **12.88***** | 0.39 |

In the model, the inertia explained by the constrained variables was 41.45% ($p<0.001$, permutations= 999), with each component explaining 32.32% for CCA1 ($p<0.001$%, permutations= 999) and 9.06% for CCA2 ($p<0.01$, permutations= 999). According to the CCA model, the nutrient concentration in the soils treated with amino acids seem to favor a higher activity of nitrogenase, which was mostly reflected in the increment of the inorganic forms of nitrogen in the soil (Figure 3.3.4). Moreover, it is important to point out that while DIN increased along with the nitrogen fixation (Figure 3.3.3A), DON diminished (F value = 27.09, $p<0.001$), and with this, also the ratio between DON:DOP (Table 3.3-1).

Additionally, the DIP concentration showed relation with the APD activity (F value = 11.38, $p<0.01$; Table 3.3-2) and abundance of the *phoD* gene (F value = 3.981, $p<0.05$). The increased APD activity seems to be negative related with the concentration of DIP, since the CCA plot shows that its mean is near the sugars and organic acids and diminish with the amino acid treatments (Table 3.3-2). Finally, it is

remarkable that while the enzymatic activity of OPMB had few linear relations with the phosphorus concentration, both phosphatases were related to all the forms of nitrogen concentration (Table 3.3-2). Furthermore, nitrogenase activity was related to DOP concentration (F value = 4.604; $p<0.05$) and the abundance of *nifH* explained some of the variation of the DIP concentration (F value = 4.377; $p<0.05$). Relations that are also reflected on the CCA plot, were the phosphatases vectors are directed to the higher concentration of the inorganic forms of nitrogen in the amino acid treatments (Figure 3.3.4).

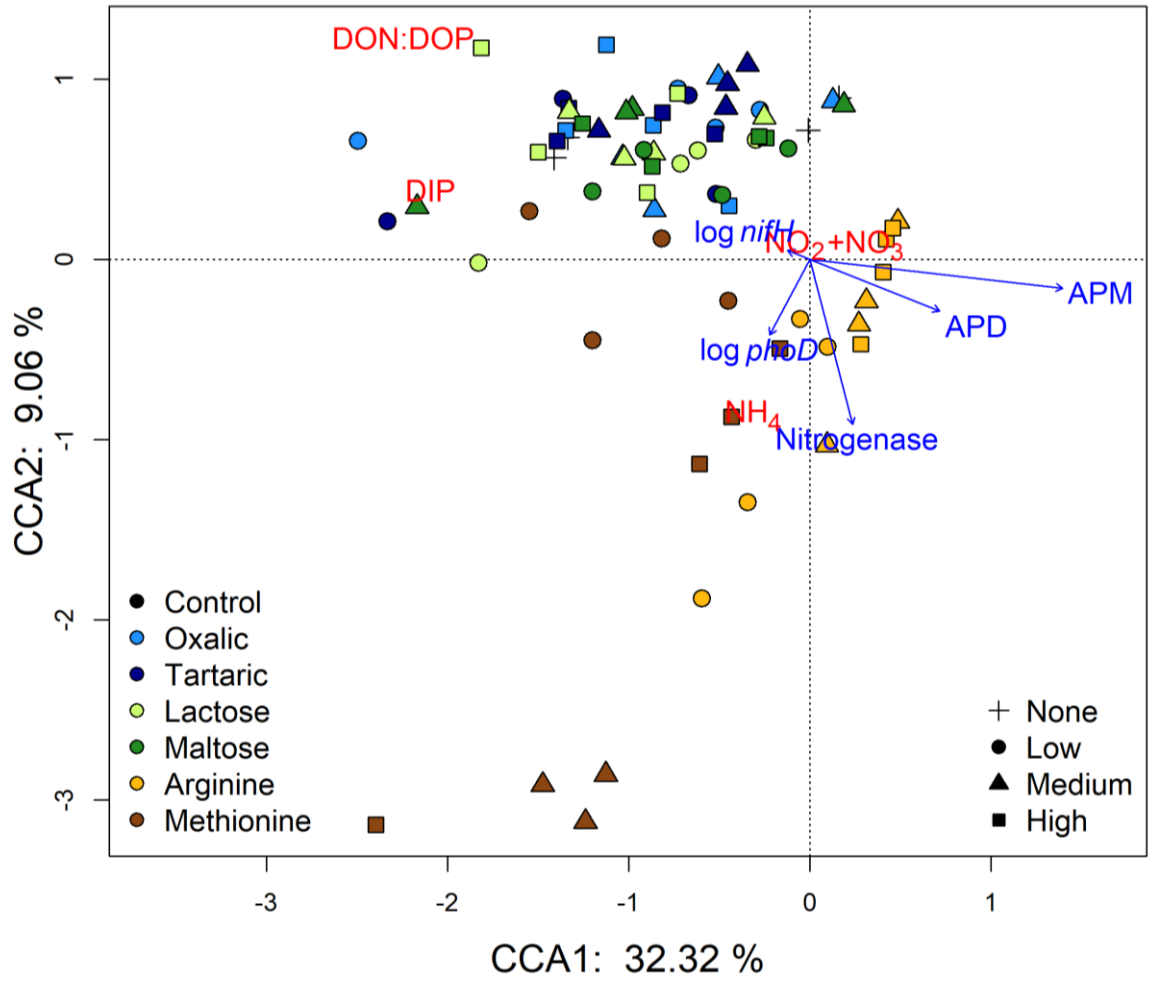



Figure 3.3.4. Canonical correspondence analysis of nutrient concentration in soil samples dosed with metabolites identified in lechuguilla. Ordinated points represent nitrogen and phosphorus concentration in soil samples with linear relation to the enzymatic activity and abundance of NFB and OPMB (Centroids are displayed in red). The color and form of the points depict the metabolite type and the concentration of the treatment. Enzymatic activity and abundance that explain the dispersion of the samples, are represented as blue arrows. Inertia explained by constrained variables = 41.45% ($p<0.001$, permutations= 999).

## 3.4 Discussion

To understand the influence of root exudates on the bacterial communities, we added metabolites identified in the rhizosphere of *A. lechuguilla* T. at different concentrations on bulk soil samples, and studied the variations in the enzymatic activity and abundance of nutrient-improvement bacteria. Furthermore, nitrogen and phosphorus concentration were quantified to analyze the influence of the microorganisms on the soil nutrients. The results suggested that amino acids enhance the enzymatic activity of NFB and OPMB, and this variation may impact the nutrient concentration in soil and its availability. The amino acids Arg and Met could be especially effective in increasing nutrient availability through the enhanced activity of nutrient-improvement bacteria.

### 3.4.1 Arginine and methionine enhanced the enzymatic activity and abundance of nutrient-improvement bacteria

The soil samples employed in this project came from the calcaric soils of the Chihuahuan Desert. These soils are poorly developed, have salinity tendencies and low nutrient concentration (Dinerstein *et al.* 2001; Houri and Machacka-Houri 2016). However, even under these conditions, *A. lechuguilla* T. has a wide distribution that may result from its symbiosis with plant-growth-promoting microorganisms recruited by its root exudates (López-Lozano *et al.* 2020; Medina-de la Rosa *et al*. 2021).

Nonetheless, little is known about the composition of root exudates in plants, especially those from arid regions or with crassulacean acid metabolism (CAM). However, in the few CAM plants studied, it has been found that some of the most abundant compounds released to the soil are sugars, amino acids, organic acids, phenolic compounds, and flavonoids (Stintzing and Carle 2005; Tao *et al.* 2016); which had been associated to the selection of beneficial microorganisms and the mobilization of soil nutrients (Finzi *et al.* 2015; Tian *et al.* 2021). The increase of the microbial activity is known as the “priming effect” and occurs by the enhanced availability of organic carbon sources in the rhizosphere (Kuzyakov and Blagodatskaya 2015), from metabolites like sugars or amino acids (Chen *et al.* 2009;

Huo *et al.* 2017). The sugars, for example, are simple forms of C compounds, easily metabolized by heterotrophic microorganisms, which permits a quick increase of biomass and activity (Chapin *et al.* 2012). As such, the presence of these compounds favors the mineralization of the organic matter since it also influences the activation of the latent microbiome in the soil (Montaño and Sánchez-Yáñez 2014). Thus, it was expected that the addition of the metabolites, increased the presence of NFB and OPMB groups, though the enhancement of the enzymatic activity was mainly with the amino acid treatments. To explain such results, we must consider that sugars and organic acids may function mainly as substrate for the production of energy and signals between the microorganisms. For example, sugars are first metabolized, the products send to the Krebs cycle, and later to the production of energy in the form of adenosine triphosphate (ATP) (Udaondo *et al.* 2018).

The organic acids such as acetic, citric malic, and oxalic are C sources with essential roles in the cellular metabolism, metal chelation, and even as chemoattractants that favor the biological control or the solubilization of organic compounds like phosphorus (Jones, 1998; Kamilova *et al.* 2006; Anandyawati *et al.* 2017), and had been associated with Proteobacteria (Chen *et al.* 2016), a phylum that, according to previous studies, is a dominant taxon in the site, where some of the most abundant genera, were *Sphingomonas*, *MND1*, and *Povalibacter* (See chapter 2). The release of organic acids may be a response to nutrient deficiency, as iron (Fe) and phosphorus (P) are released from their insoluble forms with the action of organic acids (Jones 1998; Kamilova *et al.* 2006). Nonetheless, not all sugars or organic acids increased the enzymatic activity of NFB and OPMB, and even seemed to decrease their activity, so that we may consider the negative effect of these compounds on the microorganisms.

It has been found that some NFB like *Klebsiella pneumonia*, *Enterobacter agglomerans*, *Enterobacter sp*., *Azospirillum lipoferum*, and *Pseudomonas* cannot use lactose, maltose, or oxalate as carbon sources (Haahtela *et al.* 1983), which could be related to the specific metabolic pathways of the bacteria (Tejera *et al.*

2004). Since some of these genera were the most abundant NFB in the soil used in previous studies (See chapter 2), this could explain the lack of response in the nitrogenase activity. In addition, *Pseudomonas* and *Bacillus*, also abundant in these soils, have not shown chemotaxis towards acids like lactic, oxalic, propionic, and ketoglutaric; or sugars like glucose, maltose, sucrose, xylose, or ribose (De Weert *et al.* 2002), while the enzymatic activity of nitrogenase and phosphatases was favorable, even at the low concentrations of Arg and Met (Oku *et al.* 2012; Yang *et al.* 2015; Corral-Lugo *et al.* 2016; Feng *et al.* 2021).

The amino acids can be used directly as structural components of the proteins (Czaban *et al.* 2018) or afterward their mineralization as precursor compounds, energy, carbon, nitrogen, and sulfur sources (Morris *et al.* 2004). For example, when the plants have a high concentration of amino acids, they can suppress their intake of nitrates (Nazoa *et al.* 2003), while sugars and organic acids seem to be more related to P availability (Kellermeir *et al.* 2014). Arg and Met, as other compounds in the root exudates, had a priming effect, that could be related to the change in the C:N ratio in the soil (Ma *et al.* 2017). That is, in poor soils like those from the Chihuahuan Desert with a high C:N ratio (51.35 ± 2.64), due to low N availability (See chapter 2), the addition of carbon from sugars and organic acids increases the C:N ratio and diminish the decomposition efficiency of organic matter; however, when the added compound also has nitrogen like the amino acids, the ratio goes down and the decomposition of organic matter and metabolism of the bacteria increases (Finzi *et al.* 2015). In addition, Arg has four atoms of N and Met one, which implies a higher income of nitrogen from Arg than Met, and could explain why the activity of NFB and OPMB and the mineralization of nitrogen was overall higher in the Arg treatments than any other metabolite addition. Besides, Arg is a versatile compound whose metabolization produces several compounds such as nitrogen oxide, urea, ornithine, creatinine, proline, and polyamines that could be exploited by the microorganisms (Morris *et al.* 2004).

Furthermore, with the addition of Met, nitrogenase activity had the highest rates among the treatments. This behavior could be explained since this compound is an

essential amino acid, as initiator of the protein biosynthesis (Zheng and Dean 1994; Byer *et al*. 2015), which may explain the increased enzymatic activity of nitrogenase and APD. Sulfur-containing amino acids are essential for the tolerance of proteins against oxidation, and in the nitrogenase, S is used for the biosynthesis of the molybdenum-iron cofactor (Byer *et al.* 2015; Si *et al.* 2020).

Amino acid addition promoted higher APD activity than sugars, but seemed to stall in the highest concentration. This behavior could be explained if we consider that the uptake of $^{14}$C-labelled Met was around 35 x $10^2$ disintegrations per minute and 20 x $10^3$ dpm for $^{14}$C-labelled Arg (Goodrich and Morita 1977). Also, it has been proved that the presence of sulfur-containing amino acids like Cys and Met can delay the organic matter mineralization (Hale and Fitzgerald 1988; Spohn Ermak and Kuzyakov 2013). Besides, a toxic effect has been found in amino acids over specific microorganisms (Xie *et al.* 2020). For example, the amino acids in the rhizosphere of lechuguilla, alanine, leucine, lysine, and threonine, showed a negative correlation with APM (Chapter 2). Even as these results are valuable to understand the interactions between plant and bacteria, further research is needed to corroborate this negative effect or the mechanisms behind this behavior.

Furthermore, the low reactivity of APM to treatments, could be related to the stability of the substrate for each enzyme. It is known that diesters like phospholipids, nucleic acids or teichoic acid, are more easily processed than monoesters like inositol phosphate and choline phosphate, which have high negative charge densities and strongly react with inorganic compounds in the soil, such as aluminosilicates or aluminum oxides, making them more stable (Turner *et al.* 2003; Turner *et al.* 2005; García-Oliva *et al.* 2018). In comparison, diesters may be considered labile compounds by their weak interactions, for example, with clay particles in the soil (Rheinheimer *et al.* 2002; Turner *et al.* 2003; Wei *et al.* 2014). However, since monoesters are also more abundant in the soil (Rheinheimer *et al.* 2002; von Wandruszka 2006), it could have influenced the bacterial community and reduced P availability, also controlling the increase of the abundance of OPMB microorganisms.

### 3.4.2 The synergism between nitrogen and phosphorus concentration as result of metabolites presence.

The next objective was to ascertain if the variations in the enzymatic activity and abundance promoted by the metabolites, had an effect over the nutrient availability in the treated soil. In general, we found an increase in the inorganic nitrogen forms ($NO_2$+$NO_3$ and $NH_4$). However, it is interesting to explore the relationship between nitrogen and phosphorus due to changes in enzymatic activity.

Phosphorus concentration in all forms did not report changes as an effect of the treatments compared to the control. This lack of response could be explained by the system employed or the high reactivity and immobilization susceptibility of the element. The system that we utilized was static with a single spike of the metabolites and the incubation of the experimental unit. Changes in the dynamics of this element had been observed under different conditions where mineral fertilizers with N, P, and K were added over long periods in crops (Bi *et al.* 2018). In addition, P is a reactive element that is quickly immobilized by metallic cations such as Ca, Fe, and Al, which decrease its availability for plants and the microbiome (Hinsiger 2001). This effect may be enhanced in the calcaric soils of the Chihuahuan Desert used for these tests and thus lead to low P availability (Crain *et al.* 2018). Besides, the rapid recycling of the nutrient could be reflected in the lack of variation in the P forms, since, as suggested by some authors, the small amounts of compounds may conceal rapid turnover rates (Turner *et al*. 2005).

In turn, organic nitrogen is considered the dominant form in soil (Stevenson 1982). However, it seems that the addition of the metabolites increased the mineralization of nitrogen, mainly the nitrification products, which, could be the response of the priming effect (Jones and Willett 2006; Kuzyakov and Blagodatskaya 2015). Besides, the mineralization process could be encouraged by the accumulation of nitrogen in the soil by the nitrogenase activity and the lack of sinks like the plants in the rhizosphere (Bhattacharjee *et al.* 2008). In addition, it is known that in soils where nitrogen is scarce, the nitrification process is favored (Chapin *et al.* 2011). Thus, if we consider that the action of NFB can add nitrogen and the phosphorus only is

recycled, it could explain, to some extent, that the change in the DON:DOP ratio was based on the concentration of DON (Tian *et al.* 2021). Besides, the linear models suggest a recycling and mineralization of the nutrients since we found that as the organic forms were depleted, the enzymatic activity of nitrogenase and alkaline phosphatases increased, and mineral forms of nitrogen were accumulated (Figure 3.4.1). However, further research is needed to explore directly the dynamics of the nutrients in the rhizosphere. This research is proposed as a leading step towards the understanding of this complex process.

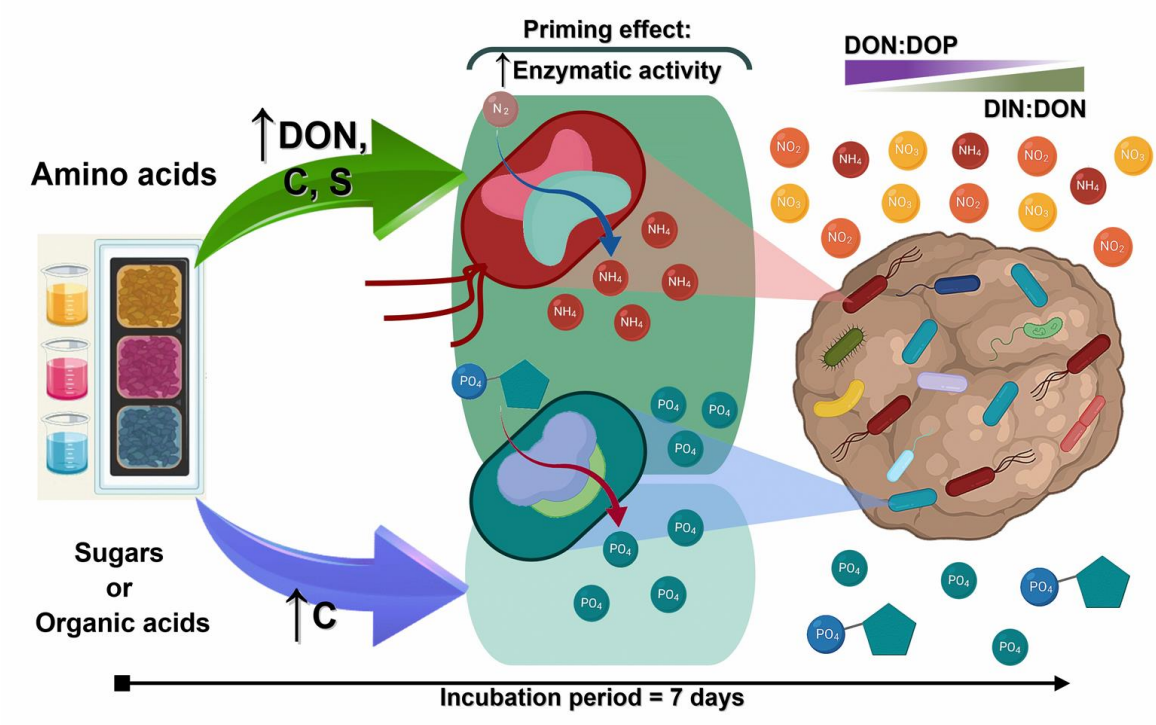



Figure 3.4.1. Theoretical changes in the enzymatic activity of NFB and OPMB, and nutrient availability after treatments with metabolites identified in the rhizosphere of lechuguilla. Amino acids add on organic nitrogen (DON), carbon ©, and, sulfur (S), for Methionine; thus, promoting a priming effect in the bacterial community. The boost in the activity of nitrogenase and phosphatases, increase the availability of nitrogen for the microbiota biomass. The excess is mineralized (DIN) and is in the form of $NO_2$+$NO_3$ at the end of the incubation period. While relation between phosphorus organic (DOP) and inorganic (DIP), remains unchanged. Sugars and organic acids enrich the soil with carbon which only improves the activity of alkaline phosphomonoesterase and -diesterase, however the dynamic of nitrogen and phosphorus remains constant.

## 3.5 Conclusion

Thus far, we explored the influence of different metabolites in the enzymatic activity and abundance of NFB and OPMB, microorganisms that may have an essential role in the adaptability of *A. lechuguilla* T. in the Chihuahuan Desert. We found that amino acids had a significant influence on the activity of nitrogenase and APD since their nitrogen-rich contribution may reduce the C:N ratio of the soil and enhance the organic matter decomposition and thus the recycling of the nutrients. In addition, the relation between N, P, and the enzymatic activity, suggest a complex mechanism for the recycling of the nutrients in the rhizosphere, where different bacteria are involved.

This work is a necessary step towards enhancing the explanation of how plant-growth-promoting bacteria aid plants even under the harsh conditions of deserts. With this knowledge, it could be possible to understand the plants and bacteria symbiosis, as well as the conditions needed to enhance such beneficial partnership. Then, we could develop strategies to improve the conditions of the plants for their establishment even in poorly conserved ecosystems. Furthermore, these studies are essential under the pressure of a growing population, climate change, and the latent need for alternatives to conventional agriculture supplies that may lead to the erosion of the soil and a decreased production. Even the reforestation programs may be benefited if we understand the selection process that favors the symbiosis with plant-growth-promoting rhizobacteria to facilitate the establishment of plants even under harsh conditions of eroded ecosystems or arid environments.

## 3.6 References


Anandyawati, Anwar S, Nugroho B *et al.* Study of Root Exudate Organic Acids and Microbial Population in the Rhizosphere of Oil Palm Seedling. *J Trop Soils* 2017;22:29–36.

Badri D V., Chaparro JM, Zhang R *et al.* Application of natural blends of phytochemicals derived from the root exudates of arabidopsis to the soil reveal that phenolic-related compounds predominantly modulate the soil microbiome. *J Biol Chem* 2013;288:4502–12.

Bahulikar RA, Torres-Jerez I, Worley E *et al.* Diversity of Nitrogen-Fixing Bacteria Associated with Switchgrass in the Native Tallgrass Prairie of Northern Oklahoma. Kostka JE (ed.). *Appl Environ Microbiol* 2014;80:5636–43.

Bezrutczyk M, Yang J, Eom JS *et al.* Sugar flux and signaling in plant–microbe interactions. *Plant J* 2018;93:675–85.

Bhattacharjee RB, Singh A, Mukhopadhyay SN. Use of nitrogen-fixing bacteria as biofertiliser for non-legumes: prospects and challenges. *Appl Microbiol Biotechnol* 2008;80:199–209.

Bi Q-F, Zheng B-X, Lin X-Y *et al.* The microbial cycling of phosphorus on long-term fertilized soil: Insights from phosphate oxygen isotope ratios. *Chem Geol* 2018;483:56–64.

Bray RH, Kurtz LT. Determination of total, organic, and available forms of phosphorus in soils. *Soil Sci* 1945;59:39–46.

Byer AS, Shepard EM, Peters JW *et al.* Radical S-Adenosyl-l-methionine Chemistry in the Synthesis of Hydrogenase and Nitrogenase Metal Cofactors. *J Biol Chem* 2015;290:3987–94.

Chapin FS, Matson PA, Vitousek PM. *Principles of Terrestrial Ecosystem Ecology*. New York, NY: Springer New York, 2011.

Chávez-Vergara B, Merino A, González-Rodríguez A *et al.* Direct and legacy effects of plant-traits control litter decomposition in a deciduous oak forest in Mexico. *PeerJ* 2018;6:e5095.

Chen L-Z, Wang G-H, Hong S *et al.* UV-B-induced Oxidative Damage and Protective Role of Exopolysaccharides in Desert Cyanobacterium *Microcoleus vaginatus*. *J Integr Plant Biol* 2009;51:194–200.

Chen Z-J, Tian Y-H, Zhang Y *et al.* Effects of root organic exudates on rhizosphere microbes and nutrient removal in the constructed wetlands. *Ecol Eng* 2016;92:243–50.

Coleman-Derr D, Desgarennes D, Fonseca-Garcia C *et al.* Plant compartment and biogeography affect microbiome composition in cultivated and native *Agave* species. *New Phytol* 2016;209:798–811.

Corral-Lugo A, De la Torre J, Matilla MA *et al.* Assessment of the contribution of chemoreceptor-based signalling to biofilm formation. *Environ Microbiol* 2016;18:3355–72.

Crain G, McLaren J, Brunner B *et al.* Biologically Available Phosphorus in Biocrust-Dominated Soils of the Chihuahuan Desert. *Soil Syst* 2018;2:56.

Czaban W, Rasmussen J, Laursen BB *et al.* Multiple effects of secondary metabolites on amino acid cycling in white clover rhizosphere. *Soil Biol Biochem* 2018;123:54–63.

Dechassa N, Schenk MK. Exudation of organic anions by roots of cabbage, carrot, and potato as influenced by environmental factors and plant age. *J Plant Nutr Soil Sci* 2004;167:623–9.

Dinerstein E, Olson D, Atchley J *et al.* eds. Ecoregion-Based Conservation in the Chihuahuan Desert. A Biological Assessment*.* 2nd ed. CONABIO-The nature Coservancy-PRONATURA-TEC, 2001.

Eder S, Shi L, Jensen K *et al.* A *Bacillus subtilis* secreted phosphodiesterase/alkaline phosphatase is the product of a Pho regulon gene, *phoD*. *Microbiology* 1996;142:2041–7.

Farrar J, Hawes M, Jones D *et al.* How roots control the flux of carbon to the rhizosphere. *Ecology* 2003;84:827–37.

Fell RD. The qualitative and quantitative analysis of insect hemolymph sugars by high performance thin-layer chromatography. *Comp Biochem Physiol Part A Physiol* 1990;95:539–44.

Feng H, Fu R, Hou X *et al.* Chemotaxis of Beneficial Rhizobacteria to Root Exudates: The First Step towards Root–Microbe Rhizosphere Interactions. *Int J Mol Sci 2021, Vol 22, Page 6655* 2021;22:6655.

Feng H, Zhang N, Du W *et al.* Identification of Chemotaxis Compounds in Root Exudates and Their Sensing Chemoreceptors in Plant-Growth-Promoting Rhizobacteria *Bacillus amyloliquefaciens* SQR9. *Mol Plant-Microbe Interact* 2018;31:995–1005.

Finzi AC, Abramoff RZ, Spiller KS *et al.* Rhizosphere processes are quantitatively important components of terrestrial carbon and nutrient cycles. *Glob Chang Biol* 2015;21:2082–94.

Flores-Núñez VM, Fonseca-García C, Desgarennes D *et al.* Functional Signatures of the Epiphytic Prokaryotic Microbiome of Agaves and Cacti. *Front Microbiol* 2020;10:3044.

Gamalero E, Bona E, Todeschini V *et al.* Saline and Arid Soils: Impact on Bacteria, Plants, and Their Interaction. *Biology (Basel)* 2020;9:116.

García-Oliva F, Merino A, Fonturbel MT *et al.* Severe wildfire hinders renewal of soil P pools by thermal mineralization of organic P in forest soil: Analysis by sequential extraction and 31P NMR spectroscopy. *Geoderma* 2018;309:32–40.

García-Robledo E, Corzo A, Papaspyrou S. A fast and direct spectrophotometric method for the sequential determination of nitrate and nitrite at low concentrations in small volumes. *Mar Chem* 2014;162:30–6.

Giri B, Giang PH, Kumari R *et al.* Microbial Diversity in Soils. In: Buscot F, Varma A (eds.). *Microorganisms in Soils: Roles in Genesis and Functions*. 1st ed. Germany: Springer, 2005, 19–58.

Goodrich TD, Morita RY. Low temperature inhibition on binding, transport, and incorporation of leucine, arginine, methionine, and histidine in Escherichia coli. *Z Allg Mikrobiol* 1977;17:91–7.

Gransee A, Wittenmayer L. Qualitative and quantitative analysis of water-soluble root exudates in relation to plant species and development. *J Plant Nutr Soil Sci* 2000;163:381–5.

Haahtela K, Kari K, Sundman V. Nitrogenase activity (acetylene reduction) of root-associated, cold-climate *Azospirillum*, *Enterobacter*, *Klebsiella*, and *Pseudomonas* species during growth on various carbon sources and at various partial pressures of oxygen. *Appl Environ Microbiol* 1983;45:563–70.

Hale DD, Fitzgerald JW. Generation of sulphate from cysteine in forest soil and litter. *Soil Biol Biochem* 1990;22:427–9.

Han NS, Robyt JF. Separation and detection of sugars and alditols on thin layer chromatograms. *Carbohydr Res* 1998;313:135–7.

Hennion N, Durand M, Vriet C *et al.* Sugars en route to the roots. Transport, metabolism and storage within plant roots and towards microorganisms of the rhizosphere. *Physiol Plant* 2019;165:44–57.

Hinsinger P. Bioavailability of soil inorganic P in the rhizosphere as affected by root-induced chemical changes: A review. *Plant Soil* 2001;237:173–95.

Houri A, Machaka-Houri N. Agave Lechuguilla as a potential biomass source in arid areas. *J Sustain Dev Energy, Water Environ Syst* 2016;4:89–93.

Huang X-F, Chaparro JM, Reardon KF *et al.* Rhizosphere interactions: root exudates, microbes, and microbial communities. *Botany* 2014;92:267–75.

Huo C, Luo Y, Cheng W. Rhizosphere priming effect: A meta-analysis. *Soil Biol Biochem* 2017;111:78–84.

Islam MJ, Jang C, Eum J *et al.* C:N:P stoichiometry of particulate and dissolved organic matter in river waters and changes during decomposition. *J Ecol Environ* 2019;43:4.

Jeyanthi V, Kanimozhi S. Plant Growth Promoting Rhizobacteria (PGPR) - Prospective and Mechanisms: A Review. *J Pure Appl Microbiol* 2018;12:733–49.

Jiang Y, Tian J, Ge F. New Insight into Carboxylic Acid Metabolisms and pH Regulations During Insoluble Phosphate Solubilization Process by Penicillium oxalicum PSF-4. *Curr Microbiol* 2020;77:4095–103.

Joergensen RG, Mueller T. The fumigation-extraction method to estimate soil microbial biomass: Calibration of the kEN value. *Soil Biol Biochem* 1996;28:33–7.

Jones D. Organic acids in the rhizosphere - a critical review. *Plant Soil* 1998;205:25–44.

Jones DL, Darrah PR. Amino-acid influx at the soil-root interface of *Zea mays* L. and its implications in the rhizosphere. *Plant Soil* 1994;163:1–12.

Jones DL, Willett VB. Experimental evaluation of methods to quantify dissolved organic nitrogen (DON) and dissolved organic carbon (DOC) in soil. *Soil Biol Biochem* 2006;38:991–9.

Kamilova F, Kravchenko L V., Shaposhnikov AI *et al.* Organic Acids, Sugars, and l -Tryptophane in Exudates of Vegetables Growing on Stonewool and Their Effects on Activities of Rhizosphere Bacteria. *Mol Plant-Microbe Interact* 2006;19:250–6.

Kellermeier F, Armengaud P, Seditas TJ *et al.* Analysis of the Root System Architecture of *Arabidopsis* Provides a Quantitative Readout of Crosstalk between Nutritional Signals. *Plant Cell* 2014;26:1480–96.

Kuzyakov Y, Blagodatskaya E. Microbial hotspots and hot moments in soil: Concept & review. *Soil Biol Biochem* 2015;83:184–99.

Landi L, Valori F, Ascher J *et al.* Root exudate effects on the bacterial communities, $CO_2$ evolution, nitrogen transformations and ATP content of rhizosphere and bulk soils. *Soil Biol Biochem* 2006;38:509–16.

Lee K-Y, So J-S, Heo T-R. Thin layer chromatographic determination of organic acids for rapid identification of bifidobacteria at genus level. *J Microbiol Methods* 2001;45:1–6.

Leyn SA, Suvorova IA, Kholina TD *et al.* Comparative Genomics of Transcriptional Regulation of Methionine Metabolism in Proteobacteria. Kuipers OP (ed.). *PLoS One* 2014;9:e113714.

Lpez-Lozano NE, Carcaño-Montiel MG, Bashan Y. Using native trees and cacti to improve soil potential nitrogen fixation during long-term restoration of arid lands. *Plant Soil* 2016;403:317–29.

López-Lozano NE, Echeverría Molinar A, Ortiz Durán EA *et al.* Bacterial Diversity and Interaction Networks of *Agave lechuguilla* Rhizosphere Differ Significantly From Bulk Soil in the Oligotrophic Basin of Cuatro Cienegas. *Front Plant Sci* 2020;11:1028.

Ma H, Pei G, Gao R *et al.* Mineralization of amino acids and its signs in nitrogen cycling of forest soil. *Acta Ecol Sin* 2017;37:60–3.

Ma W, Guinel FC, Glick BR. *Rhizobium leguminosarum* Biovar viciae 1-Aminocyclopropane-1-Carboxylate Deaminase Promotes Nodulation of Pea Plants. *Appl Environ Microbiol* 2003;68:4396–402.

Marini F, Speranza S, del Gallo M. Influence of the carbon substrate on the composition of the exocellular polysaccharides by *Azospirillum brasilense*. In: Fendrik I, del Gallo M, Vanderleyden J, et al. (eds.). *Azospirillum VI and Related Microorganisms*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1995, 467–77.

Mattsson T, Kortelainen P, Laubel A *et al.* Export of dissolved organic matter in relation to land use along a European climatic gradient. *Sci Total Environ* 2009;407:1967–76.

Medina-de la Rosa G, García-Oliva F, Alpuche-Solís ÁG *et al.* The nutrient-improvement bacteria selected by *Agave lechuguilla* T. and their role in the rhizosphere community. *FEMS Microbiol Ecol* 2021;97:1–14.

Meier IC, Finzi AC, Phillips RP. Root exudates increase N availability by stimulating microbial turnover of fast-cycling N pools. *Soil Biol Biochem* 2017;106:119–28.

Montaño NM, Sánchez-Yañez JM. Nitrification in tropical soils linked to microbial competition: a model based on Lotka-Volterra theory. *Ecosistemas* 2014;23:98–104.

Morris SM. Enzymes of Arginine Metabolism. *J Nutr* 2004;134:2743S-2747S.

Nazoa P, Vidmar JJ, Tranbarger TJ *et al.* Regulation of the nitrate transporter gene AtNRT2.1 in *Arabidopsis thaliana*: responses to nitrate, amino acids and developmental stage. *Plant Mol Biol 2003 523* 2003;52:689–703.

Nelson DW. Determination of ammonium in KCl extracts of soils by the salicylate method. *Commun Soil Sci Plant Anal* 1983;14:1051–62.

Oku S, Komatsu A, Tajima T *et al.* Identification of Chemotaxis Sensory Proteins for Amino Acids in *Pseudomonas fluorescens* Pf0-1 and Their Involvement in Chemotaxis to Tomato Root Exudate and Root Colonization. *Microbes Environ* 2012;27:462–9.

Olanrewaju OS, Glick BR, Babalola OO. Mechanisms of action of plant growth promoting bacteria. *World J Microbiol Biotechnol* 2017;33:197.

Paungfoo-Lonhienne C, Lonhienne TGA, Rentsch D *et al.* Plants can use protein as a nitrogen source without assistance from other organisms. *Proc Natl Acad Sci* 2008;105:4524–9.

Pereira e Silva MC, Schloter-Hai B, Schloter M *et al.* Temporal Dynamics of Abundance and Composition of Nitrogen-Fixing Communities across Agricultural Soils. Shah V (ed.). *PLoS One* 2013;8:e74500.

Pétriacq P, Williams A, Cotton A *et al.* Metabolite profiling of non-sterile rhizosphere soil. *Plant J* 2017;92:147–62.

Phillips DA, Fox TC, King MD *et al.* Microbial Products Trigger Amino Acid Exudation from Plant Roots. *Plant Physiol* 2004;136:2887–94.

Poly F, Monrozier LJ, Bally R. Improvement in the RFLP procedure for studying the diversity of *nifH* genes in communities of nitrogen fixers in soil. *Res Microbiol* 2001;152:95–103.

Ragot SA, Kertesz MA, Bünemann EK. phoD Alkaline Phosphatase Gene Diversity in Soil. Voordouw G (ed.). *Appl Environ Microbiol* 2015;81:7281–9.

Rheinheimer DS, Anghinoni I, Flores AF. Organic and inorganic phosphorus as characterized by phosphorus-31 nuclear magnetic resonance in subtropical soils under management systems. *Commun Soil Sci Plant Anal* 2002;33:1853–71.

Richardson AE, Simpson RJ. Soil microorganisms mediating phosphorus availability. *Plant Physiol* 2011;156:989–96.

Rodríguez H, Fraga R, Gonzalez T *et al.* Genetics of phosphate solubilization and its potential applications for improving plant growth-promoting bacteria. *Plant Soil* 2006;287:15–21.

Santos AFJ, Moreira ZPM, Souza JT *et al.* Culturable diazotrophic bacterial community associated with *Agave sisalana* P. plants from semi-arid regions in Brazil. *Rev Bras Ciências Agrárias - Brazilian J Agric Sci* 2019;14:1–10.

Dos Santos PC, Fang Z, Mason SW *et al.* Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics* 2012;13:162.

Selvakumar G, Panneerselvam P, An G. Biopesticides Biosafety of Novel Bioinoculants. *J Biofertilizers Biopestic* 2014;5:1–5.

Shi S, Richardson AE, O'Callaghan M *et al.* Effects of selected root exudate components on soil bacterial communities. *FEMS Microbiol Ecol* 2011;77:600–10.

Simonet P, Grosjean MC, Misra AK *et al.* Frankia genus-specific characterization by polymerase chain reaction. *Appl Environ Microbiol* 1991;57:3278–86.

Spohn M, Ermak A, Kuzyakov Y. Microbial gross organic phosphorus mineralization can be stimulated by root exudates – A 33P isotopic dilution study. *Soil Biol Biochem* 2013;65:254–63.

Stevenson FJ ed. *Nitrogen in Agricultural Soils*. Madison, WI, USA: American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, 1982.

Stintzing FC, Carle R. Cactus stems (Opuntia spp.): A review on their chemistry, technology, and uses. *Mol Nutr Food Res* 2005;49:175–94.

Stout LM, Joshi SR, Kana TM *et al.* Microbial activities and phosphorus cycling: An application of oxygen isotope ratios in phosphate. *Geochim Cosmochim Acta* 2014;138:101–16.

Tabatabai MA, Bremner JM. Use of p-nitrophenyl phosphate for assay of soil phosphatase activity. *Soil Biol Biochem* 1969;1:301–7.

Tao Q, Hou D, Yang X *et al.* Oxalate secretion from the root apex of *Sedum alfredii* contributes to hyperaccumulation of Cd. *Plant Soil* 2016;398:139–52.

Tao Q, Zhao J, Li J *et al.* Unique root exudate tartaric acid enhanced cadmium mobilization and uptake in Cd-hyperaccumulator *Sedum alfredii*. *J Hazard Mater* 2020;383:121177.

Tejera NA, Ortega E, Rodés R *et al.* Influence of carbon and nitrogen sources on growth, nitrogenase activity, and carbon metabolism of *Gluconacetobacter diazotrophicus*. *Can J Microbiol* 2004;50:745–50.

Tian J, Ge F, Zhang D *et al.* Roles of Phosphate Solubilizing Microorganisms from Managing Soil Phosphorus Deficiency to Mediating Biogeochemical P Cycle. *Biology (Basel)* 2021;10:1–19.

Trame A-M, Coddington AJ, Paige KN. Field and genetic studies testing optimal outcrossing in *Agave schottii*, a long-lived clonal plant. *Oecologia* 1995;104:93–100.

Turner B, Cade-Menum B, Condron L *et al.* Extraction of soil organic phosphorus. *Talanta* 2005;66:294–306.

Turner BL, Mahieu N, Condronc LM. Organic Geochemistry The phosphorus composition of temperate pasture soils determined by NaOH—EDTA extraction and solution 31 P NMR spectroscopy. *Org Geochem* 2003;34:1199–210.

Turner TR, Ramakrishnan K, Walshaw J *et al.* Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *ISME J* 2013;7:2248–58.

Udaondo Z, Ramos JL, Segura A *et al.* Regulation of carbohydrate degradation pathways in *Pseudomonas* involves a versatile set of transcriptional regulators. *Microb Biotechnol* 2018;11:442–54.

Vives-Peris V, de Ollas C, Gómez-Cadenas A *et al.* Root exudates: from plant to rhizosphere and beyond. *Plant Cell Rep* 2020;39:3–17.

von Wandruszka R. Phosphorus retention in calcareous soils and the effect of organic matter on its mobility. *Geochem Trans* 2006;7:6.

De Weert S, Vermeiren H, Mulders IHM *et al.* Flagella-driven chemotaxis towards exudate components is an important trait for tomato root colonization by *Pseudomonas fluorescens*. *Mol Plant-Microbe Interact* 2002;15:1173–80.

Wei K, Chen Z, Zhu A *et al.* Application of 31P NMR spectroscopy in determining phosphatase activities and P composition in soil aggregates influenced by tillage and residue management practices. *Soil Tillage Res* 2014;138:35–43.

Xie Y, Wang F, Wang K *et al.* Responses of bacterial *phoD* gene abundance and diversity to crop rotation and feedbacks to phosphorus uptake in wheat. *Appl Soil Ecol* 2020;154:103604.

Yang Y, Pollard AM, Höfler C *et al.* Relation between chemotaxis and consumption of amino acids in bacteria. *Mol Microbiol* 2015;96:1272–82.

Yoneyama T, Terakado-Tonooka J, Minamisawa K. Exploration of bacterial $N_2$-fixation systems in association with soil-grown sugarcane, sweet potato, and paddy rice: a review and synthesis. *Soil Sci Plant Nutr* 2017;63:578–90.

Zavala-Hurtado JA, Jiménez M. Diversity and Uniqueness at Its Best: Vegetation of the Chihuahuan Desert. *Plant Di*. Springer, Cham, 2020, 1–17.

Zhang L, Peng Y, Zhou J *et al.* Addition of fructose to the maize hyphosphere increases phosphatase activity by changing bacterial community structure. *Soil Biol Biochem* 2020;142:107724.

Zheng L, Dean DR. Catalytic formation of a nitrogenase iron-sulfur cluster. *J Biol Chem* 1994;269:18723–6.

## Conclusions and perspectives

The interaction plant-microbiome has been extensively studied in several environments (Truu *et al.* 2017; Eida *et al.* 2018; Kong *et al.* 2019; Mosqueira *et al.* 2019), though all the factors involved are not entirely understood. With this work, we offer an insight into biotic and abiotic factors that influence the shape, activity, and abundance of bacterial communities in the rhizosphere of *Agave lechuguilla* Torrey from the Chihuahuan Desert in Mexico. This work was divided in three sections, encompassing the objectives of the project: First, to understand the role of the plant stage and soil properties in the composition and interactions of the bacterial communities, mainly over the nitrogen-fixing bacteria (NFB) and organic phosphorus mineralizing bacteria (OPMB). Then, to deepen our study, we focused our analyses on the functions of the bacterial communities and identified the compounds in the rhizosphere of lechuguilla as, beyond the soil properties, the root exudates seemed to prevail in the bacteria recruitment. Finally, along with sugars and organic acids also found in the rhizosphere of lechuguilla, the amino acids that showed influence over NFB and OPMB, we studied the changes in the abundance and enzymatic activity of these two groups and the dynamics of the nutrients in the soil. In this way, we had a global perspective of the interaction between the metabolites released by plants and bacteria within the soil matrix.

For developing this project, we choose the Chihuahuan Desert, because it is the biggest desertic region in North America and the second most diverse in the world (Granados-Sánchez *et al.* 2011; Villarreal-Quintanilla *et al.* 2017). Among its inhabitants, *A. lechuguilla* T. spreads successfully and widely over 20 million hectares over poorly developed soils (Castillo, Mares and Villavicencio 2011; Ramos *et al.* 2012). Such success is proposed to be due to their symbiosis with the microbiome (López-Lozano *et al.* 2020; Medina-de la Rosa *et al.* 2021), an interaction known as holobiont (Berg *et al.* 2014), which encompasses all the microorganisms associated with the host (Vandenkoornhuyse *et al.* 2015). Previous tudies have considered plant bacteria from different compartments and found a significant reduction from the dominant Bacillales and Actinomycetales in the

rhizosphere to an increased abundance of Rhizobiales and Xanthomonadales in the endosphere (Coleman-Derr *et al.* 2016; Fonseca-García *et al.* 2018). A selection that results from the change in the conditions from the soil to the endosphere. Thus, it may be crucial to the understanding of plant-bacteria interactions, to explore the bacteria associated with the different compartments in lechuguilla.

In this work, we proposed the study of the rhizosphere as it is one of the most diverse among those associated with the plant and the buffer against pathogens and soil properties (Berendsen, Pieterse and Bakker 2017; Dutta and Podile 2017). We found that the rhizobacterial community associated with lechuguilla included several taxa with mechanisms related to nutrient availability, such as *env. OPS17* (Bacteroidetes), *Gemmatimonadaceae uncultured, S0134 terrestrial group, BD211 terrestrial group* (Gemmatimonadetes), *Chthoniobacteracea*, and *Candidatus Udaeobacter* (Verrucomicrobia), which were also the genera with more interactions in the bacterial communities of the rhizosphere of lechuguilla (Chapter 1). Thus, these taxa or 'hubs' may be decisive genera for the stability and communication in the network that represent the associations between bacteria (Poudel *et al.* 2016). In addition, since these communities are exposed to different conditions, with the co-occurrence network models, we could compare the theoretical dynamics of the interactions in bacterial communities from bulk soil and the rhizosphere of young and mature lechuguillas. This way, we found similarities between the three communities and the gradual recruitment of the taxa in the plant stages, where young plants had extensive taxa shared with those in bulk soil but already with specialized plant-growth-promoting rhizobacteria (PGPR).

Moreover, the community composition of the rhizosphere of lechuguilla concurred with the functional prediction based on the 16S rRNA gene sequencing; abundant PGPR and mineralization traits in the rhizobacterial community, and tolerance stress and solubilization functions in the bulk soil. Considering such traits in the diverse bacterial community of bulk soil, it is not surprising that the most abundant phylum was Actinobacteria. Though for deep knowledge about the active functions in the

rhizosphere, an approach with metatranscriptomic or metaproteomic sequencing may be necessary.

Furthermore, with the comparison of the bacterial communities of lechuguilla from two subregions in the Chihuahuan Desert, we could identify the influence of soil properties in their structure and enzymatic activity. For example, the network from the Central Subregion in the north of Mexico was mainly influenced by the sodic soil, alkaline pH, and low nutrient availability. Under these conditions, the important genera for the communication and stability of the network were carbon fixing bacteria like *Candidatus Chloroploca, Kouleothrix,* and *FFCH7168* (Chloroflexi), (Garrity *et al.* 2001; Grouzdev *et al.* 2018); and tolerant genera like *Omnitrophicaeota* (Omnitrophicaeota) which has been isolated from saline estuaries (Baricz *et al.* 2019). In contrast, the bacteria from the Meseta subregion in the northeast of Mexico had richer nutrient-wise conditions than the Central subregion, where the most abundant genera belong to Gemmatiomonadetes and Proteobacteria. Some of which can mineralize organic phosphorus and intervene in the cycling of carbon, nitrogen, and sulfur (Spain, Krumholz and Elshahed 2009; Karp *et al.* 2019). We also contrasted the behavior of Actinobacteria genera like *0319-7L14, Atopobium, Gaiellales uncultured*, and *MB A2 108* in each subregion and found that under the harsh conditions of the Central subregion, these genera had more negative interactions with other bacteria, which could represent competence or depredation (Shi *et al.* 2016). In the Meseta, where the nutrients were less restricted, the same genera had a higher proportion of positive interactions than in limited conditions. This shift probably responds to modulations in the behavior of the bacteria that has been found in strains from the same phylum as *Streptomyces* (Ibrahimi *et al.* 2020).

Overall, the contrast of these two conditions allows us to identify essential factors that shape bacterial communities. Thus, the study at a regional level is essential to understand patterns based on climate, topography, altitude, or soil parental origin, which could shed more light on the microbial communities (Ma *et al.* 2016), especially with co-occurrence networks that represent the interactions inside the communities (Zhang *et al.* 2018).

Along with the selection of the rhizobacterial communities, plant growth was a relevant factor for enhancing the enzymatic activity and abundance of nutrient-improvement bacteria. The copy numbers of genes *nifH* and *phoD* were higher in the rhizosphere of mature plants than the young plants or bulk soil, a pattern found in other works (Hai *et al.* 2009; Jingang *et al.* 2018); however, further research is needed to explain these results as no direct pathways had been described. The enhancement of bacterial biomass according to the plant stage has been found in other species like leguminous, tubers, cultivars, among others (İnceoğlu *et al.* 2010; Haldar and Sengupta 2015; Qiao *et al.* 2017); but there are few studies about crassulacean acid metabolism (CAM) plants (Stintzing and Carle 2005; Tao *et al.* 2016). In this work, we provide insight into the composition and changes of the root exudates of *A. lechuguilla* T. and the effect on the bacterial communities, as more information is needed to unravel the operation of the arid wild ecosystems.

Up to this point, we demonstrated that mature lechuguillas had different recruitment patterns than young plants and bulk soil. As root exudates are the primary form of communication between plants and the microbiome (Broeckling *et al.* 2008; Micallef *et al.* 2009; Liu *et al.* 2017), we identified the composition of amino acids from the rhizosphere of lechuguilla. The amino acids are abundant compounds in the root exudates and have multiple essential functions in living organisms (Jaeger *et al.* 1999; Carvalhais *et al.* 2011), like structural components of proteins, communication signals, and siderophores precursors (Byer *et al.* 2015; Saijo, Loo and Yasuda 2017; Khan, Singh and Srivastava 2018; Si *et al.* 2020). In addition, it has been proposed that their concentration could reflect the nutrimental state of the plant (Jaeger *et al.* 1999).

In this regard, we found that arginine (Arg) and methionine (Met) were significant amino acids either for the composition of the bacterial community or the enzymatic activity of NFB and OPMB. A community where Proteobacteria and Bacteroidetes were more abundant in the rhizosphere of lechuguilla than bulk soil. However, *Bacillus* was the dominant genus of OPMB within this community, while diazotrophic bacteria with the *nifH* gene were mostly among Proteobacteria and Firmicutes. We

propose that Arg and Met may have an essential role for the bacterial community as nutrient sources for the metabolism of the bacteria and the promotion of the enzymes as structural components of the nitrogenase and alkaline phosphomonoesterase (APM). Moreover, the interaction between lechuguilla and its rhizobacterial community could have a two-sided effect were plant and bacteria affect each other (Bakker *et al.* 2014). Recruited NFB, for example, had been shown to enhance the production of amino acids in the rhizosphere of plants due to the increased availability of nitrogen (Hutapea *et al.* 2018; Rahmoune *et al.* 2019). However, this area remains scarcely explored, highlighting that a deeper investigation is needed to understand the full extent of the interactions in the holobiont. With this information, however, we have the first steps towards the understanding of the interaction process between plant-bacteria, which was explored by evaluating the effects of these metabolites in NFB and OPMB. After all, beyond the valuable and needed knowledge, science must be applied to actual problems, like climate change, the desertification process, and the ever-growing demand for food in increasingly eroded crop fields.

Finally, after testing different metabolites identified in the rhizosphere of lechuguilla, we could measure the variations in the enzymatic activity of NFB and OPMB, along with the affected nutrients, nitrogen, and phosphorus. We found that amino acids influenced more the activity of nitrogenase and alkaline phosphodiesterase and the mineralization of nitrogen, probably due to the addition of nitrogen, included in the amino groups. In turn, the supplemented nitrogen reduced the C:N ratio, improved the efficiency of organic matter decomposition and provided the bacteria with carbon, energy, and structural components for the biosynthesis of proteins. We also found that phosphorus may be a highly recalcitrant nutrient in the calcareous soil because the ratio between organic and inorganic forms did not change under any treatment (Deubel *et al.* 2005; von Wandruszka 2006). Once it is liberated by mineralization or solubilization, phosphorus is quickly absorbed by microorganisms and incorporated into their biomass (Crain *et al.* 2018). Therefore, its cycling may remain unchanged without new additions to the system like in agricultural soils where the addition of

mineral fertilizer over a long period, did change the dynamic of the element (Bi *et al.* 2018).

This experiment grants a glimpse into the complexity of the interaction plant-soil-microorganisms. While the use of a mix of several compounds identified in root exudates may be more accurate to their real composition, the single incorporation of metabolites permits us to identify the changes in the community, test hypotheses, and formulate new experiments to explore different settings.

Furthermore, we may even venture to propose that despite the recommendations of the guides for lechuguilla cultivation that estate that this plant do not require a great amount of organic matter in the soil (Martinez, Castillo and Mares 2011), it may be recommended to amend the soils selected for lechuguilla cultivation, in order to increase the nutrient availability and, thus, the bacterial community activity. Considering that PGPR associated to lechuguilla require a lower C:N ratio to increase their activity and the benefits towards the growth and protection of lechuguilla.

The interactions in the holobiont are complex and need to be studied at depth to unravel its mechanisms and understand the processes that sustain the ecosystems, and the strategies plants have developed to adapt to the changing conditions. Thus, if we understand the mechanisms of their interactions and their symbionts, we may be able to develop programs for their application in several areas, from the conservation of the ecosystems to agricultural production.

## References


Bakker MG, Schlatter DC, Otto-Hanson L *et al.* Diffuse symbioses: Roles of plant-plant, plant-microbe and microbe-microbe interactions in structuring the soil microbiome. *Mol Ecol* 2014;23:1571–83.

Baricz A, Chiriac C, Bulzu P-A *et al.* Patterns of prokaryotic diversity in sediments of brackish to hypersaline salt lakes in Romania. *Stud Univ Babes-Bolyai, Biol* 2019;64:96.

Berendsen RL, Pieterse CMJ, Bakker PAHM. The rhizosphere microbiome and plant health. *Trends Plant Sci* 2012;17:478–86.

Berg G, Grube M, Schloter M *et al.* Unraveling the plant microbiome: Looking back and future perspectives. *Front Microbiol* 2014;5:1–7.

Bi Q-F, Zheng B-X, Lin X-Y *et al.* The microbial cycling of phosphorus on long-term fertilized soil: Insights from phosphate oxygen isotope ratios. *Chem Geol* 2018;483:56–64.

Broeckling CD, Broz AK, Bergelson J *et al.* Root exudates regulate soil fungal community composition and diversity. *Appl Environ Microbiol* 2008;74:738–44.

Byer AS, Shepard EM, Peters JW *et al.* Radical S-Adenosyl-l-methionine Chemistry in the Synthesis of Hydrogenase and Nitrogenase Metal Cofactors. *J Biol Chem* 2015;290:3987–94.

Carvalhais LC, Dennis PG, Fedoseyenko D *et al.* Root exudation of sugars, amino acids, and organic acids by maize as affected by nitrogen, phosphorus, potassium, and iron deficiency. *J Plant Nutr Soil Sci* 2011;174:3–11.

Castillo D, Mares O, Villavicencio EE. Lechuguilla (*Agave lechuguilla* Torr.), planta suculenta de importancia económica y social de las zonas áridas y semiáridas de México. *Boletín la Soc Latinoam y del Caribe Cactáceas y otras Suculentas* 2011;8:6–9.

Coleman-Derr D, Desgarennes D, Fonseca-Garcia C *et al.* Plant compartment and biogeography affect microbiome composition in cultivated and native *Agave* species. *New Phytol* 2016;209:798–811.

Crain G, McLaren J, Brunner B *et al.* Biologically Available Phosphorus in Biocrust-Dominated Soils of the Chihuahuan Desert. *Soil Syst* 2018;2:56.

Deubel A, Deubel A, Merbach W *et al.* Influence of Microorganisms on Phosphorus Bioavailability in Soils. *Microorganisms in Soils: Roles in Genesis and Functions*. 2005, 177–91.

Dutta S, Podile AR. Plant Growth Promoting Rhizobacteria (PGPR): the bugs to debug the root zone. *Crit Rev Microbiol* 2017;36:232–44.

Eida AA, Ziegler M, Lafi FF *et al.* Desert plant bacteria reveal host influence and beneficial plant growth properties. Berta G (ed.). *PLoS One* 2018;13:e0208223.

Fonseca-García C, Desgarennes D, Flores-Núñez VM *et al.* The Microbiome of Desert CAM Plants: Lessons From Amplicon Sequencing and Metagenomics. In: Nagarajan M (ed.). *Metagenomics. Perspectives, Methods, and Applications*. Academic Press, 2018, 231–54.

Garrity GM, Holt JG, Castenholz RW *et al.* Phylum BVI. Chloroflexi phy. nov. *Bergey's Manual® of Systematic Bacteriology*. New York, NY: Springer New York, 2001, 427–46.

Granados-Sánchez D, Sánchez-González A, Granados-Victorino RL *et al.* Ecología de la vegetación del Desierto Chihuahuense. *Rev Chapingo Ser Ciencias For y del Ambient* 2011;XVII:111–30.

Grouzdev DS, Rysina MS, Bryantseva IA *et al.* Draft genome sequences of '*Candidatus Chloroploca asiatica*' and '*Candidatus Viridilinea mediisalina*', candidate representatives of the Chloroflexales order: phylogenetic and taxonomic implications. *Stand Genomic Sci* 2018;13:24.

Hai B, Diallo NH, Sall S *et al.* Quantification of key genes steering the microbial nitrogen cycle in the rhizosphere of sorghum cultivars in tropical agroecosystems. *Appl Environ Microbiol* 2009;75:4993–5000.

Haldar S, Sengupta S. Impact of plant development on the rhizobacterial population of *Arachis hypogaea*: a multifactorial analysis. *J Basic Microbiol* 2015;55:922–8.

Hutapea PS, Abdullah L, Karti PDMH *et al.* Improvement of *Indigofera zollingeriana* Production and Methionine Content through Inoculation of Nitrogen-Fixing Bacteria. *Trop Anim Sci J* 2018;41:37–45.

Ibrahimi M, Korichi W, Hafidi M *et al.* Marine Actinobacteria: Screening for Predation Leads to the Discovery of Potential New Drugs against Multidrug-Resistant Bacteria. *Antibiot (Basel, Switzerland)* 2020;9, DOI: 10.3390/antibiotics9020091.

İnceoğlu Ö, Salles JF, van Overbeek L *et al.* Effects of Plant Genotype and Growth Stage on the Betaproteobacterial Communities Associated with Different Potato Cultivars in Two Fields. *Appl Environ Microbiol* 2010;76:3675–84.

Jaeger CH, Lindow SE, Miller W *et al.* Mapping of Sugar and Amino Acid Availability in Soil around Roots with Bacterial Sensors of Sucrose and Tryptophan. *Appl Environ Microbiol* 1999;65:2685–90.

Jingang L, Ying L, Yue J *et al.* High-methionine soybean has no significant effect on nitrogen-transforming bacteria in rhizosphere soil. *Plant, Soil Environ* 2018;64:108–13.

Karp PD, Billington R, Caspi R *et al.* The BioCyc collection of microbial genomes and metabolic pathways. *Brief Bioinform* 2019;20:1085–93.

Khan A, Singh P, Srivastava A. Synthesis, nature and utility of universal iron chelator – Siderophore: A review. *Microbiol Res* 2018;212–213:103–11.

Kong P, Hong C. Soil bacteria as sources of virulence signal providers promoting plant infection by Phytophthora pathogens. *Sci Rep* 2016;6:33239.

Liu Y, Chen L, Wu G *et al.* Identification of Root-Secreted Compounds Involved in the Communication Between Cucumber, the Beneficial *Bacillus amyloliquefaciens*, and the Soil-Borne Pathogen *Fusarium oxysporum*. *Mol Plant-Microbe Interact* 2017;30:53–62.

Ma B, Wang H, Dsouza M *et al.* Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. *ISME J* 2016;10:1891–901.

Martinez OU, Castillo D, Mares O. Caracterización y selección de sitios para plantaciones de lechuguilla (Agave lechuguilla Torr.) en el estado de Coahuila. 2011:41.

Micallef SA, Channer S, Shiaris MP *et al.* Plant age and genotype impact the progression of bacterial community succession in the *Arabidopsis* rhizosphere. *Plant Signal Behav* 2009;4:777–80.

Morris SM. Enzymes of Arginine Metabolism. *J Nutr* 2004;134:2743S-2747S.

Mosqueira MJ, Marasco R, Fusi M *et al.* Consistent bacterial selection by date palm root system across heterogeneous desert oasis agroecosystems. *Sci Rep* 2019;9:4033.

Poudel R, Jumpponen A, Schlatter DC *et al.* Microbiome Networks: A Systems Framework for Identifying Candidate Microbial Assemblages for Disease Management. *Phytopathology*® 2016;106:1083–96.

Qiao Q, Wang F, Zhang J *et al.* The Variation in the Rhizosphere Microbiome of Cotton with Soil Type, Genotype and Developmental Stage. *Sci Rep* 2017;7:3940.

Rahmoune B, Zerrouk IZ, Bouzaa S *et al.* Amino acids profiling in *Datura stramonium* and study of their variations after inoculation with plant growth promoting Rhizobacteria. *Biologia (Bratisl)* 2019;74:1373–83.

Ramos F, Oranday A, Rivas C *et al.* Cytotoxic activity of *Agave lechuguilla* Torr. *AFRICAN J Biotechnol* 2012;11:12229–31.

Saijo Y, Loo EP, Yasuda S. Pattern recognition receptors and signaling in plant-microbe interactions. *Plant J* 2017:592–613.

Shi S, Nuccio EE, Shi ZJ *et al.* The interconnected rhizosphere: High network complexity dominates rhizosphere assemblages. Johnson N (ed.). *Ecol Lett* 2016;19:926–36.

Si Z, Guan N, Zhou Y *et al.* A Methionine Sulfoxide Reductase B Is Required for the Establishment of *Astragalus sinicus–Mesorhizobium* Symbiosis. *Plant Cell Physiol* 2020;61:1631–45.

Spain AM, Krumholz LR, Elshahed MS. Abundance, composition, diversity and novelty of soil Proteobacteria. *ISME J* 2009;3:992–1000.

Stintzing FC, Carle R. Cactus stems (*Opuntia* spp.): A review on their chemistry, technology, and uses. *Mol Nutr Food Res* 2005;49:175–94.

Tao Q, Hou D, Yang X *et al.* Oxalate secretion from the root apex of *Sedum alfredii* contributes to hyperaccumulation of Cd. *Plant Soil* 2016;398:139–52.

Truu M, Ostonen I, Preem J-K *et al.* Elevated Air Humidity Changes Soil Bacterial Community Structure in the Silver Birch Stand. *Front Microbiol* 2017;8:1–15.

Vandenkoornhuyse P, Quaiser A, Duhamel M *et al.* The importance of the microbiome of the plant holobiont. *New Phytol* 2015;206:1196–206.

Villarreal-Quintanilla JA, Bartolomé-Hernández JA, Estrada-Castillón E *et al.* El elemento endémico de la flora vascular del Desierto Chihuahuense. *Acta Bot Mex* 2017;Enero:65–96.

von Wandruszka R. Phosphorus retention in calcareous soils and the effect of organic matter on its mobility. *Geochem Trans* 2006;7:6.

## Supplementary materials

Table 0-1. Alpha diversity indices based on gene 16S rRNA sequencing on Central subregion samples. Sample ID provides indicate if locations correspond to Los Álamos (AL) or Tres Coronas (TC), type (MP- Mature plant, YP- Young plant, and S- Bulk soil), and replicates. DNA sample concentration and number of reads are also depicted.

| Central Subregion | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Sample type** | **Sample ID** | **DNA yield µg $g^{-1}$ soil** | **Initial reads** | **Reads after filtering** | **Observed OTU** | **Chao 1** | **Shannon** | **Simpson** |
| **Mature plant** | **ALMP01** | 14.59 | 281,961 | 111,693 | 2,073 | 3,144 | 5.00 | 0.03 |
| | **ALMP02** | 16.38 | 289,282 | 115,693 | 2,494 | 3,844 | 5.35 | 0.02 |
| | **ALMP03** | 13.54 | 297,543 | 115,882 | 2,277 | 3,597 | 5.13 | 0.02 |
| | **ALMP04** | 10.83 | 343,583 | 134,956 | 2,146 | 3,189 | 5.00 | 0.02 |
| | **TCMP01** | 12.00 | 261,837 | 103,505 | 2,382 | 3,525 | 5.42 | 0.01 |
| | **TCMP02** | 18.63 | 390,446 | 122,657 | 2,152 | 3,306 | 5.34 | 0.01 |
| | **TCMP03** | 13.09 | 335,636 | 131,848 | 2,280 | 3,593 | 5.33 | 0.02 |
| | **TCMP04** | 10.90 | 28,421 | 113,131 | 2,237 | 3,532 | 5.14 | 0.02 |
| **Young plant** | **ALYP01** | 17.25 | 287,912 | 114,783 | 2,296 | 3,479 | 5.30 | 0.01 |
| | **ALYP02** | 17.73 | 265,478 | 102,973 | 2,291 | 3,676 | 5.28 | 0.02 |
| | **ALYP03** | 12.77 | 275,109 | 107,330 | 2,072 | 3,246 | 5.05 | 0.02 |
| | **ALYP04** | 13.00 | 185,964 | 74,083 | 2,216 | 3,443 | 5.06 | 0.02 |
| | **TCYP01** | 14.31 | 307,800 | 118,795 | 2,316 | 3,655 | 5.33 | 0.01 |
| | **TCYP02** | 9.60 | 300,322 | 116,921 | 2,215 | 3,313 | 5.30 | 0.01 |
| | **TCYP03** | 12.13 | 284,506 | 110,960 | 2,266 | 3,741 | 5.29 | 0.01 |
| | **TCYP04** | 10.82 | 204,286 | 80,365 | 1,845 | 2,907 | 5.06 | 0.02 |
| **Bulk soil** | **ALS01** | 10.61 | 252,598 | 100,546 | 2,381 | 3,773 | 5.39 | 0.01 |
| | **ALS02** | 12.76 | 247,587 | 98,095 | 2,336 | 3,388 | 5.40 | 0.01 |
| | **ALS03** | 14.63 | 303,533 | 117,637 | 2,205 | 3,190 | 5.37 | 0.01 |
| | **ALS04** | 12.36 | 240,461 | 95,611 | 2,162 | 3,371 | 5.12 | 0.02 |
| | **TCS01** | 8.66 | 365,077 | 145,819 | 2,029 | 3,094 | 5.11 | 0.02 |
| | **TCS02** | 9.95 | 94,431 | 1,744 | - | - | - | - |
| | **TCS03** | 8.66 | 236,781 | 96,698 | 1,876 | 2,698 | 5.00 | 0.02 |
| | **TCS04** | 9.62 | 441,968 | 132,489 | 2,013 | 2,918 | 5.23 | 0.02 |
| **Mock community** | | 8.30 | 231,379 | 92,711 | - | - | - | - |
| **Reagent only** | | Out of range | 58,451 | 4,775 | 576 | 1,061 | 1.12 | 0.55 |

Table 0-2. Alpha diversity indices based on gene 16S rRNA sequencing on Meseta Central subregion samples. Sample ID provides indicate if locations correspond to Charco Blanco (CB) or Los Amoles (AM), type (MP- Mature plant, YP- Young plant, and S- Bulk soil), and replicates. DNA sample concentration and number of reads are also depicted.

| Meseta Central Subregion | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Soil sample type** | **Sample ID** | **DNA yield (µg $g^{-1}$ soil)** | **Initial reads** | **Reads after filtering** | **Observed OTU** | **Chao 1** | **Shannon** | **Simpson** |
| **Mature plant** | **CBMP01** | 6.69 | 288,147 | 94,054 | 2,199 | 3,378 | 5.20 | 0.01 |
| | **CBMP02** | 9.01 | 269,394 | 98,086 | 2,106 | 3,233 | 5.25 | 0.01 |
| | **CBMP03** | 9.84 | 256,200 | 118,433 | 2,028 | 3,324 | 5.20 | 0.01 |
| | **CBMP04** | 6.37 | 436,918 | 91,892 | 2,052 | 2,993 | 5.17 | 0.01 |
| | **AMMP01** | 8.17 | 240,187 | 114,295 | 2,221 | 3,470 | 5.19 | 0.02 |
| | **AMMP02** | 6.13 | 242,810 | 105,086 | 2,362 | 3,641 | 5.10 | 0.02 |
| | **AMMP03** | 5.83 | 302,294 | 102,028 | 2,381 | 3,819 | 5.15 | 0.02 |
| | **AMMP04** | 5.75 | 241,752 | 175,745 | 2,528 | 3,768 | 5.37 | 0.01 |
| **Young plant** | **CBYP01** | 14.27 | 321,385 | 149,058 | 2,048 | 3,383 | 5.07 | 0.02 |
| | **CBYP02** | 10.52 | 299,852 | 151,788 | 2,289 | 3,781 | 5.14 | 0.01 |
| | **CBYP03** | 12.95 | 293,667 | 145,718 | 2,281 | 3,723 | 5.23 | 0.01 |
| | **CBYP04** | 11.28 | 231,613 | 127,314 | 2,185 | 3,306 | 5.11 | 0.01 |
| | **AMYP01** | 7.06 | 381,716 | 125,056 | 2,412 | 3,704 | 5.22 | 0.02 |
| | **AMYP02** | 4.87 | 383,713 | 118,341 | 2,364 | 3,568 | 5.22 | 0.01 |
| | **AMYP03** | 3.98 | 375,139 | 115,846 | 1,976 | 3,205 | 4.85 | 0.02 |
| | **AMYP04** | 6.41 | 322,442 | 91,072 | 2,329 | 3,346 | 5.23 | 0.02 |
| **Bulk soil** | **CBS01** | 8.84 | 409,317 | 106,582 | 1,924 | 3,107 | 4.89 | 0.02 |
| | **CBS02** | 9.66 | 418,361 | 97,568 | 1,918 | 3,232 | 4.95 | 0.02 |
| | **CBS03** | 2.96 | 353,214 | 95,406 | 2,028 | 3,147 | 4.94 | 0.02 |
| | **CBS04** | 4.82 | 318,958 | 100,048 | 2,025 | 3,206 | 5.01 | 0.02 |
| | **AMS01** | 3.37 | 271,155 | 155,797 | 2,286 | 3,610 | 5.22 | 0.02 |
| | **AMS02** | 3.17 | 254,556 | 161,960 | 2,577 | 4,044 | 5.24 | 0.02 |
| | **AMS03** | 2.42 | 244,024 | 137,488 | 2,310 | 3,472 | 5.06 | 0.02 |
| | **AMS04** | 3.18 | 254,516 | 121,656 | 2,435 | 3,950 | 5.16 | 0.02 |
| **Mock community** | | 8.30 | 231,379 | 92,711 | - | - | - | - |
| **Reagent only** | | Out of range | 58,451 | 4,775 | 576 | 1,061 | 1.12 | 0.55 |

Table 0-3. Percent increase of mean squared error (%IncMSE) of soil variables. Adjusted to the linear regression model (randomForest), to explain the variation of enzymatic activity/abundance of nitrogen fixing bacteria and organic phosphate mineralizing bacteria. (No. trees = 1000, Variables tested in each split = 7). MSR= Mean squared residuals, Exp. Var.= Explained variation.

| Variable | Organic phosphorus mineralizing bacteria | | Nitrogen fixing bacteria | |
|---|---|---|---|---|
| | Alkaline phosphomonoesterase (μg pNP $g^{-1}$ $h^{-1}$ soil) | *phoD* (log copies *phoD* $g^{-1}$ soil) | Nitrogenase (nmol $C_2H_4$ $kg^{-1}$ $day^{-1}$) | *nifH* (log copies *nifH* $g^{-1}$ soil) |
| | %IncMSE | | | |
| **Sample type** | 3.77 | 4.81 | 7.47 | 11.97 |
| **Total Org. C** | 13.04 | 8.40 | 14.38 | 12.04 |
| **N** | 7.31 | -0.63 | 3.47 | -1.19 |
| **C:N ratio** | 16.55 | 7.17 | 18.66 | 12.03 |
| **$NH_4$** | 8.60 | -1.53 | 6.64 | 2.11 |
| **$NO_2$+$NO_3$** | 13.31 | 4.93 | 9.30 | 7.55 |
| **P** | 1.38 | -1.23 | 1.29 | -3.43 |
| **Ca** | 0.83 | 2.81 | -0.90 | 1.51 |
| **Na** | 19.49 | 5.29 | 12.73 | 14.47 |
| **K** | 2.12 | -0.28 | 2.17 | 1.49 |
| **Mg** | 4.65 | 4.62 | 6.68 | 2.35 |
| **Cu** | 2.08 | 0.18 | 1.23 | 2.67 |
| **Fe** | 2.35 | -0.28 | 6.56 | 3.31 |
| **Mn** | 5.84 | 6.01 | 2.60 | -0.07 |
| **Zn** | 0.87 | 2.14 | 1.39 | -1.62 |
| **pH** | 5.60 | 3.48 | 4.70 | 1.33 |
| **EC** | 3.81 | 0.49 | 0.90 | 3.58 |
| **Clay** | 2.89 | 0.83 | 2.49 | 3.33 |
| **Sand** | 4.71 | 5.44 | 3.81 | 3.12 |
| **Silt** | 4.33 | 4.42 | 4.23 | 5.05 |
| **MSR (Exp. Var.)** | 1169 (96.51%) | 0.39 (13.04%) | 760.55 (68.71%) | 0.18 (74.59%) |

Table 0-4. Topological properties of bacterial empiric subregion networks and comparison with random networks. Empiric networks were made with relative abundance from bacterial genera found in the rhizosphere of *A. lechuguilla* T., from Central and Meseta subregions inside the Chihuahuan desert. Data from random networks corresponds to the mean and standard error from permutations (n = 1000) of each empiric network. Scale-free distribution was evaluated with the fitting of the degree to the power-law model. $R^2$ and exponent α are displayed.

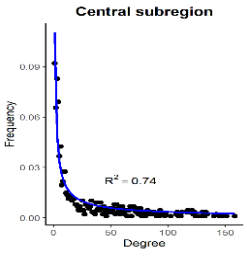
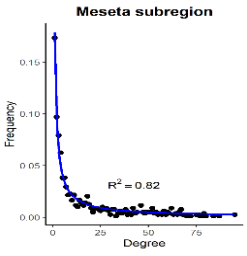
| | | Central Subregion | Meseta Central Subregion |
|---|---|---|---|
| **Topological properties (undirected network)** | **Vertices (Edges)** | 868 (12698) | 785 (6626) |
| | **Diameter** | 7.24 | 8.06 |
| | **Mean distance** | 3.34 | 3.66 |
| | **Mean degree** | 29.26 | 16.88 |
| | **Eigen value** | 0.87 | 0.87 |
| | **Betweeness** | 0.02 | 0.02 |
| | **Hub score** | 0.12 | 0.12 |
| **Hubs** | **Genera (Max degree)** | *Omnitrophicaeota* (158) | *MBNT15 (Proteobacteria)* (95) |
| **Modularity** | **Value (No. Modules)** | 0.35 (215) | 0.44 (168) |
| | **Modules >4 nodes** | 20 | 10 |
| **Max degree** | **Size Mod 1** | 181 | 221 |
| | **Node max degree** | *Omnitrophicaeota (158)* | *MBNT15, Paraoerskovia (95)* |
| | **Size Mod 2** | 149 | 122 |
| | **Node max degree** | *Gaiellales uncultured (140)* | *Candidatus Jorgensenbacteria* |
| | **Size Mod 3** | 134 | 119 |
| | **Node max degree** | *Subgroup 5 (105)* | *A4b (84)* |
| | **Size Mod 4** | 73 | 24 |
| | **Node max degree** | *Bacteriap25 (73)* | *Oceanobacillus (20)* |
| | **Size Mod 5** | 8 | 9 |
| | **Node max degree** | *Oxalophagus (19)* | *Blastomonas (14)* |
| **Nodes Centrality measures** | **Degree** | 29.26 ± 1.20 | 16.88 ± 0.73 |
| | **Betweenness** | 412.08 ± 20.93 | 303.10 ± 20.61 |
| | **Eigen vector** | 0.12 ± 0.01 | 0.12 ± 0.01 |
| | **Hub score** | 0.09 ± 0.01 | 0.08 ± 0.01 |
| **Random Networks** | **Distance** | 2.6 ± 0.01 | 2.87 ± 0.01 |
| | **Eigen vector** | 0.82 ± 0.01 | 0.83 ± 0.01 |
| | **Betweenness** | 0.02 ± 0.00 | 0.02 ± 0.00 |
| **Scale free distribution** | **α** | 1.26 | 1.35 |
| | **$R^2$** | 0.74 | 0.81 |
| **Degree distribution (free-scale model) [In black node degree distribution of empirical networks, random distribution in blue]** | |   |   |

Table 0-5. Topological properties of bacterial empiric sample type networks and comparison with random networks. Empiric networks were made with relative abundance from bacterial genera found in the rhizosphere of mature and young plants of *A. lechuguilla* T., from Central and Meseta subregions inside the Chihuahuan desert. Data from random networks corresponds to the mean and standard error from permutations (n = 1000) of each empiric network. Scale-free distribution was evaluated with the fitting of the degree to the power-law model. $R^2$ and exponent α are displayed.

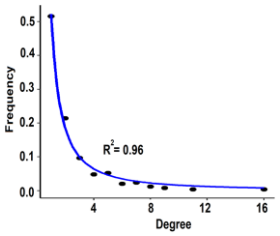
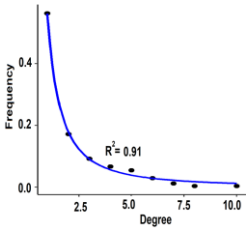
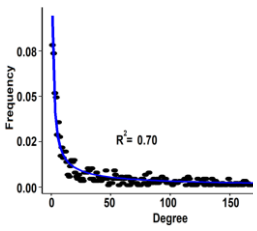
| | | Mature plant | Young plant | Bulk soil |
|---|---|---|---|---|
| **Topological properties (undirected network)** | **Vertices (Edges)** | 248 (280) | 238 (246) | 869 (17275) |
| | **Diameter** | 15.89 | 14.07 | 10.68 |
| | **Mean distance** | 6.44 | 5.99 | 3.43 |
| | **Mean degree** | 2.26 | 2.07 | 39.76 |
| | **Eigen value** | 0.96 | 0.95 | 0.83 |
| | **Betweenness** | 0.12 | 0.08 | 0.01 |
| | **Hub score** | 0.05 | 0.06 | 0.17 |
| **Hubs** | **Genera (Max degree** | *S0134 terrestrial group (Gemmatimonadetes)*16 | *env.OPS 17 (Bacteroidetes)*10 | *Roseiflexus (Chloroflexi)*173 |
| **Modularity** | **Value (Modules)** | 0.77 (66) | 0.83 (60) | 0.3 (201) |
| | **Modules >4 nodes** | 16 | 12 | 22 |
| **Max degree** | **Size Mod 1** | 46 | 27 | 255 |
| | **Node max degree** | *S0134 terrestrial group (16)* | *env.OPS 17 (10)* | *Roseiflexus (173)* |
| | **Size Mod 2** | 10 | 24 | 136 |
| | **Node max degree** | *Rhodocytophaga (5)* | *0319_6G20 (7)* | *DS 100 (164)* |
| | **Size Mod 3** | 10 | 17 | 72 |
| | **Node max degree** | *AKIW781, Kallotenue (7)* | *Isosphaeraceae unc. (6)* | *RB41 (77)* |
| | **Size Mod 4** | 10 | 15 | 63 |
| | **Node max degree** | *OM190, Subgroup 5 (6)* | *Blastocatella (6)* | *Thermobaculum (95)* |
| | **Size Mod 5** | 7 | 9 | 28 |
| | **Node max degree** | *TK17 (5)* | *TSBb06* | *Blyi10 (42)* |
| **Nodes Centrality measures** | **Degree** | 2.26 ± 0.13 | 2.07 ± 0.11 | 39.76 ± 1.54 |
| | **Betweenness** | 2.16 ± 0.53 | 1.62 ± 0.33 | 412.52 ± 19.59 |
| | **Eigen vector** | 0.05 ± 0.01 | 0.06 ± 0.01 | 0.17 ± 0.01 |
| | **Hub score** | 0.03 ± 0.01 | 0.03 ± 0.01 | 0.12 ± 0.01 |
| **Random Networks** | **Distance** | 5.42 ± 0.20 | 6.1 ± 0.32 | 2.48 ± 0.00 |
| | **Eigen vector** | 0.95 ± 0.01 | 0.93 ± 0.02 | 0.78 ± 0.01 |
| | **Betweenness** | 0.19 ± 0.04 | 0.14 ± 0.03 | 0.01 ± 0.00 |
| **Scale free distribution** | **α** | 2.02 | 2.09 | 1.21 |
| | **$R^2$** | 0.96 | 0.91 | 0.7 |
| **Degree distribution (free-scale model) [In black node degree distribution of empirical networks, random distribution in blue]** | |   |   |   |

Table 0-6. Alpha diversity indices based on 16S rRNA gene from the bacterial communities of the rhizosphere of lechuguilla and bulk soil. *The number of reads post filtering of the blank sample may correspond to sequences formed during the reaction and not to bacteria, which discarded the contamination of the samples during their manipulation.

| Type | Sample ID | DNA yield (µg $g^{-1}$ soil) | Initial reads | Reads post-filtering | Observed OTU | Chao 1 | Shannon | Simpson |
|---|---|---|---|---|---|---|---|---|
| **Mature plant** | MP01 | 8.17 | 240,187 | 114,295 | 2,199 | 3,378 | 5.20 | 0.01 |
| | MP02 | 6.13 | 242,810 | 105,086 | 2,106 | 3,233 | 5.25 | 0.01 |
| | MP03 | 5.83 | 302,294 | 102,028 | 2,028 | 3,324 | 5.20 | 0.01 |
| | MP04 | 5.75 | 241,752 | 175,745 | 2,052 | 2,993 | 5.17 | 0.01 |
| **Young plant** | YP01 | 7.06 | 381,716 | 125,056 | 2,048 | 3,383 | 5.07 | 0.02 |
| | YP02 | 4.87 | 383,713 | 118,341 | 2,289 | 3,781 | 5.14 | 0.01 |
| | YP03 | 3.98 | 375,139 | 115,846 | 2,281 | 3,723 | 5.23 | 0.01 |
| | YP04 | 6.41 | 322,442 | 91,072 | 2,185 | 3,306 | 5.11 | 0.01 |
| **Bulk soil** | S01 | 3.37 | 271,155 | 155,797 | 1,924 | 3,107 | 4.89 | 0.02 |
| | S02 | 3.17 | 254,556 | 161,960 | 1,918 | 3,232 | 4.95 | 0.02 |
| | S03 | 2.42 | 244,024 | 137,488 | 2,028 | 3,147 | 4.94 | 0.02 |
| | S04 | 3.18 | 254,516 | 121,656 | 2,025 | 3,206 | 5.01 | 0.02 |
| | Mock | 8.30 | 231,379 | 92,711 | 576 | 1,061 | 1.12 | 0.55 |
| | Blank | Out of range | 58,451 | 4,775 | 0* | | | |

Table 0-7. *nifH* and *phoD* genes sequencing data from the rhizosphere of *A. lechuguilla* T. and bulk soil. Observed ASV, initial reads and post-filtering. ID provides information about the sample type (MP: Mature plant, YP: Young plant, and S: Bulk soil) and replicates number.

| Type | Sample ID | *nifH* | | | *phoD* | | |
|---|---|---|---|---|---|---|---|
| | | Initial reads | Reads post-filtering | Observed ASV | Initial reads | Reads post-filtering | Observed ASV |
| **Mature plant** | MP01 | 171,419 | 18,834 | 89 | 37,936 | 35,353 | 347 |
| | MP02 | 142,718 | 12,460 | 79 | 61,987 | 57,869 | 343 |
| | MP03 | 203,557 | 25,325 | 75 | 55,063 | 51,400 | 374 |
| | MP04 | 162,180 | 6,591 | 97 | 72,798 | 67,099 | 397 |
| **Young plant** | YP01 | 113,882 | 4,313 | 85 | 53,752 | 50,488 | 337 |
| | YP02 | 206,345 | 38,577 | 78 | 53,501 | 46,806 | 370 |
| | YP03 | 122,658 | 9,538 | 90 | 44,184 | 40,358 | 363 |
| | YP04 | 76,393 | 4,953 | 61 | 58,833 | 49,457 | 363 |
| **Bulk soil** | S01 | 132,047 | 28,183 | 75 | 57,189 | 51,953 | 343 |
| | S02 | 98,548 | 13,755 | 73 | 70,004 | 62,647 | 334 |
| | S03 | 103,764 | 11,101 | 71 | 52,841 | 47,073 | 256 |
| | S04 | - | - | - | 56,197 | 52,777 | 347 |
| **Mock** | | - | - | - | 108,236 | 104,447 | 12 |

## R scrips for data analyses

All scripts were tested in R version 4.1.0 (2021-05-18) -- “Camp Pontanezen”. Copyright © 2021 The R Foundation for Statistical Computing; Platform: x86_64-w64-mingw32/x64 (64-bit). In the platform Rstudio version 1.4.1717© 2009-2021 RStudio PBC, “Juliet Rose” (df86b69e, 2021-05-24) for Windows.

### a) Functional prediction with Tax4Fun2 (Tax4Fun2 V 1.1.5)

**## Load packages**

```
“”{r}
X= c(“Tax4Fun2”, “phyloseq”, “seqinr”)
lapply(X, require, character.only = TRUE)

rm(list=ls())
#Reference data pathway
path <- “C:/Users/gpe_/Documents/Tax4Fun2_ReferenceData_v2”
“”
```

Reference on the web: https://github.com/bwemheu/Tax4Fun2

**# Package installation and loading of reference database**

```
“”{r}
buildReferenceData(path_to_working_directory = “.”, use_force = FALSE,
install_suggested_packages = TRUE)
testReferenceData(path_to_reference_data = “Tax4Fun2_ReferenceData_v2”)
buildDependencies(path_to_reference_data = “./Tax4Fun2_ReferenceData_v2”,
install_suggested_packages = TRUE, use_force = T)
“”
```

Data preparation.

- OTU sequences or ASV representatives (.fasta)
- Relative abundance of sequences (.txt o .csv)

Reference database: Representative sequences are aligned with the reference sequences with BLAST. We can choose from database Ref99NR or Ref100NR. Information is saved in as .txt file in TEMP folder, following the pathway specified at the beginning.

**# Alignment**

```
“”{r}
#Load OTU representative sequences
fasta <- read.fasta(“SEC_GPE/lechuquilla_rep.fasta”)

#Test with 99NR database
runRefBlast(path_to_otus = “SEC_GPE/lechuquilla_rep.fasta”,
path_to_reference_data = path,
path_to_temp_folder = “SEC_GPE/TEMP”,
database_mode = “Ref99NR”,
use_force = T, num_threads = 2)
```

```
#Check file
ref<- read.delim(file = "SEC_GPE/TEMP/ref_blast.txt")
""
```

After the functional prediction, we get two files:
1. Relative abundance of the predicted functions in txt file
2. Abundance of metabolic pathways per sample (txt file)

**#Functional prediction**

```
""{r}
makeFunctionalPrediction(path_to_otu_table = "SEC_GPE/rep_count.txt",
path_to_reference_data = path,
path_to_temp_folder = "SEC_GPE/TEMP",
database_mode = "Ref99NR",
normalize_by_copy_number = T,
min_identity_to_reference = 0.97,
normalize_pathways = F)
""

#Levels check
""{r}
pathways <- read.delim(file = "SEC_GPE/TEMP/pathway_prediction.txt")
level1 <- unique(pathways$level1)
level2 <- unique(pathways$level2)
level3 <- unique(pathways$level3)

read.delim(file = "SEC_GPE/TEMP/functional_prediction.txt")
""
```

**#Functional redundance prediction**

```
""{r}
calculateFunctionalRedundancy(path_to_otu_table= "SEC_GPE/rep_count.txt",
path_to_reference_data = path,
path_to_temp_folder = "SEC_GPE/TEMP",
database_mode = "Ref99NR",
min_identity_to_reference = 0.97)
""
```

## b) Heatmap for relative abundance of bacteria (pheatmap V 1.0.12)

**## Load packages**

```
""{r }
X = c("pheatmap", "RColorBrewer", "colorspace")
lapply(X, require, character.only = TRUE)

hcl_palettes(plot = TRUE) #paleta de colores
knitr::opts_chunk$set(comment = NA)

rm(list = ls())
""
```

**## Functions**

```
""{r}
braycurtis <- function(mat, rowcol=1) {
mat <- as.matrix(mat)
if(rowcol==1) {
mat.sums <- apply(mat, 1, sum)
n <- dim(mat)[1]
BC <- dist(mat, method="manhattan")
BC <- as.matrix(BC)/(matrix(mat.sums, nrow=n, ncol=n) + matrix(mat.sums,
nrow=n, ncol=n, byrow=T))
 }
if(rowcol==2) {
mat.sums <- apply(mat, 1, sum)
n <- dim(mat)[2]
BC <- dist(t(mat), method="manhattan")
BC <- as.matrix(BC)/(matrix(mat.sums, nrow=n, ncol=n) + matrix(mat.sums,
nrow=n, ncol=n, byrow=T))
 }
as.dist(BC)
}

callback = function(braycurtis, mat){
sv = svd(t(mat))$v[,1]
dend = reorder(as.dendrogram(braycurtis), wts = sv)
as.hclust(dend)
}

#Normalization
function(x){
(x-min(x))/(max(x)-min(x))
}
""
```

**## Genera selection based on the relative abundance (>1%)**

```
""{r}
# Load data base
otus_short<- read.csv("Database.csv",row.names=1)[, c(15:26)]

# Relative Abundance >1%
```

```
BD<-                    otus_short[rowSums(apply(otus_short,2,function(x)
x/sum(x)*100) >=1)>=1,]
# Normalization
BD_corta<- apply(BD,2,function(x) x/sum(x)*100)
""
```

**## Heatmap**

```
""{r }
#Left annotation:
#Variable vector
annotation = data.frame(Type = factor(rep(c("Mature", "Soil", "Young"),
each=4)))
#Annotation vector
rownames(annotation)<-colnames(BD_corta)
#Chose colors
anno_col = list(Type = c(Mature = "green4", Young="blue", Soil="orange4"))

#Function to change font to italics
newnames <- lapply(
colnames(t(BD_corta)),
function(x) bquote(italic(.(x))))

#Colors
tr <- c("aliceblue", "cornflowerblue", "royalblue", "blue", "mediumblue",
"navyblue", "black")
#Save Figure
tiff("heatmap_amoles_00_class01.tiff", units="in", width=13, height=5,
res=200)
pheatmap(t(BD_corta),
#Function for Bray distance clusters
clustering_callback= callback,
#Trees options
treeheight_row = 20, treeheight_col = 30,
cluster_row = T, cluster_col = F,
#Cells
cellwidth = 15, cellheight = 14,
#Column and row names
show_rownames = T, show_colnames = T,
#Leyendas
legend = T,
fontsize = 10,
annotation_legend = T,
annotation_row = annotation,
annotation_colors = anno_col,
#Genera names in italics
labels_col = as.expression(newnames),
#Heatmap legend colors
color = colorRampPalette(colors = tr)(50))
""
```

## c) Non-metric multidimensional scaling and analysis of similarities (vegan V 2.5-7)

**## Load packages**

```
""{r message=FALSE, warning=FALSE}
X= c("ggplot2", "vegan")
lapply(X, require, character.only = TRUE)

library("RColorBrewer")
rm(list = ls())
""
```

**## Load database**

```
""{r}
#Absolute abundance of genera
otus<- read.csv("database.csv",row.names=1)

#Variables database
metadata<- read.csv("VARdatabase.csv", row.names = 1)

#Qualitative variable order
metadata$Type <- factor(metadata$Type, levels=c("Mature", "Young", "Soil"))

""
```

**## Sample selection**

```
""{r}
#Selection based on abundance
good_samples <- colnames(otus[(colSums(decostand(otus,"pa")) >= 1)])
otus = otus[,good_samples]
metadata = metadata[good_samples,]
""
```

**## Samples on row and genera names as columns**

```
""{r}
t_otus <- as.data.frame(t(otus))
""
```

**## Random rarefied community data**

```
""{r}
min_depth = min(colSums(otus))
t_otus_rarefied <- as.data.frame(round(rrarefy(t_otus, min_depth)))
""
```

**## Selection of distance method**

```
""{r}
sqrt_t_otus_rarefied = sqrt(t_otus_rarefied)
rank.totus <- rankindex(as.matrix(sqrt_t_otus_rarefied), t_otus_rarefied,
indices = c("bray", "euclid", "manhattan", "horn"), method = "spearman")
print(paste("The highest rank was given by the", names(sort(rank.totus,
decreasing = TRUE)[1]), "method."))
""
```

**## Distance matrix**

```
""{r}
otus_dist = as.matrix((vegdist(t_otus_rarefied, "bray")))
```

```
""
```

**## NMDS from distance matrix**

```
""{r}
NMDS = metaMDS(otus_dist);NMDS
stressplot(NMDS)
""
```

**## Envfit with environmental variables**

```
""{r }
#Selection of environmental variables
envdata<-metadata[,35:51]

#Envfit with 999 permutations
set.seed(200)
fit <- envfit(NMDS, envdata, perm = 999, na.rm = TRUE)
scores(fit, "vectors")
plot(NMDS, type = "t")
plot(fit)

#Qualitative variable for color selection
grp1 <- metadata$Type

#Colors:
cols <- c('red', 'orange', 'blue')
cols[grp1]

#Save figure
jpeg(filename = "NMDS_envfit.jpeg", units="in", width=8, height=6, res=400,
quality = 100)
#Margin adjust
par(mar=c(5,6,1,1))
plot(NMDS, display = "site", type = "n", cex.axis=1.8, cex.lab = 1.9)
abline(v=0, col="gray", lty=2, lwd=2)
abline(h=0, col="gray", lty=2, lwd=2)
points(NMDS, bg = cols[grp1], cols= "black", cex = 3.5, pch = 21)
legend('topleft', legend = tools::toTitleCase(levels(grp1)),
fill = cols, bty = 'o', cex = 1.5)
plot(fit, p.max = 0.05, col = "black", cex=1.8)

#Envfit data (variables, loadings and P values)
fit_tab <-as.data.frame(fit$vectors$arrows)
fit_tab$r2 <- fit$vectors$r
fit_tab$pvalue <- fit$vectors$pvals

""
```

**## Genera envfit**

```
""{r }
set.seed(4)
amoles.spp.fit <- envfit(NMDS, t_otus_rarefied, permutations = 999)
table_fitamoles <- as.data.frame(amoles.spp.fit$vectors$arrows)
table_fitamoles$r <- amoles.spp.fit$vectors$r
```

```
table_fitamoles$pvalue <- amoles.spp.fit$vectors$pvals

#Write results table
write.csv(table_fitamoles, "amoles.spp.fitSP.csv")

#Save figure
jpeg(filename = "NMDS_envfit_generos.jpeg", units="in", width=8, height=6,
res=400, quality = 100)
par(mar=c(5,6,1,1))
plot(NMDS, display = "site", type = "n", cex.axis=1.8, cex.lab = 1.9)
abline(v=0, col="gray", lty=2, lwd=2)
abline(h=0, col="gray", lty=2, lwd=2)
points(NMDS, bg = cols[grp1], cols= "black", cex = 3.5, pch = 21)
legend('topright', legend = tools::toTitleCase(levels(grp1)),
fill = cols, bty = 'o', cex = 1.5)
plot(amoles.spp.fit, p.max = 0.001, col = "black", cex = 0.8)

""
```

**# ANOSIM**

**## Distance matrix**

```
""{r}
dist1 = as.matrix((vegdist(t_otus_rarefied, "bray")))
set.seed(100)
#ANOSIM based on qualitative variable
anosim_tipo = anosim(dist1, metadata$Type, permutations=999)
#Summary and plot
summary(anosim_tipo)
plot(anosim_tipo)
""
```

**# Paired ANOSIM**

```
""{r}
# Database division by variable levels
MP_YP = t_otus_rarefied[c(1:4, 9:12),]
MP_S = t_otus_rarefied[c(1:8),]
YP_S = t_otus_rarefied[c(5:12),]

# Qualitative variable
a = metadata[c(1:4, 9:12),]
b = metadata[c(1:8),]
c = metadata[c(5:12),]

# Paired ANOSIM
set.seed(100)
#test
anosim(MP_YP, a$Type)
anosim(MP_S, b$Type)
anosim(YP_S, c$Type)
""
```

## d) Principal component analysis (factoextra V 1.0.7 and FactoMineR V 2.4)

**## Load packages**

```
""{r message=FALSE, warning=FALSE}
X("FactoMineR", "factoextra")
lapply(X, require, character.only = TRUE)

knitr::opts_chunk$set(comment = NA)
rm(list = ls())
""
```

**## Load database**

```
""{r}
MPcomp<- read.csv("database.csv", row.names = 1)

#Factor selection
MPcomp<-MP[,c(8:28) ]

#Qualitative variable order
MP$Type <- factor(MP$Type, levels=c("Mature", "Young", "Soil"))
""
```

**#Selection of variables**

```
""{r}
MPcut <- MPcomp[, -c(1, 13, 18, 19, 20)]
MP$Type <- factor(MP$Type, levels = c("Mature", "Young", "Soil"))
names(MPcut)
paste(names(MPcut), collapse= ", ")
""
```

**## PCA**

```
""{r }
PC_mcs<- PCA(MPcut, graph = FALSE)

#Summary
summary(PC_mcs)

#Eigen value
get_eig(PC_mcs)

#Explained variance by each dimension
fviz_screeplot(PC_mcs, addlabels = TRUE, ylim = c(0, 50))
#Variable contribution to each dimension
fviz_contrib(PC_mcs, choice="var", axes = 1:2 )

#Variables
var <- get_pca_var(PC_mcs)
contrib<- var$contrib
contrib <-contrib
con<- data.frame(contrib[, 1:2])

#Save data as new file
```

```
write.csv(con, "contrib_pca.csv")

#Colors for variable levels
jco <- c("#33A02C","#1F78B4", "#B15928")

#Save figure
jpeg(filename = "PCA_figure.jpeg", units="in", width=8, height=6, res=400,
quality = 100)
fviz_pca_biplot(PC_mcs, labelsize = 6, repel = TRUE,
#Individual variables
#Bullet point form
geom.ind = "point",
#Color for level variable
fill.ind = MP$Type, col.ind ="black",
#Size and shape of bullets
pointshape = 21, pointsize = 4,
#Colors
palette = jco,
#Add ellipse
addEllipses = TRUE, ellipse.level=0.50,
# Variables
col.var = "contrib",
#Legends
legend.title = list(fill = "Type", color= "% Contrib")) +
#Color scale for contribution of variables
scale_color_gradient2(low = "gray87", mid = "gray50", high = "gray0",
midpoint = 6)+
#Font size
theme(text = element_text(size = 12),
axis.title = element_text(size = 12),
axis.text = element_text(size = 12))

#Loadings from dimension 1 y 2
dim<-as.data.frame(PC_mcs$ind$coord)[, 1:2]
dim
write.csv(dim, "dims_pca_AA.csv")
""
```

## e) Random forest (randomForest V 4.6-14)

**## Load packages**

```
""{r}
library(randomForest)
rm(list = ls())
""
```

**# Load database**

```
""{r}
BD<- read.csv("database.csv", row.names = 1)
attach(BD)

#Variable selection
var <-c("Type", "CN", "C", "N", "Ca", "Na", "K", "Mg", "Cu", "Fe", "Zn",
"Mn", "OM", "NH4", "NO2_NO3", "P", "pH", "EC", "Clay", "Silt", "Sand")
BD_short <- BD[, var]
BD_short$Type <- as.factor(BD_short$Type)
""
```

**# Model**

```
""{r}
# Train and validation sets (70:30).
set.seed(100)
train <- sample(nrow(BD_short), 0.7*nrow(BD_short), replace = FALSE)
TrainSet <- BD_short[train,]
ValidSet <- BD_short[-train,]

#Independent variables
feat <-c("CN", "N", "Ca", "Na", "K", "Mg", "Cu", "Fe", "Zn", "Mn", "NH4",
"NO2_NO3", "P", "pH", "EC", "Clay", "Silt", "Sand")
#Response variable
resp <- "Type"

#Model
formula <- as.formula(paste(resp, "~", paste(feat, collapse = "+")))
#Test model
model1 <- randomForest(formula, data = TrainSet, importance = TRUE)
model1
#Error rate
model1$mse[[500]]
plot(model1, type = "h")

# Model tuning. Adjust tree number and number of variables to use on each
classification
ntree <- 1000
mtry <- 6
model2 <- randomForest(formula, data = TrainSet, ntree = ntree, mtry= mtry,
importance = TRUE)
model2
#Error rate
model2$mse[[ntree]]
```

```
#Test model
resumen <- as.data.frame(importance(model2)) #MeanDecreaseAccuracy
varImpPlot(model2, type = 1)
abline(v=mean(resumen$`%IncMSE`), col= "red")
""
```

**#Model verification**

```
""{r}
#Prediction on training data set
predTrain <- predict(model2, TrainSet, type = "class")
#Verification
table(predTrain, TrainSet$Type)

#Prediction on validation data set
predValid <- predict(model2, ValidSet, type = "class")
#Verification
mean(predValid == ValidSet$Type)
table(predValid,ValidSet$Type)
""
```

## f) Redundancy analysis (vegan V 2.5-7)

**## Load packages**

```
""{r}
library(vegan)
rm(list=ls())
""
```

**## Load database**

```
""{r}
#Genera
gen <- read.csv("genero.csv", row.names = 1)

#Variables
env<- read.csv("variables.csv", row.names = 1)

#Qualitative variable order
env$Type <- factor(env$Type, levels=c("Mature", "Young", "Soil"))
""
```

**## Hellinger transformation for relative abundances of the genera**

```
""{r}
spe.log <- log1p (gen[, -1])
spe.hell <- decostand (spe.log, 'hell')
""
```

**## Uncosntrained Model**

```
""{r}
modelo0 <- rda(spe.hell); modelo0

# Obtention of model axis
PCA12 <- scores (modelo0, display = 'sites', choices = 1:6)

tbRDA_PCA12 <- rda (spe.hell ~ PCA12)
#Variance explained by the model
RsquareAdj (tbRDA_PCA12)$r.squared*100
""
```

**#RDA model**

```
""{r}
# Model
tbRDA <- rda (spe.hell ~ Leu+Tyr+Arg, data = env);tbRDA

#Variance explained by the new model
R2.obs <- RsquareAdj (tbRDA)$r.squared; R2.obs*100
""
```

**## Monte Carlo test**

```
""{r}
# Model variables
env_short <- env[, c("Leu", "Tyr", "Arg")]

#Permutations
set.seed(200)
n.perm <- 999
```

```
# Function to perform permutations
R2.rand <- replicate (n = n.perm, expr = {
env.rand <- env_short[sample (1:12),]
#Model
tbRDA.rand <- rda (spe.hell ~ Leu+Tyr+Arg, data = env.rand)
RsquareAdj (tbRDA.rand)$r.squared
 })

#1000 permutations in vector R2
R2 <- c (R2.rand, R2.obs)

#Plot
hist (R2, nclass = 1000, xlim = c(0,1))
#Red line depicts observed variance against permutations values
abline (v = R2.obs, col = 'red')
range (R2.rand)

#Monte Carlo test
P <- sum (R2 >= R2.obs)/(n.perm + 1);P

#ANOVA to check Monte Carlo test
set.seed(200)
anova (tbRDA, permutations = 999)
""
```

**## Explained variance by model (red) and not explained (black)**

```
""{r fig.height=4, fig.width=6}
#Variation as percentage
constrained_eig <- tbRDA$CCA$eig/tbRDA$tot.chi*100
unconstrained_eig <- tbRDA$CA$eig/tbRDA$tot.chi*100
expl_var <- c(constrained_eig, unconstrained_eig)

#Plot
barplot (expl_var[1:20], ylim=c(0,100), #ylim ajusta límites del eje
col = c(rep ('red', length (constrained_eig)),
rep ('black', length (unconstrained_eig))),
las = 2, ylab = '% variation')
""
```

**## RDA plot**

```
""{r}
#Simple plot
ordiplot(tbRDA, choices = c(1,2))

#RDA axis summary
R.sum <- summary(tbRDA)

#Explained variance by each axis
R.sum$cont

#AS percentage
```

```
RDA1 <- round(R.sum$cont$importance[2, "RDA1"]*100, 2)
RDA2 <- round(R.sum$cont$importance[2, "RDA2"]*100, 2)

#Enfit variables
ef <- envfit (tbRDA, env_short, choices = c(1,2), permutations = 999)
#Colors
cols <- adjustcolor(cols, alpha.f = 0.9)
grp1 <- env$Type

#Save figure
jpeg(file="RDA_nmds_AA_text.tiff", units="in", width=8, height=6, quality
= 100, res = 400)
#Plot margins
par(mar=c(5,5,1.5,1))
#Final plot
ordiplot (tbRDA, choices = c(1,2), type = 'n',
xlab= paste0("RDA1 (", RDA1, "%)"),
ylab= paste0("RDA2 (", RDA2, "%)"))
#Bullet points
points (tbRDA, choices = c(1,2), display = 'sites', bg = cols[grp1], cex =
2.1, pch = 21)
#Envfit (vectors)
plot(ef, cex=1)
#Legend
legend('topright', legend = tools::toTitleCase(levels(grp1)),
fill = cols, bty = 'o', cex = 1)
""
```

**## Test with more variables**

```
""{r}
#RDA plot
ordiplot (tbRDA, choices = c(1,2), type = 'n')
points (tbRDA, choices = c(1,2), display = 'sites', pch = as.character
(env$Type), col = env$Type)
#Data base with new variables
BD<- env[, 35:51]
set.seed(200)
ef <- envfit (tbRDA, BD, choices = c(1,2), permutations = 999)

#Plot with envift of new variables
plot (ef, p.max= 0.05)

#R and P values
ef1 <- as.data.frame(ef$vectors$arrows)
ef1$r2 <- ef$vectors$r
ef1$pvals <- ef$vectors$pvals
""
```

## g) Canonical correspondence analysis (vegan V 2.5-7)

**## Load packages**

```
""{r message=FALSE, warning=FALSE}
library(vegan)
rm(list=ls())
""
```

**## Load database**

```
""{r}
BD <- read.csv(("enzimas_transformado_prom03_junio21.csv"), row.names =
2)[-71,]
BD <- BD[, -c(1,5,7:9)]
# Change into factors the variables for further use:
BD$Concentration <- as.factor(BD$Concentration)
BD$Metabolite_Type <- as.factor(BD$Metabolite_Type)

# Desired order of the variables to use:
BD$Concentration <- factor(BD$Concentration,
levels = c("None", "Low", "Medium", "High"))
BD$Metabolite <- factor(BD$Metabolite,
levels=c("Control",
"Oxalic", "Tartaric",
"Lactose", "Maltose",
"Arginine", "Methionine"))
""
```

**## Selection and rename of the variables**

```
""{r}
#Two data set with the variables selected for the model
ACTABU <- BD[, c(5,6, 11:13)]
NUT <- BD[,c(7,8, 14,17)]

#Rename variables columns
colnames(ACTABU) <- c('Nitrogenase', 'log nifH', 'APM', 'APD', 'log phoD')
colnames(NUT) <- c('NH4', 'NO2+NO3', 'DIP', 'DON:DOP')
""
```

**## CCA Model**

```
""{r }
# Formula of the model (Y~X, data=database)
vare.cca <- cca(NUT ~ Nitrogenase+log.nifH+APM+APD+log.phoD, data=BD)

# Complete results
vare.cca

# Summary of the results
summary(vare.cca)
""
```

**## Validation of the model**

```
""{r }
# Significance of the model: Permutation test (permutations 999)
anova (vare.cca)

# Significance by axis: Permutation test (permutations 999)
anova (vare.cca, by = 'axis')
""
```

**## Variance/Inertia explained by the model**

```
""{r}
sum <- summary(vare.cca)
#Total inertia of the model
vare.cca$tot.chi

#Inertia explained by constrained variables
sum$constr.chi

#Inertia explained by unconstrained variables
sum$unconst.chi

#Proportion of inertia explained by the constrained variables:
paste0(round((sum$constr.chi*100)/vare.cca$tot.chi,  2),  "%  of  inertia
explained by the constrained variables")
#Variance explained by components
#CCA1
round(sum$cont$importance[2,1]*100, 2)

#CCA2
round(sum$cont$importance[2,2]*100, 2)

""
```

**## Base plot**

```
""{r fig.height=7, fig.width=9}
plot(vare.cca, display = c("sites","species"),
type="t", scaling = "symmetric")
""
```

**## Edited plot to save in high definition**

```
""{r fig.height=7, fig.width=9}
#Parameters for colors and forms
tipos <- BD$Metabolite
conx <- BD$Concentration
pch <- c(3,21,24,22)
cols <- c("black",
"dodgerblue", "blue4",
"darkolivegreen1", "forestgreen",
"darkgoldenrod1", "chocolate4")

# tiff("CCA.tiff", units="in", width=8, height=7, res=300)
#Margins
```

```
par(mar = c(5,4,2,2))
#Base plot without points or title axis
plot(vare.cca, type="n", scaling = "symmetric",
xlab="",ylab="")
#Title axis
mtext(side=1, line=3,
paste('CCA1: ', round(sum$cont$importance[2,1]*100, 2), '%'),
col="black", cex=1.5)
mtext(side=2, line=2,
paste('CCA2: ', round(sum$cont$importance[2,2]*100, 2), '%'),
col="black", cex=1.5)
#Points
points(vare.cca, scaling = "symmetric",
bg = cols[tipos],
col= "black",
pch = pch[conx],
cex=1.5)
#Vectors
text(vare.cca, display="bp", scaling = "symmetric", col= "blue", cex=1.2)
#Species or response variables
text(vare.cca, display = "species", col= "red", scaling = "symmetric",
cex=1.2)
#Legends
with(BD, legend("bottomleft", cex=1.2,
legend = levels(Metabolite), bty = "n",
col = "black", pch = 21, pt.bg = cols))
with(BD, legend("bottomright", cex=1.2,
legend = levels(Concentration), bty = "n",
col = "black", pch = c(3,21,24,22), pt.bg = "black"))

""
```

## h) Sparse partial least square discriminant analysis (mixOmics V 6.16.1)

**# Load packages**

```
""{r }
X= c("mixOmics", "vegan")
lapply(X, require, character.only = TRUE)

rm(list=ls())
""
```

**# Load database**

```
""{r}
#Absolute abundance
otus<- read.cs("database.csv",header = TRUE, row.names = 1)
""
```

**# Genera selection**

```
""{r}
#Above 1% of relative abundance and with two or more counts per sample
BD<- otus[rowSums(apply(otus,2,function(x) x/sum(x)*100) >=1)>=2,]
head(BD)
#Normalization
BD_corta<- apply(BD,2,function(x) x/sum(x)*100)

#Write new file with selected taxa
write.csv(x = BD_corta, file = "ABREL_generos1_Redes.csv")
""
```

# Load new file

```
""{r}
# New database
BD_taxas<- read.csv("newdatabase.csv", row.names = 1)

# If needed transposition. Samples as row names and taxa as columns names
head(BD_taxas)
BD_taxas <- t(BD_taxas)
# Categoric variables
metadata<- read.csv("VAR_data.csv", header = TRUE, check.names = FALSE,
row.names = 1)
metadata<-metadata[-22, 1:28]
BD_var <- metadata[, c(7, 4)]

# Merge databases
BDCom <- merge(BD_var, BD_taxas, by= "row.names")
rownames(BDCom) <- BDCom[,1] #Rename rownames
BDCom[,1] <- NULL #delate new columne
head(BDCom) #Check

# Qualitative variable order
BDCom$Type <- factor(BDCom$Type,levels = c("Mature", "Young", "Soil"))
""
```

**# sPLS-DA by type**

```
""{r}
```

```
#Vector with type levels
BD_Lugar <- BDCom[, -1]
""
```

**## Test model**

```
""{r}
#Y variable (Response) and X variables (explication variables):
BD_Lugar <- BD_Lugar[c(1:12, 37:47), ]#Central
BD_Lugar <- BD_Lugar[-c(1:12, 37:47), ]#Meseta

spe.log <- log1p (BD_Lugar[, -1]) #Taxa only
x <- decostand (spe.log, 'hell') #Hellinger transformation for relative
abudances

x <- BD_Lugar[, -1] #TAXA
y <- BD_Lugar$Type

#Colors
colors <- c('red', 'orange', 'blue')

#Default values for first model
MyResult.splsda <- splsda(x, y, keepX = c(50,50), ncomp = 2, scale = TRUE)
#Plot
plotIndiv(MyResult.splsda, ind.names = FALSE, legend=TRUE,
ellipse = TRUE, star = TRUE, title = 'sPLS-DA on type',
X.label = 'PLS-DA 1', Y.label = 'PLS-DA 2',
col = colors)
#Significant taxa cutoff =>0.7
plotVar(MyResult.splsda, var.names = TRUE, cutoff = 0.7)

#Performance plot
auc.plsda <- auroc(MyResult.splsda)
# Values correspond to Wilcoxon test

#Significant variables by component
selectVar(MyResult.splsda, comp=1)$value
plotLoadings(MyResult.splsda, contrib = 'max', method = 'mean', comp = 1)
""
```

**### New model**

```
""{r}
#Model
MyResult.plsda2 <- plsda(x,y, ncomp=5)

#Test number of components
set.seed(30)
MyPerf.plsda <- perf(MyResult.plsda2, validation = "Mfold", folds = 3,
progressBar = TRUE, nrepeat = 50)
#Plot
plot(MyPerf.plsda, col = color.mixo(5:7), sd = TRUE, legend.position =
"horizontal")
#Error rate
```

```
MyPerf.plsda
MyPerf.plsda$error.rate
""
```

**### Model tuning**

```
""{r}
# Vector to test number of classification variables. As big as the number
of variables.
list.keepX <- c(1:10, seq(15, 200, 20))
list.keepX

set.seed(30)
tune.splsda.srbct <- tune.splsda(x, y, ncomp = 4,
validation = 'Mfold',
folds = 3, dist = 'max.dist', progressBar = T,
measure = "BER", test.keepX = list.keepX,
nrepeat = 10)  #
# Error rate
error <- tune.splsda.srbct$error.rate

ncomp <- tune.splsda.srbct$choice.ncomp$ncomp; ncomp
# Suggested number of components

select.keepX <- tune.splsda.srbct$choice.keepX[1:ncomp]; select.keepX
#Suggested number of taxa for classification
""
```

**### Final model**

```
""{r , test-rgl, webgl=TRUE}
#Tunned model
MyResult.splsda.final <- splsda(x, y, ncomp = ncomp, keepX = select.keepX)

tiff("PLSDA_redes_13.tiff", units="in", width=7, height=4.5, res=300)
plotIndiv(MyResult.splsda.final, ind.names = F, pch = c(15,16,17,18),
legend=TRUE, legend.position= "bottom", legend.title = "",
ellipse = TRUE, title = 'sPLS-DA',
X.label           =           paste("Component           1           (",
round(MyResult.splsda.final$prop_expl_var$X[[1]], 3)*100, "%)"),
Y.label           =           paste("Component           2           (",
round(MyResult.splsda.final$prop_expl_var$X[[2]], 3)*100, "%)"),
Z.label           =           paste("Component           3           (",
round(MyResult.splsda.final$prop_expl_var$X[[3]], 3)*100, "%)"),
size.xlabel = rel(1.4),
size.ylabel = rel(1.4),
size.zlabel = rel(1.4),
size.axis = rel(1.2),
ellipse.level = 0.95,
col = colors)

tiff("Type_PLSDA_redes_VAR.tiff", units="in", width=8, height=5, res=300)
plotVar(MyResult.splsda.final, var.names = TRUE, cutoff = 0.7)  #variables
(*)
```

```
auc.plsda <- auroc(MyResult.splsda.final) #Curva ROC

#Performance results
vars <- selectVar(MyResult.splsda.final, comp=1)$value
# Change component to review results

# Variables plot
tiff("Types_plsda_redes_comp3.tiff",   units="in",   width=6,   height=15,
res=300)
plotLoadings(MyResult.splsda.final, contrib = 'max', method = 'mean', comp
= 3,
size.title = rel(1), size.name = 0.6,
xlim = c(-0.4, 0.4), border = TRUE)

#Loadings
loadings <-MyResult.splsda.final$loadings$X
write.csv(loadings, "Type_loadings_redes_X.csv")

vips <- vip(MyResult.splsda.final)
write.csv(vCips, "Type_vips_redes.csv")

#Performance of tunned model
MyPerf.plsda <- perf(MyResult.splsda.final, validation = "Mfold", folds =
3,
progressBar = TRUE, nrepeat = 50)
plot(MyPerf.plsda)
MyPerf.plsda$error.rate #Error rate

#Cross Validation (Classification Error Rate)
MVA.test(x, y, model = "PLS-DA", ncomp = ncomp, K=47,
kout = 7, kinn = 6, Q2diff = 0.05,
p.method = "fdr", nperm = 999)
""
```