Cinvestav
Unidad Irapuato

Seed-transmitted bacteria and their contribution to the cacti holobiont

Tesis presentada por:

QBP Víctor Manuel Flores Núñez

Para obtener el título de Maestro en Ciencias en Biología Integrativa

Directora de tesis:

Dra. Laila Pamela Partida Martínez

Irapuato, Gto. Febrero 2017

Este trabajo se realizó en el Centro de Investigación y de Estudios Avanzados del IPN - Unidad Irapuato, Departamento de Ingeniería Genética, Laboratorio de Interacciones Microbianas bajo la dirección de la Dra. Laila Pamela Partida Martínez gracias al apoyo del Consejo Nacional de Ciencia y Tecnología (CONACyT) a través de los Proyectos CB-2010-01-1510007 y INFR-2012-01-187799, además del convenio con el US Department of Energy Joint Genome Institute a través del JGI Community Science Program bajo el número de contrato DE-AC02-05CH11231.

Este trabajo fue realizado del 31 de Agosto de 2015 al 20 Enero de 2017.

Miembros del Comité de Asesores:
Dra. June Kilpatrick Simpson Williamson
Dr. Alfredo Heriberto Herrera Estrella

Fecha de examen de grado: 10 – Marzo - 2017

A la voz en mi cabeza que insiste en repetir que
no soy suficientemente bueno
Lo soy

AGRADECIMIENTOS

A Laila Partida, quien confió en mis capacidades para realizar este proyecto, por su continuo apoyo y motivación que fueron fundamentales para concluirlo con éxito. Sin duda, es un ejemplo de que se puede hacer ciencia de calidad cuando es prioridad el desarrollo académico y personal de los alumnos.

A los integrativos Gaby, Pavel, Arely, Edder, Gil y Daniel, cuya amistad hizo que valiera la pena cada minuto de está maestría y de quienes he aprendido grandes lecciones de vida sobre voluntad, convicción, libertad, valor, gentileza y perseverancia.

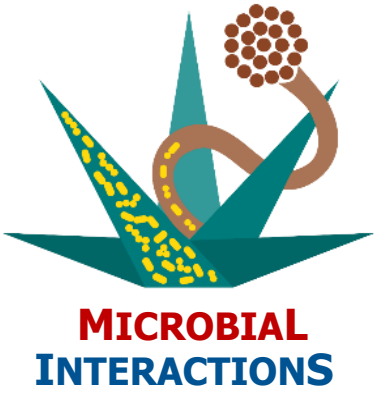
A todos los miembros del Laboratorio de Interacciones Microbianas, Lety, David, Astrid, quienes hemos formado un equipo asombroso; en especial a Neli quien hace una labor excepcional en el laboratorio.

A June Simpson y Alfredo Herrera por participar proactivamente en la evaluación y por sus aportaciones a esté trabajo.

A mi papás por su apoyo incondicional a pesar de la dificultad de estar separados.

A Citlalli Fonseca por sentar el precedente para desarrollar este trabajo, y a Beto y Gabriel Moreno por su apoyo para realizar los análisis bioinformáticos.

A CONACYT por la beca de maestría de Marzo 2015 a Febrero 2017



RESUMEN

Los simbiontes microbianos afectan la supervivencia, desarrollo, adecuación y evolución de sus hospederos. Un holobionte es una unidad biológica formada por el conjunto de estos microorganismos más el hospedero. Estudios recientes muestran que el holobionte de los cactus está formado por un abundante y diverso microbioma, el cual podría ser importante para su adaptación a ambientes áridos. Para revelar la capacidad funcional del microbioma de los cactus nos enfocamos en los endófitos bacterianos de semillas, que son transmitidos verticalmente a través de éstas y podrían contribuir a la adecuación de los cactus. Para evaluar la composición de endófitos transmitidos en semillas de cactus (subfamilias *Opuntioideae* y *Cactoid*eae) se utilizaron técnicas dependientes e independientes de cultivo. Se utilizó *Arabidopsis thaliana* para evaluar el efecto de las bacterias transmitidas por semilla en la adecuación de las plantas bajo estrés hídrico. Finalmente, se analizaron los genomas de estas bacterias para identificar funciones enriquecidas relacionadas con la promoción del crecimiento vegetal y la tolerancia a estrés. Las bacterias transmitidas por semillas de cactus pertenecen a los géneros *Bacillus*, *Paenibacillus, Psychrobacillus*, *Agrococcus*, *Nocardiopsis*, *Staphylococcus* y *Leclercia,* siendo las bacterias formadoras de esporas las más abundantes. La región V4 del gen 16S rRNA de *Bacillus* sp*., Staphylococcus hominis* and *Leclercia* sp. coincidió con OTUs abundantes y frecuentes de la endósfera de cactus maduros, sugiriendo que son importantes para el holobionte. Los OTUs de otras cepas estuvieron presentes en otros compartimentos, pero en baja frecuencia y abundancia. La mayoría de las cepas incrementaron el número de raíces laterales de *A. thaliana* en condiciones estándar de cultivo, mientras que otras como *S. hominis*, *Paenibacillus* sp. y *Nocardiopsis prasina* incrementaron la tasa de germinación en bajo potencial de agua, sugiriendo que podrían tener un papel durante la germinación de los cactus en ambientes áridos. El genoma de *S. hominis* no mostró ningún signo de reducción genómica cómo otros simbiontes de trasmisión vertical, pero reveló genes relacionados con la promoción de crecimiento vegetal y tolerancia a estrés. Esta cepa muestra un enriquecimiento en transposasas y presentó otros genes que podrían ser reflejo de la trasmisión transgeneracional y la adaptación a vivir dentro de la planta, respectivamente. En conjunto, estos hallazgos sugieren que los endófitos bacterianos trasmitidos por semillas de cactus podrían contribuir a la germinación y establecimiento de los cactus en ambientes áridos y su potencial para ser usadas como biofertilizantes en la agricultura en sistemas áridos.

ABSTRACT

Microbial symbionts account for survival, development, fitness and evolution of most eukaryotic hosts. These microorganisms together with their host form a biological unit known as holobiont. Recent studies had revealed that the cacti holobiont comprises a diverse and abundant microbiome, which might be important for its adaptation to arid ecosystems. To dissect the functional capabilities of the cacti microbiome, we focused on seed-borne bacterial endophytes that are vertically transmitted through seeds and might contribute to cacti fitness Our strategy included culture-dependent and independent techniques to evaluate the composition of seed-borne bacterial endophytes in cacti (subfamilies *Opuntioideae* and *Cactoid*eae). *Arabidopsis thaliana* was used to assess the impact of seed-borne strains on plant fitness under drought stress. Then, we made use of genomic data from isolated bacteria to identify enriched functions related to plant growth promotion and stress tolerance. Our results showed that cultivable cacti seed-borne bacteria are represented by members of *Bacillus*, *Paenibacillus, Psychrobacillus*, *Agrococcus*, *Nocardiopsis*, *Staphylococcus* and *Leclercia,* being spore-forming bacteria the most abundant strains. Seed-borne *Bacillus* sp*., Staphylococcus hominis* and *Leclercia* sp. strains matched with abundant OTUs from the endosphere of mature cacti, suggesting their importance for the cacti holobiont. Other strains were present in other compartments, but in low abundance and frequency. Most of the strains increased the number of lateral roots of *A. thaliana* under standard conditions while others like *S. hominis*, *Paenibacillus* sp. and *Nocardiopsis prasina* increased germination rate under low water potential suggesting that they might have a role during germination of cacti seeds in drylands. The genome of seed-borne *S. homini* did not show any signs of genome reduction common in other vertically transmitted symbionts, but did revealed genes related to plant growth promotion and stress tolerance. This strain was enriched in unique transposases and other genes that may reflect the transgenerational transmission and the adaptation to plant lifestyle. Altogether, these findings indicate that cacti seed-borne strains might contribute to the germination and establishment of cacti in arid environments and they might have a potential use as biofertilizers for desert farming.

CONTENTS

I. INTRODUCTION

THE HOLOGENOME THEORY OF EVOLUTION

Symbiosis concept was first defined by Anton de Bary in 1978 as “the living together of unlike organism”, implying the association of two or more species regardless of the outcome of the symbiosis (Garcia & Gerardo 2014). Symbiosis is an evolutionary force within all living organisms that has played a key role through the history of life on Earth, from microbial associations during the origin of life to the elaborate microbiota of plants and animals (Guerrero & Berlanga 2015). Nowadays, it’s almost impossible to picture an organism living in complete isolation from others, organisms are no longer considered as autonomous entities, but rather as holobionts, biological units composed by the host plus its microbial symbionts that play a fundamental role in health, development, fitness and evolution of their host (Rosenberg & Zilber-Rosenberg 2013; Bordenstein & Theis 2015).

The hologenome concept was introduced by Jefferson (1994) and Rosenberg et al. (2007), it’s defined as the genomic content of a holobiont including the host, organelles and microbial symbionts genomes (Figure I1) (Rosenberg & Zilber-Rosenberg 2013; Bordenstein & Theis 2015). The hologenome theory of evolution originally implied that the holobiont was the unit of selection in evolution and it considered four principles (Rosenberg & Zilber-Rosenberg 2013): (1) Plants and animals harbor abundant and diverse microbial communities, (2) the microbial symbionts and their genomes together with the host genome can be transmitted to the next generation of holobionts, (3) the interaction of the host with its microbial symbionts affects the physiology, health and fitness of the holobiont and (4) genetic variation of the hologenome occurs by changes either in the host genomes or in the microbial populations. Recently, Bordenstein & Theis (2015) have expanded the theoretical framework of this theory, they suggest that the variation in the hologenome leads to variation in the holobiont’s phenotypes upon which natural selection and drift could operate in agreement with the modern synthesis of evolution.

Variation in the hologenome can occur either in the host or microbial symbionts genomes by mutation, chromosome rearrangements, sexual reproduction and recombination. In addition, hologenome variation can arise by tree additional processes (Rosenberg & Zilber-Rosenberg 2013): (1) by microbial amplification or an increase in the number of certain microbial symbionts, (2) by adding novel strains from the environment that might have novel

functions for the holobiont and (3) by horizontal gene transfer between microorganisms of the same or different species, and with the host.

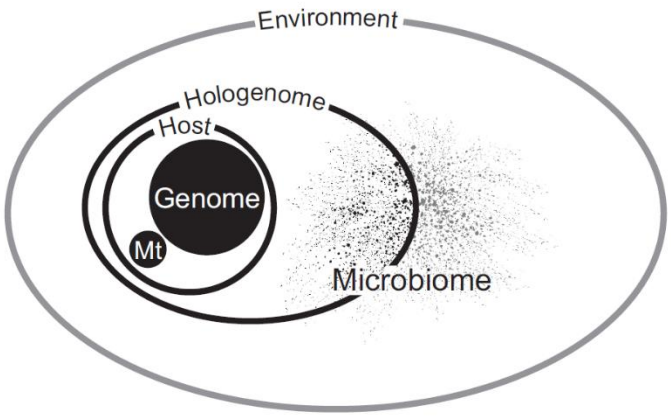


Figure I1. The hologenome. The hologenome is comprised by the host genome (nuclear and organelles) plus the genome of all its symbiotic microorganisms (black dots). Most of the host-associated microbiome derives from the environment and it's limited by the host. (Brucker & Bordenstein 2013)

Hologenomic variation must be inherited to holobiont's offspring to account for its evolution. Symbiotic microorganisms can be transmitted either vertically or horizontally. Symbionts are vertically transmitted when they are directly transfer from the progenitor to their progeny for example: transmammary, transgametic, transplacental, intrauterine transmission and others involving social behaviors (Ebert 2013). During horizontal transmission, microbial symbionts are acquired from the environment (free-living populations) like soil and marine water (Bright & Bulgheresi 2010), it can occur via vectors, sexual transmission and during delivery. The mode of transmission impacts on the dispersion, persistence and genome features of the symbionts, as well the possibility of co-evolving with their hosts (Ebert 2013).

VERTICALLY TRANSMITTED MICROORGANISMS IN PLANTS

Plants harbor a great diversity and abundance of symbiotic microorganisms (bacteria, fungi, virus, protists), both inside (endosphere) and outside (episphere) their tissues (Vandenkoornhuyse et al. 2015; Hacquard 2016). Plant phenotype is influenced by the co-expression and co-regulation of plant and microbial genes (Partida-Martínez & Heil 2011). It is unquestionable that microorganisms (fungi and bacteria) have a beneficial influence on plants by promoting their growth or inducing tolerance against environmental stresses (pathogens, salinity etc.). The mechanisms involved have been extensively reviewed by

several authors (Glick & Glick 2012; de Souza et al. 2015; Hakeem et al. 2016), being the most studied: nitrogen fixation, phosphate solubilization, siderophore production, the modulation of phytohormone levels (ACC deaminase, indole acetic acid, cytokines and gibberellins), competition, hydrolytic enzymes production, induction of systemic resistance among others.

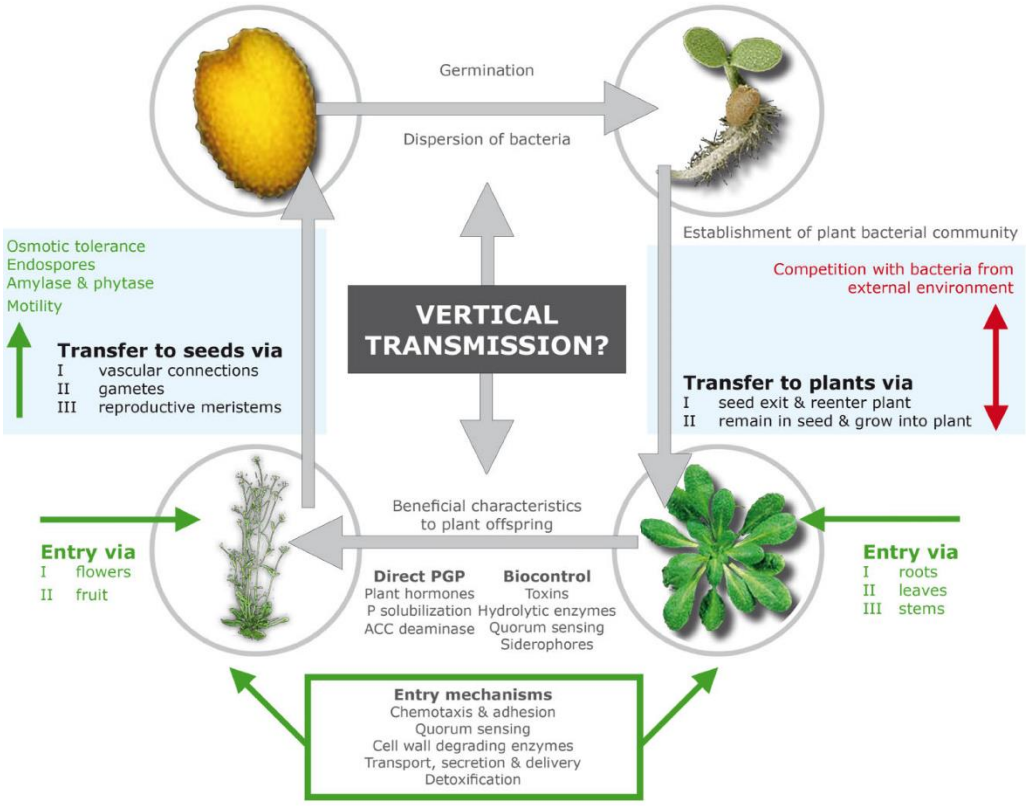


Figure I2. Vertical and horizontal transmission of bacterial seed endophytes. After germination environmental bacteria colonize the plant's endosphere. Endophytes can be transfer to the seeds, this might ensure the dispersal of the endophytes and the presence of beneficial microorganisms to the offspring (Truyens et al. 2015)

Seeds are structures derived from sexual reproduction of spermatophytes (plants with seed) that under certain condition can germinate and generate a new plant individual (Barret et al. 2015). Vertical transmission of plant-associated microorganisms consists in the transfer of the microbial symbionts through the seeds. The occurrence of this phenomenon has been demonstrated in several spermatophytes especially in angiosperms (plants with flowers). This mode of transmission might not be restricted to seeds, *Sphagnum* sp. mosses vertically transmit *Burkholderia* spp. populations (Bragina et al. 2013). Also vegetative reproduction of plants (e. g. cacti stems can give birth to a new plant, when falling into the soil) might be

included in this category since it ensures a direct transmission of the microbiome of the original plant (Zilber-Rosenberg & Rosenberg 2008).

As reviewed by Truyens et al. (2015), bacterial symbionts can be transferred to the seeds via vascular connections, gametes or reproductive meristems and they might have some traits like motility, endospore production etc. to be able to colonize the seeds. When the seeds germinate, this community is the first to colonize the plant tissues and compete with environmental microorganisms (Figure I2). This process might ensure the transmission of a subset of bacteria that could be beneficial during the first stages of the development of plants and also the dispersal of bacteria. The location of seed-transmitted bacteria has been demonstrated in seeds for example in the seed cortex of *Picea abies* and *Pachycereus pringlei* seeds (Puente et al. 2009b; Cankar et al. 2005) and in the endospermic and embryonic tissues of rice seeds (Mukhopadhyay et al. 1996).

Cultivable seed-transmitted bacteria comprise common plant-associated phyla as *Proteobacteria*, *Firmicutes*, *Actinobacteria* and *Bacteroidetes*, the most isolated genera inhabiting the seeds are *Bacillus* and *Pseudomonas*, *Paenibacillus*, *Micrococcus*, *Staphylococcus*, *Pantoea* and *Acinetobacter* that are also found as endophytes (Truyens et al. 2015). Seed-transmitted fungi have been reported less frequently, the canonical examples are *Epichloë* spp. obligate endophytes of cold grasses that are horizontally and vertically transmitted and have co-evolved with their host (Saikkonen et al. 2016). Recent work has focused on the whole microbial community using amplicon sequencing. Using this approach, the presence of non-cultivable bacterial OTUs of the phyla *Verrucomicrobia* and *Acidobacteria* and several fungal OTUs from *Ascomycota* and *Basidiomycota* phyla such as *Fusarium* and *Cryptococcus* genera in *Phaseolus vulgaris* seeds has been shown (Klaedtke et al. 2015).

CACTI: OUR STUDY MODEL

Cacti are plants native to the New World where they originated 32.11 million years ago, they are present in most biomes, but are major components of arid and semiarid environments (Nobel 1988; Hernández-Hernández et al. 2014). Cacti belong to the monophyletic *Cactaceae* family, it's a core member of the *Caryophyllales* order and comprises approximately 1500-1800 species. Most cacti possess perennial photosynthetic succulent stems, leaf spines in aureoles (modified axillary buds), but no green leaves; flowers are

colorful, with separated perianth parts, many stamens and an inferior ovary with numerous ovules; although several basal cacti might have primitive or highly reduced features (Nobel 2002).

Two subfamilies are considered as core members of *Cactaceae* family : *Cactoideae* and *Opuntioideae* (Majure et al. 2012). *Opuntioideae* have 2 characteristic synapomorphies: short deciduous, barbed spines called glochids and a bony aril surrounding a campylotropous ovule, the subfamily includes 15% of the cacti species (approximately 250 species). *Cactoideae* subfamily is characterized by the presence of tubercles or ribs in the stems and the lack of glochids and leaves, it's the most diverse subfamily and comprises more than 80% of the cacti species (Nobel 2002; Nyffeler 2002).

In the present work, we sampled 3 genera of cacti: *Opuntia, Cylindropuntia* and *Myrtillocactus* (Figure I3) that are widely distributed in Mexico. *Opuntia* spp. are characterized by their flat photosynthetic stem segments (cladodes) (Majure et al. 2012) while *Cylindropuntia* spp. have terete-stem segments (Griffith & Porter 2009), both genera can have tree-like, shrub or creeping lifestyles and their fruits have areoles with glochids or spines. *Myrtillocactus* genus have tree-like or shrub lifestyles with a well-defined trunk, stems are highly branched like chandeliers with 4-8 ribs and rigid spines (Rzedowski & Rzedowski 2005).

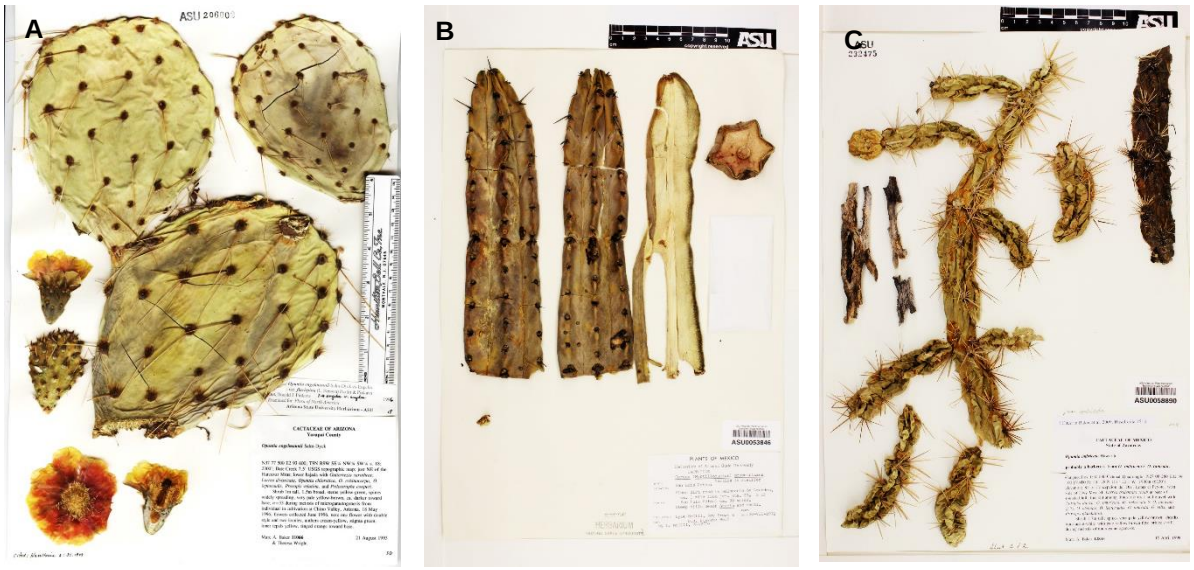


Figure I3. Cacti species of this work. A. *Opuntia engelmannii*. B. *Myrtillocactus geometrizans*. C. *Cylindropuntia imbricata*. Images: Arizona State University Vascular Plant Herbarium.

Due to the fact that cacti are native to the New World, they were very important for the establishment of prehispanic civilization in north and center of Mexico, together with other crops like maize, common bean and agave (Anaya-Pérez & Bautista-Zane 2008), also *Opuntia* genus is recognized in the national flag and in the beautiful landscapes of Mexico. The stems and fruits of many cacti are eaten raw, cocked or preserved since prehispanic civilizations. In Mexico, the consumption of the young cladodes of *Opuntia ficus–indica* and its hybrids (collectively named nopales) and their fruits (tunas and xoconostles) is very common, but many other fruits are known to be consumed by indigenous populations, such as of *Carnegiega gigantea* (saguaro), *Stenocereus gummosus* (Nobel 1988) and *M. geometrizans*.

The uses of cacti are unlimited, Seri populations on the north of Mexico use woody tissues of some columnar cacti like *C. gigantea* and *Lophocereus schotii* (senita) for food storage, construction and toys, several other cacti species have also been used as supplement for human health and nutrition (Shetty et al. 2012). Nowadays *Opuntia* spp. have many agro-industrial uses including: the pharmaceutical and food supplements industry (mucilage, fiber, etc), the cosmetic industry (production of shampoos, creams etc.), the natural additives industry (colorants), the energy sector (production of biogas and biofuels), and as forage for animal feeding, among others (Sáenz et al. 2013).

II. BACKGROUND

THE CACTI MICROBIOME

Most of the studies have focused on the cultivable microbiota of cacti, but none of them attempted to make a comprehensive isolation from all the plant compartments and different cacti species. Several studies reported the isolation of various microorganisms with plant growth promoting traits such as *Bacillus* spp., *Citrobacter* spp., and *Actinomadura oligospora* in the rhizoplane of *Pachycereus pringlei* and *Opuntia chola* (Puente, Bashan, et al. 2004); *Bacillus* spp., *Klebsiella* spp., *Staphylococcus* spp., and *Pseudomonas* spp. from the seeds, pulp of the fruits, seedlings and small, mature plants of *P. pringlei* from volcanic areas of Baja California Sur, Mexico (Puente et al. 2009b)*; Azotobacter vinelandii, Pseudomonas putida, Enterobacter sakazakii, and Bacillus megaterium* in the root and shoot endosphere of *Mammillaria fraileana* from Sonoran Desert (Lopez et al. 2011)*; Enterobacter* spp., *Burkholderia* spp., *Pseudomonas* sp., *Pantoea* sp. and *Rhizobium* sp. in the endosphere of *Cereus jamacaro* and *Melocactus zehntneri* from Brazilian semi-arid regions (Lima et al. 2015)*; Ochrobactrum* spp., *Bacillus* spp., *Arthrobacter* spp. among others from the rhizosphere of *Mammillaria carnea*, *Opuntia pilifera* and *Stenocereus stellatus* from semi-arid highlands in central Mexico (Aguirre-Garrido et al. 2012); and *Bacillus* spp. *S. hominis*, *Leclercia* sp. *Psychrobacillus* spp, among others in seeds and seedlings of *Opuntia robusta* and *Mirtyllocactus geometrizans* from Guanajuato, Mexico (Fonseca-García et al. 2016).

Culture approaches only account only for 1-10% of the total microbial diversity (Nannipieri et al. 2003), therefore high-throughput methods such as amplicon sequencing of different markers are suitable to study abundant and diverse microbial communities. Aguirre-Garrido et al. (2012) studied the rhizospheric bacterial communities of tree different cacti species on rainy and dry seasons using the low-resolution DGGE-microbial profiling of 16S genes. These authors found that each plant species harbored a different microbial community and highly influenced by seasonality. Contrary Fonseca-García et al. (2016) made an holistic study of the composition and diversity of the prokaryotic and fungal communities using amplicon sequencing of 16S genes and ITS region, respectively, in sympatric *O. robusta* and *M. geometrizans*. They found that cacti species shared most of their OTUs, and the composition of their microbial communities was influenced by the plant compartment (rhizosphere, phyllosphere, endosphere), while seasonality and geography had a minor influence.

GROWTH PROMOTING BACTERIA OF CACTI

Traditionally, plant growth promotion potential of microorganisms has been demonstrated using biochemical tests, the most studied traits are the production of indole acetic acid (IAA), solubilization of inorganic phosphorous, nitrogen fixation, ACC deaminase and siderophore production. Interestingly, cacti-associated microorganisms have shown many of these capabilities. Such is the case of bacteria isolated from rhizoplane and seeds of *P. pringlei* and the endosphere of *M. fraileana* that solubilized insoluble phosphorous and several types of rocks (Puente, Li, et al. 2004; Puente et al. 2009b; Lopez et al. 2011). Moreover the characterized bacteria promoted the growth of young cardon cactus growing on rocky substrates, demonstrating the role of this bacteria in the development of rock-weathering cacti (Puente, Bashan, et al. 2004; Puente et al. 2009a). Similarly, mycorrhizal fungi have shown to influence, nutrient status, water uptake and gas exchange in *Ferocactus acanthodes* and *Opuntia ficus-indica* (Cui & Nobel 1992). Seed-borne bacteria from *O. robusta* and *M. geometrizans* did not only show canonical plant growth promoting traits, but also drought tolerance traits like the production of exopolysaccharides, osmotic resistance and high temperature resistance that might be important for the holobiont's fitness (Fonseca-García et al. 2016), but functional assays are needed to test this hypothesis.

ESTABLISHMENT OF CACTI IN ARID ECOSYSTEMS

Seed germination and seedling establishment are critical stages during plant development. The rage of germination is conditioned by osmotic and matrix potentials (Hegarty & ROSS 1978), therefore drought and salinity decrease water potentials in soil and limit seed germination (Vallejo et al. 2010). Cacti seeds are no exception because drylands receive low and sporadic amounts of precipitation (UNCCD & Unep 2011) and optimal conditions may never occur. Additionally, several species of cacti have specific light and temperature requirements (Nobel 1988).

As mentioned above, cacti possess many adaptive traits to survive on arid environments such as big succulent bodies, CAM metabolism, thin and superficial roots, prickles and thick cuticles. Contrary, cacti seedlings have less water stored, are smaller and some species have C3 metabolism during this stage. Consequently, seedlings are very sensitive to drought, high solar radiation and high temperatures and experience high mortality rates in arid environments (Nobel 1988; De la Barrera & Smith 2009)

Seed germination and seedling establishment of cacti can occur under the shade of nurse plants that protect them from high solar radiation and provide enough moisture (Rojas-Aréchiga & Vázquez-Yanes 2000). Interestingly, bacterial endophytes of cardon seeds can promote the growth of cardon seedling on rocky-substrates. Therefore we wonder if seed-borne bacteria can also be helpful during germination and seedling survival in arid environments.

III. JUSTIFICATION

Dry-lands cover 40% of the Earth's land surface and based on their aridity index, are classified into: hyper-arid deserts, arid, semi-arid, and dry sub-humid environments. Dry-lands are characterized by the low and variable amount of precipitation, high solar radiation, extreme temperatures, and high potential of evaporation (UNCCD & Unep 2011; Reynolds et al. 2007), high salinity and acidity, and low nutrient availability (Soussi et al. 2015). Drylands are expected to increase up to 50-56% by the end of the twenty-first century as a consequence of global warming, rapid economic development, urbanization, population growth (Huang et al. 2015), and over exploitation of land and natural resources.

Desertification is characterized by soil erosion, nutrient and water depletion, the increase of salinity and the disruption of biological cycles (UNCCD & Unep 2011). A direct consequence of the increase of drylands is the decline in animal and crop productivity (Reynolds et al. 2007). The effects of global warming and desertification on agriculture will be more severe in developing countries were 78% of the dryland expansion and 50% of the population growth will occur by 2100, and because they rely more on agricultural rather than industrial economies (Cline 2008; Huang et al. 2015).

Microorganisms associated with plants that are adapted to arid and semiarid environments (e.g. cacti and agaves) might have an important role in fitness of these plants, especially those adapted to the unique seed environment and to be vertically transmitted. Also, many authors have shown the role of microorganisms to alleviate drought stress in plants, suggesting that they can be used as fertilizers to enhance crop productivity under arid and semi-arid environments. This approach is very important and should be taken in account in transdisciplinary efforts to combat desertification and ensure a sustainable development in drylands.

Functional studies are often difficult because plants harbor an abundant and diverse microbiota in all their tissues and many of them are uncultivable yet. Seed-transmitted microorganisms represent small and less diverse communities that are transmitted from parents to offspring. This small community might be the first to colonize plant tissues after germination and will allow us to obtain more holistic conclusions regarding the contribution of each member of the community to plants' fitness in drylands.

IV. HYPOTHESIS

Seed-borne endophytic bacteria of cacti contribute to plant fitness by promoting seed germination and seedling growth under drought.

v. OBJECTIVES

MAIN OBJECTIVE

Evaluate the contribution of seed-borne endophytic bacteria of cacti to plant fitness under drought in early stages of plant development.

SPECIFIC OBJECTIVES

Assess the composition of seed-borne endophytic bacteria in natural populations of *Opuntia* spp. and *M. geometrizans*

Evaluate the effect of seed-borne endophytic bacteria on *Arabidopsis thaliana* under hydric stress.

Identify genomic differences between cacti seed-borne bacteria and their relatives with different lifestyles

VI. METHODS

ASSESSMENT OF MICROBIAL COMMUNITY COMPOSITION

- Sampling sites

Cacti fruits were harvested in two different locations in Guanajuato State, Mexico: El Magueyal and San Felipe, on May of 2016 (Figure M1). Flat-padded *Opuntia* spp. (including *O. robusta*), cylindrical-jointed *Cylindropuntia* sp. and *M. geometrizans* individuals were selected on each site based on their general appearance, presence of unspoiled fruits, and proximity between them (Table M1). Three or more fruits were collected using tweezers (without damaging their integrity) kept in plastic bags and brought to the lab.

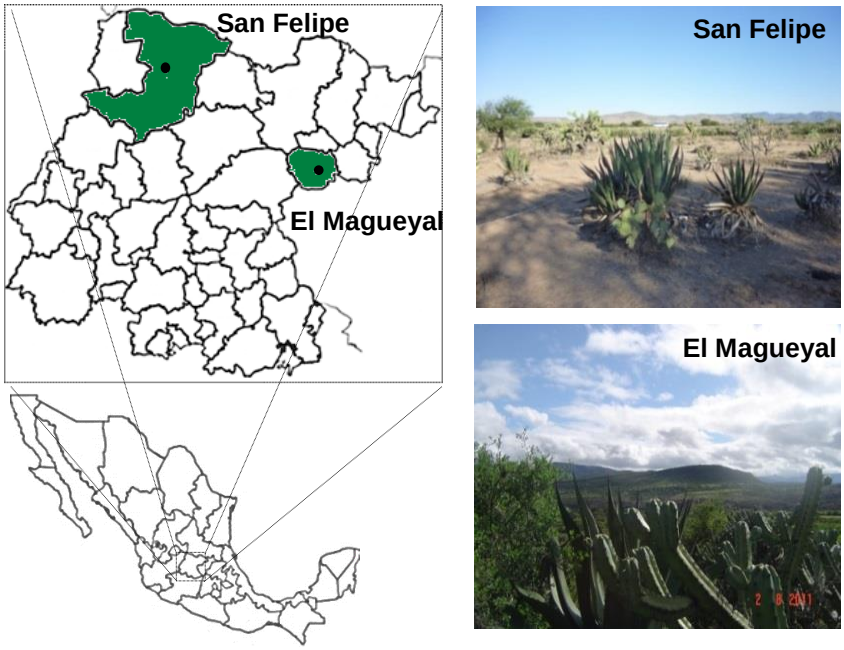


Figure M1. Sampling sites. Both sites are on Guanajuato State, Mexico and have many sympatric CAM plants as cacti and agaves.

- Sample processing

Seeds were extracted from fruits and submerged in 1% H_2SO_4 to hydrolyze the remaining mesocarp (pulp). After acid treatment, the remaining pulp and seeds were sieved and washed with abundant tap water until the seeds were clean and pulp-free. Clean seeds were air-dried and stored in 50 ml Falcon tubes.

Table M1. Experimental design

Identification in site	Sampling site	Number of individuals	Morphological characteristics	Sample key
O. robusta	El Magueyal	3	Flattened-Padded	Anor
Opuntia sp.	El Magueyal	3	Flattened-Padded	Anob
Opuntia sp.	San Felipe	3	Flattened-Padded	Sfob
Opuntia sp.	El Magueyal	3	Tree-like	Anot
Opuntia sp.	San Felipe	3	Tree-like	Sfot
Cylindropuntia sp.	El Magueyal	1	Shrubby	Ancy
Cylindropuntia sp.	San Felipe	4*	Shrubby	Sfcy
M. geometrizans	El Magueyal	3	Tree-like	Anmy
M. geometrizans	San Felipe	3	Tree-like	Sfmy

Mature and immature fruits were harvested from 1 individual

Table M2. Environmental conditions and soil characteristics of the study sites

Sites	El Magueyal	San Felipe
Coordinates	N21°05.106 W100°17.653	N21°39.626 W100°02.959
Altitude (masl)	2175	2089
Environmental conditions in 2012*		
Annual mean temperature (◦C)	18.3	17
Annual mean precipitation (mm)	485	204.5
Precipitation during rainy season (mm)	355	136.5
Precipitation during dry season (mm)	130	68
Soil characteristics		
Texture	Sandy loam	Sandy loam
pH	5.54	6.67
Organic matter (%)	3.26	0.57
Nitrogen (ppm)	1.44	1.44
Phosphorous (ppm)	29.37	4.29
Potassium (ppm)	246.85	351.15
Iron (ppm)	114.76	7.49
Salinity	Very low	Medium low

***CONABIO

Seeds were surface-sterilized in 2ml tubes with 70% ethanol during 3 minutes, NaHCO (Cloralex 15%) during 20 minutes with soft shacking and 3 times washed with sterilized distilled water. A 50 µL drop of the supernatant was inoculated in TSA plates and incubated during 48 h at 28°C to test the effectiveness of the sterilization. 0.1 g of sterilized seeds were vigorously grinded with liquid Nitrogen in a mortar. Grounded tissue was transferred to 1.5 ml tubes and total DNA was extracted following the modified CTAB method of

Edwards (2001), whereby we substituted chloroform with phenol/chloroform/isoamyl alcohol (25:24:1) for extraction. DNA was measured in a nanoDrop and stored at -20°C.

- Molecular characterization of plants

We amplified the internal transcribed spacer region (ITS) of nuclear rRNA using the primers ITS1 and ITS5 (White et al. 1990) and the trnL-F non-coding region of chloroplast using the primers trnE and trnF (Taberlet et al. 1991). PCR runs were performed in 20 µL reactions, containing 2 µL of 10X PCR buffer (Qiagen), 0.8 µL of 25 µM $MgCl_2$, 0.4 µL of 10 µM dNTPs (Qiagen), 0.5 µL of DMSO, 0.8 µL of both 5 µM reverse and forward primers (10 µM for trnE and trnF), 1 unit of Taq DNA polymerase (Qiagen) and 1 µL of 40 ng/µL of template DNA. Thermal cycling conditions are summarized in Table M3, PCR products were visualized in 1% agarose gels. Amplicons were purified using the GenAll Expin™ purification kit and sequenced at the Genomic Services facility of the Advanced Genomic Unit-LANGEBIO using Sanger technology.

ITS and tnrL-F partial gene sequences were aligned using Muscle (Edgar 2004). Gene alignments were concatenated and a two-gene phylogeny was constructed with the MEGA7 software (Kumar et al. 2016) based on the Jukes-Cantor substitution model (Jukes & Cantor 1969) and the Maximum Likelihood method. Plants were identified based on morphological clues.

Table M3. Thermal cycling parameters

Parameter	16S rRNA	ITS	trnL-F	16S rRNA V4	ITS2
Initial denaturalization (°C / min)	95 / 2	94 / 3	94 / 4	94 / 3	94 / 3
Denaturalization (°C / sec)	95 / 40	94 / 45	94 / 45	94 / 45	94 / 45
Initial annealing (°C / sec)	-	-	-	78 / 10	-
Annealing (°C / sec)	55 / 30	53 / 45	55 / 45	50 / 60	50 / 60
Amplification (°C / sec)	72 / 90	72 / 60	72 / 45	72 / 90	72 / 90
Cycles	25	35	30	30	35
Final amplification (°C / min)	72 / 7	72 / 7	72 / 5	72 / 10	72 / 10

- Amplicon sequencing of microbial communities associated with cacti seeds

Libraries for HiSeq Illumina amplicon sequencing were prepared using the ready-to-use Hot Master Mix (5 PRIME) kit. For prokaryotic communities, we amplified the V4 region of the 16S rRNA gene with the primers 341F (5'-CCTACGGGNBGCASCAG-3') and 785R (5'-GACTACNVGGGTATCTAATCC-3'). For the eukaryotic community, we amplified the ITS2 region of nuclear rRNA with the primers ITS9F (5'-GAACGCAGCRAAIIGYGA-3') and ITS4R (5'-TCCTCCGCTTATTGATATGC-3'); every primer had with a linker, unique index, pad and spacer sequences in the 5' end. PCR runs were performed in 25 µL reactions, containing 10 µL of 2.5X of hot master mix (5 PRIME), 1 µL of 10 mg/mL BSA, 0.5 µL each 10 µM reverse and forward primers and 1 µL of 40ng/µL of template DNA, 0.31 µL of 10 µM PNA1 and PNA2 primers were added to 16S PCR reactions. Thermal cycling conditions are summarized in Table M3, PCR products were visualized in 1% agarose gels and the amount of product was quantitatively measured in a Qubit® fluorometer. 100ng of each sample were pooled and purified using the Agencourt AMPure XP kit. Quality of the purified products was analyzed at the Advanced Genomic Unit of LANGEBIO with a CHIP Bioanalyzer system.

ISOLATION AND CHARACTERIZATION OF CACTI SEED-BORNE MICROORGANISMS

- Isolation of cacti seed-borne microorganisms

For isolation of seed-borne endophytes, we used seeds and seedlings of *O. robusta* and *M. geometrizans* propagated *in vitro* by Fonseca-García et al. (2016). Seedlings and surface-sterilized seeds were grinded with 0.5-1 mL of sterilized water. 100 µL of supernatant were plated onto TSA plates for isolation of heterotrophic aerobic bacteria, 100 µL of supernatant and tissue leftovers were inoculated onto Winogradsky plates and semisolid Winogradsky medium, respectively, for the isolation of Nitrogen fixing bacteria. Additionally, tissue leftovers were inoculated on PDB media and seedlings were cut lengthwise and immersed on 0.5X PDA medium for the isolation of fungi. Cultures were incubated at 28°C for 3 weeks. Isolated bacteria were sub-cultured in TSA plates until axenic and pure cultures were obtained. No filamentous fungi nor yeast were obtained.

Total genomic DNA was isolated following the same methodology as Desgarennes et al. (2014) . We amplified the 16S rRNA gene from every isolate using primers 16rna27F (5'-AGAGGTTTGATCCTGGCTCAG-3') and 16rna1492R (5'- GGTTACCTTGTTACGACTT-

3'). PCR runs were performed as mentioned above, thermal cycling conditions are summarized in Table M3. PCR products were visualized in 1% agarose gel electrophoresis and cloned with pJet1.2/blunt in *Escherichia coli.* One clone per strain was sequenced at the Genomic Services facility of the Advanced Genomic Unit-LANGEBIO using Sanger technology.

16S rRNA gene sequences were analyzed using BLAST tool of NCBI to retrieve homologous reference sequences with the best coverage and similarity. Sequences were aligned in Muscle tool (Edgar 2004) using default parameters and the phylogeny was constructed with MEGA7 (Kumar et al. 2016) software using Kimura 2 –parameter model (Kimura 1980) and the Maximum Likelihood method. Seed-borne strains were identified base on the sequence similarity against reference strains.

- Identification of seed-borne bacteria in the cacti microbiome data

OTU sequences assigned to the same genus or family of characterized seed-borne strains were extracted from the cacti microbiome database (Fonseca-García et al. 2016). Sequences were aligned using Muscle tool (Edgar 2004) with default parameters and the phylogeny was constructed with MEGA7 (Kumar et al. 2016) software using Kimura 2 –parameter (Kimura 1980) model and the Maximum Likelihood method. OTU sequences matching 16S rRNA sequences of se-borne strains with more than 97.0% of sequence similarity were considered as the same OTU and present in the microbiome of cacti. Relative abundance vs frequency plots were made using ggplot2 (Wickham 2009) package of R.

EFFECT OF CACTI SEED-BORNE BACTERIA *IN PLANTA*

PEG-infused plates (-0.5 MPa) were prepared following van der Weele et al. (2000) and Verslues et al. (2006) with some modifications. 0.5X MS solution was filtered through conventional filter paper before adding PEG 8000 (no MES buffer was added), then PEG solutions were filter-sterilized using 0.45 µm DURAPORE membrane filters (MILLIPORE). 24 ml of PEG solution was poured onto the top of plates with 16 ml of 0.5X MS plates (15 g/L agar), plates were wrapped in plastic paper and allowed to equilibrate during 24h. After equilibration, PEG solution was carefully poured-off the plates making sure no solution remained on the surface. For standard water potential plates, 0.5X MS solution was used instead of PEG solution.

Per each bacterial strain tested, 4 tubes containing 100-120 seeds of *A. thaliana* col-1 were used. Seeds were surface-sterilized with 500 µL of 96% ethanol during 5 minutes with shaking, 700 µL NaHCO (Cloralex 20%) during 7 minutes with shaking and washed 6 times with 700 µL of sterilized distilled water. Disinfected seeds were vernalized at 4°C during 48 h to synchronize germination. A 50 µL drop of the supernatant was inoculated in TSA plates and incubated during 48 h at 28°C to test the effectiveness of the sterilization.

Seed-borne strains were grown overnight in TSB at 28°C, except *N. prasina* L17 who was incubated 5-7days prior to experiments. Cultures were pelleted by centrifugation, washed once with 0.85% NaCl and based on the number of disinfected seeds in each tube, they were suspended to 10, 1000 and 100000 cfu/seed (*N. prasina* was suspended to 10000 cfu/seed) in 1 ml of 0.85% NaCl. Each suspension was pelleted, the supernatant was discarded and the pellet suspended on the remaining solution using the vortex.

Sterilized seeds were transferred to their corresponding bacterial suspension and gently shaken, 0.85% NaCl was used for non-inoculated seeds. 12 seeds per treatment were transferred onto the surface of low and standard water potential plates following a straight line on one side of the plate with 0.5 cm of distance between seeds. Plates were incubated vertically in a growth chamber at 21°C with a day/night cycle of 16/8 h during 12 days. Remaining seeds were transferred to TSA plates and incubated at 28°C during 48 h to corroborate the presence or absence of bacteria.

Germination rate was recorded within the first 4 days, then root length was measured every two days. At the end of the experiment, lateral roots were counted, plants were transferred to 1.5 ml tubes and dried at 60°C during 24 h. Dry weight and humidity were determine. A picture of the rosettes was taken to measure leaf area using ImageJ. Plots and statistical analysis were performed with R software (R Core Team 2015).

COMPARATIVE GENOMICS OF CACTI SEED-BORNE BACTERIA

- Genomic overview and plant growth promotion traits

Seed-borne strains genomes were sequenced at the DOE Joint Genome Institute (JGI) using Illumina HiSeq 2500 technology. Only 3 draft genomes were available during the course of this thesis. 24 other genomes were selected from IMG/M data base from the same species (*S. hominis*) and genera (*A. baldri* and *P. quisquiliarum*) for comparison and are

summarized in Table M4. Hierarchical clustering was performed on the IMG/M portal (Chen et al. 2016) and plots were made using R software (R Core Team 2015). Protein prediction for each genome can be searched manually in the IMG/M portal, but the name of the products can differ in uppercases, spaces, hyphens etc. To overcome this situation, we retrieved all the possible product names for proteins related to plant growth promotion and stress tolerance traits, then we downloaded the protein prediction for all the genomes and searched for the complete list of product names in each genome. Finally we counted the presence of the predicted proteins and plotted the data using ggplot2 (Wickham 2009) package of R.

Table M4. Selected genomes for comparative genomics

Taxon_id	Lifestyle	Genome Name
651285007	Human microbiome	*Staphylococcus hominis hominis* C80
2654587786	Poison ivy endophyte	*Staphylococcus hominis* RIT-PI-k
2599185353	Switchgrass root endophyte	*Staphylococcus* sp. NFPP34
2623620554	*O. robusta* seed	*Staphylococcus hominis* hominis ZBW5
651285006	Human microbiome	*Staphylococcus caprae* C87
643886010	Human microbiome	*Staphylococcus capitis* SK14
645058825	Human microbiome	*Staphylococcus warneri* L37603, SK66
2648501903	Rice seed	*Staphylococcus warneri* SA9
2600254943	Switchgrass root endophyte	*Staphylococcus pasteuri* NFIX07
2654587624	Rice seed	*Staphylococcus epidermidis* SA8
2534681825	Human microbiome	*Staphylococcus epidermidis* BVS058A4
2537562024	Human microbiome	*Staphylococcus lugdunensis* ACS-027-V-Sch2
2675903256	Rice seed	*Staphylococcus xylosus* NS341
2537562023	Human microbiome	*Staphylococcus simulans* ACS-120-V-Sch1
2558860399	Human microbiome	*Staphylococcus aureus* KPL1845
2623620615	*O. robusta* seed	*Paenisporosarcina quisquiliarum* SK 55
2551306413	Sediment-laden stratified basal ice	*Paenisporosarcina* sp. TG-14
2551306364	Debris-rich basal ice	*Paenisporosarcina* sp. TG20
2541046999	Human microbiome	*Paenisporosarcina* sp. HGH0030
2634166279	Dried seaweed	*Agrococcus jejuensis* DSM 22002
2634166154	Soil from a coal mine	*Agrococcus carbonis* DSM 22965
2597489936	Cold thrombolytic microbialites	*Agrococcus pavilionensis* RW1
2524023149	Soil below an ice glacier	*Agrococcus lahaulensis* DSM 17612
2623620613	*O. robusta* seed	*Agrococcus baldri IAM* 15147

To dissect the differences between lifestyles we compared protein domains among a set of genomes to assess for reduction or expansion of functions within a statistical framework using edgeR (Robinson et al. 2009) package of R as performed by (Bermúdez-Barrientos 2016). Significant enriched or reduced COG groups were selected based on p-value and FDR values.

- Orthology analysis of seed-borne strains

Orthology analysis was performed in the Laboratory of Genomic conSequences at Wilfrid Laurier University under the guidance of Gabriel Moreno-Hagelsieb. We selected only 3 bacterial genomes for each lineage of seed-borne strains with different lifestyles (plant, human or environmental). First we re-annotated the genomes using rpsblast, this tool compares gene sequences with a collection of conserved domains, then we cleaned the non-significant and overlapping hits using R software (R Core Team 2015) and translated CDD domains to COG groups. We identified putative orthologues based on the Best Reciprocal Hits, whereby two genes in two different genomes are orthologs if they found each other as the best hit (Tatusov et al. 1997; Bork et al. 1998) when comparing them using BLAST tool as described by Moreno-Hagelsieb & Latimer (2008). COG groups and categories from accessory genomes of seed-borne strains were retrieved and plotted using ggplot2 (Wickham 2009) package of R.

VII. RESULTS

COMPOSITION OF CACTI SEED-BORNE MICROBIAL COMMUNITIES

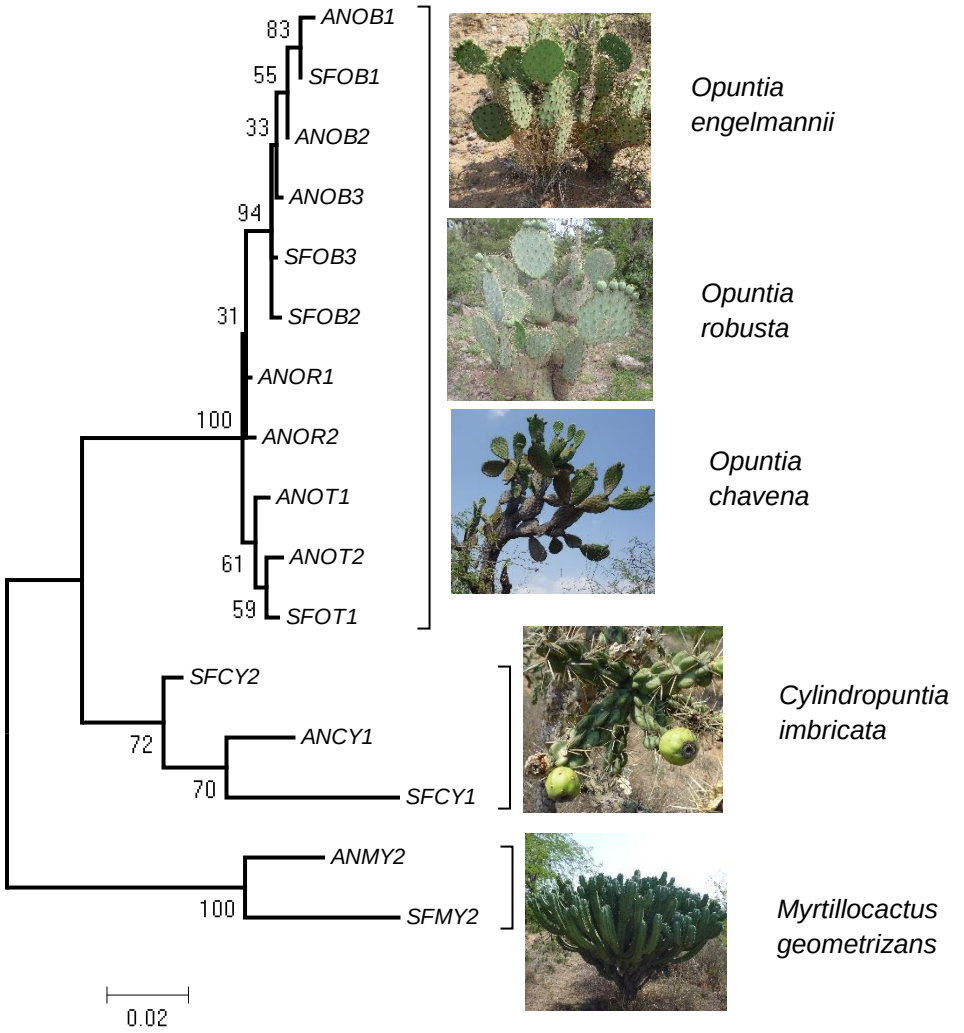


Figure R1. Phylogenetic analysis of chloroplast trnL-F and nuclear ITS partial sequences of samples. The evolutionary history was inferred by using the Maximum Likelihood method based on the Jukes-Cantor model.

The morphological characterization of the plants on site correlated with the phylogeny based on trnL-F and ITS sequences (Figure R1). *Opuntia* sp. individuals with green medium-sized, heart-shaped cladodes, dark brown areoles and big yellow prickles (Anob and Sfob samples) clustered together regardless of the site. These plants were very similar to *Opuntia robusta* (big, round and pale green cladodes), whose phylogenetic position was not solved. *Opuntia* sp. individuals with tree-like structures, oval cladodes and small white

prickles (Anot and Sfot samples) also clustered together regardless the sampling site. Flat-stemmed *Opuntia* spp., terete-stemmed *Cylindropuntia* sp. and *M. geometrizans* recapitulate the phylogeny of cacti (Hernández-Hernández et al. 2011; Griffith & Porter 2009)

16S (V4) and ITS2 libraries for amplicon sequencing were analyzed with CHIP Bioanalyzer (Figure R2). After purification, we obtained an abundant amplicon of around 450 bp in the ITS2 library while many non-specific bands were present in higher abundance that the expected amplicon in the 16S library (600 bp approximately). ITS2 library has been already sequenced by now, but a complete analyses of the fungal taxa associated with cacti seeds remains to be done.

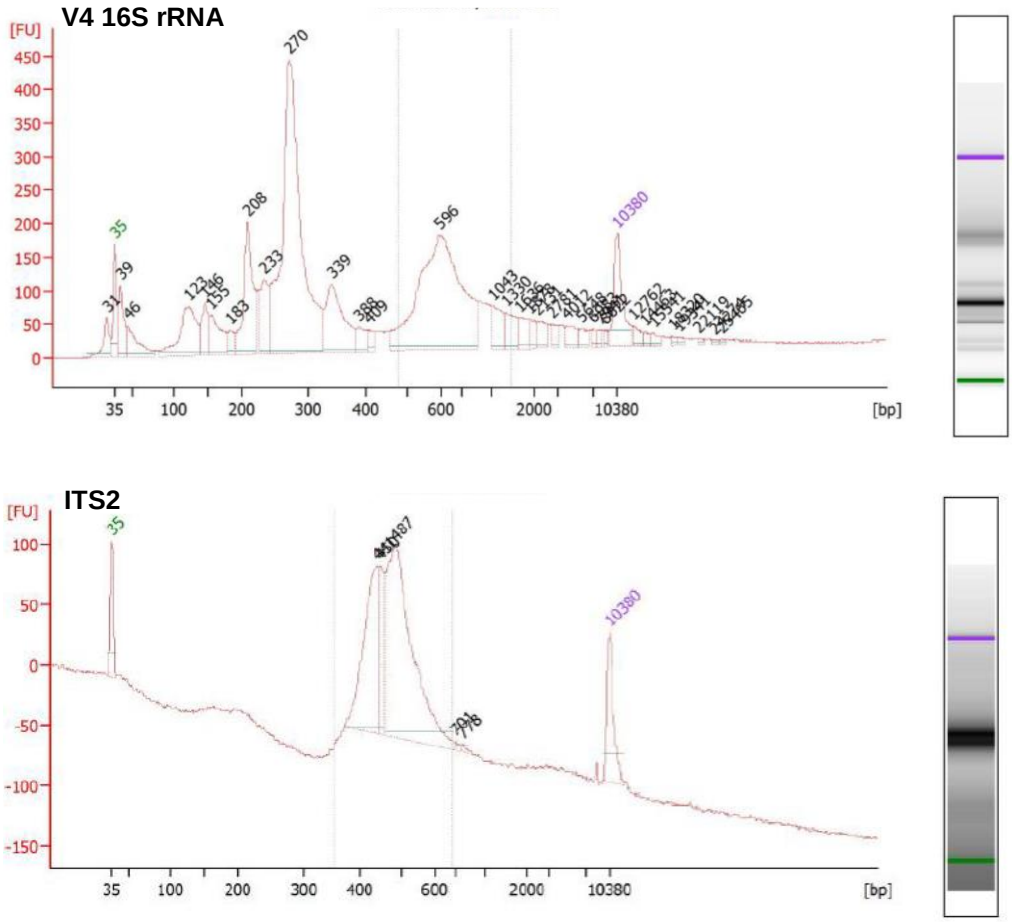


Figure R2. CHIP Bioanalyzer analysis of V4 16S rRNA and ITS2 libraries. The electropherogram is shown in the left and the electrophoresis gel is shown in the right.

CHARACTERIZATION OF CACTI SEED-BORNE BACTERIA

- Taxonomic characterization of isolated seed-borne bacteria

Isolation frequency of seed-borne bacteria was very low. Only eleven bacterial strains were isolated from *O. robusta* and *M. geometrizans* seeds and seedlings (Table S1). These strains were grouped in four morphotypes based on their colony and microscopic morphology (Table R1). According to their 16S rRNA gene partial sequence analysis, one of the morphotypes was identified as *Paenibacillus* sp. (D1 strain) and the rest as *Bacillus* sp (A, B and C strains). *Bacillus* strains could not be differentiated from each other based on their 16S rRNA sequence, as they were very similar to that of many other *Bacillus* species (Figure R3) including *B. subtilis* and comprised a different clade from the *Bacillus* strains isolated in 2013 (Fonseca-García, *et al*. 2016). *Paenibacillus* sp. was very different from the reference sequences with the best BLAST hit and might represent a different species.

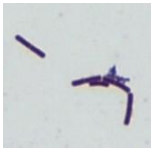
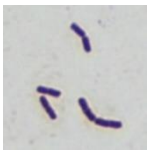
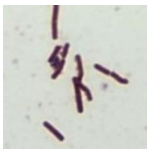
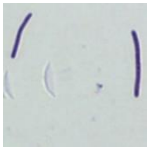
These strains together with the ones from Fonseca-Garcia, *et al.* (2016) composed a collection of 28 cacti seed-borne strains encompassing 8 different genera: *Agrococcus, Nocardiopsis, Leclercia, Staphylococcus, Bacillus*, *Paenibacillus*, *Psychrobacillus* and *Paenisporosarcina.* Interestingly, *Firmicutes* that are known for producing endospores (e. *Bacillus* spp., *Psychrobacillus* spp.) were frequently isolated from cacti seeds and seedlings.

- Presence of seed-transmitted bacteria in the cacti microbiome

Partial 16S rRNA sequences of the 28 seed-borne strains were compared against the measurable OTU data base of the cacti microbiome using a phylogenetic approach. Figure R4 shows that all the strains matched at least 1 OTU from the cacti microbiome (Fonseca-García, *et al*. 2016) with more than 97% of sequence similarity which was the cut off value for prokaryotic OTU clustering. For example seed-borne *Leclercia* sp. L16, *Paenibacillus* sp. D1, *S. hominis* L12 and *N. prasina* L17 matched OTU 4 (*Leclercia*), OTU 1942 (*Paenibacillus*), OTU 99 (*Staphylococcus*) and OTU 680 (*Nocardiopsaceae*), respectively. Seed-borne *Bacillus* formed a large clade matching more than one OTU, but they had the highest similarity (97.7%-100%) with OTU 19 (*Bacillus*). Also *Psychrobacillus/Paenisporosarcina* strains also formed a single clade with OTU 29 that is assigned to *Paenisporosarcina* genus (96.9-98.4% of sequence similarity). Seed-borne

Agrococcus baldri L15 matched OTU 2692 (*Curtobacterium*) with 96.9% of similarity so it could be considered a different OTU that is not present in the microbiome of adult cacti.

Table R1. Morphological characterization of bacteria strains isolated from cacti seeds and seedlings

Morphotype A		White, 5-7mm, amorphous, rough and acuminate surface, wavy edge, matte, translucent, dry appearance without diffusible pigment
		Single Gram positive bacilli and attached as pairs or strains
Morphotype B		Beige, 8-9mm, amorphous, smooth and convex surface, entire edge, bright, translucent, moist and mucous-like, without diffusible pigment
		Single Gram positive bacilli and attached as pairs
Morphotype C		Beige, 3mm, amorphous, rough surface, flat, slightly tortuous, matte, opaque, dry appearance, without diffusible pigment
		Single Gram positive bacilli and attached as pairs
Morphotype D		Beige, 2mm, amorphous, flat, smooth surface, entire edge, glossy, translucent wet look
		Single and long Gram positive bacilli

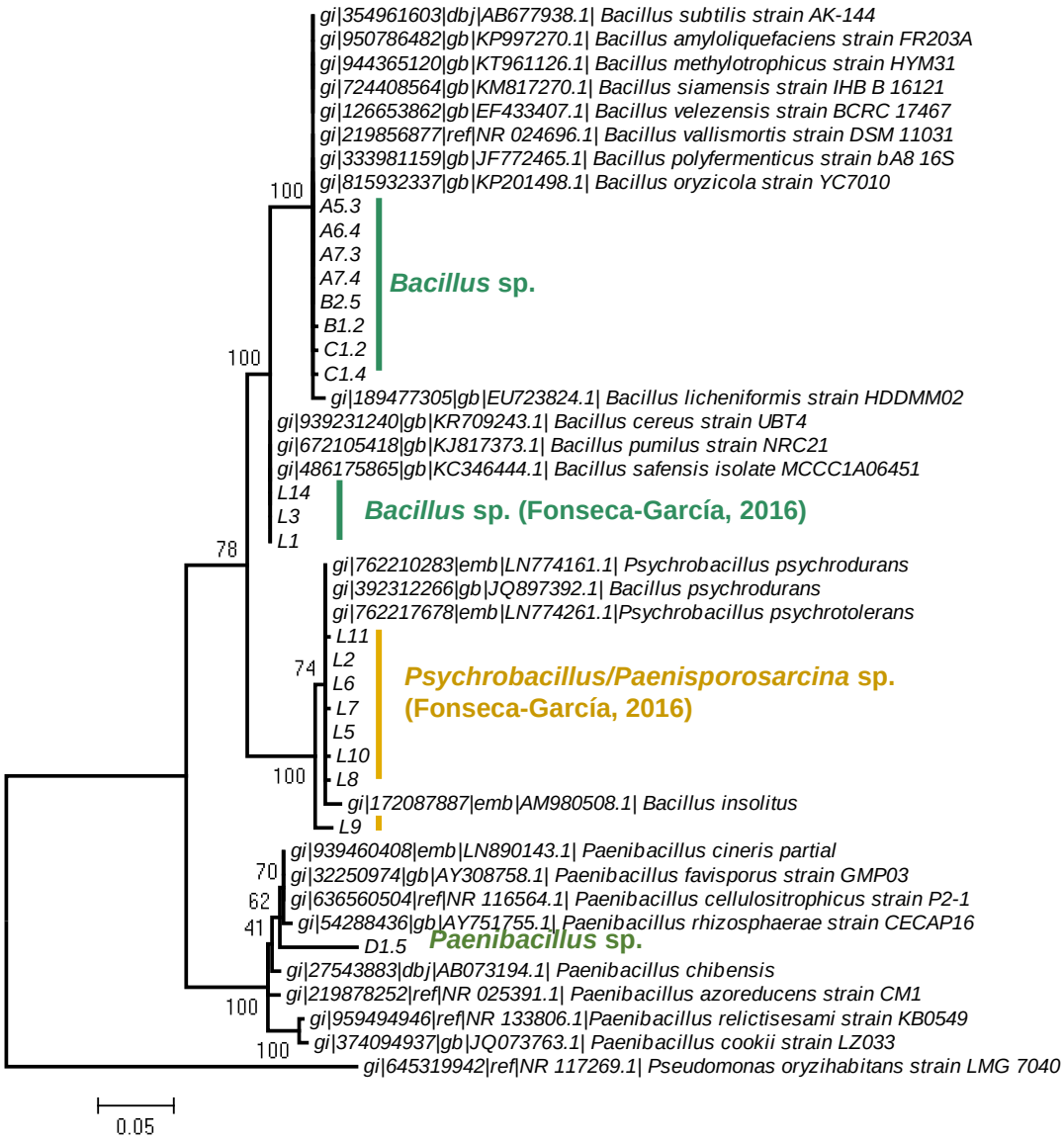
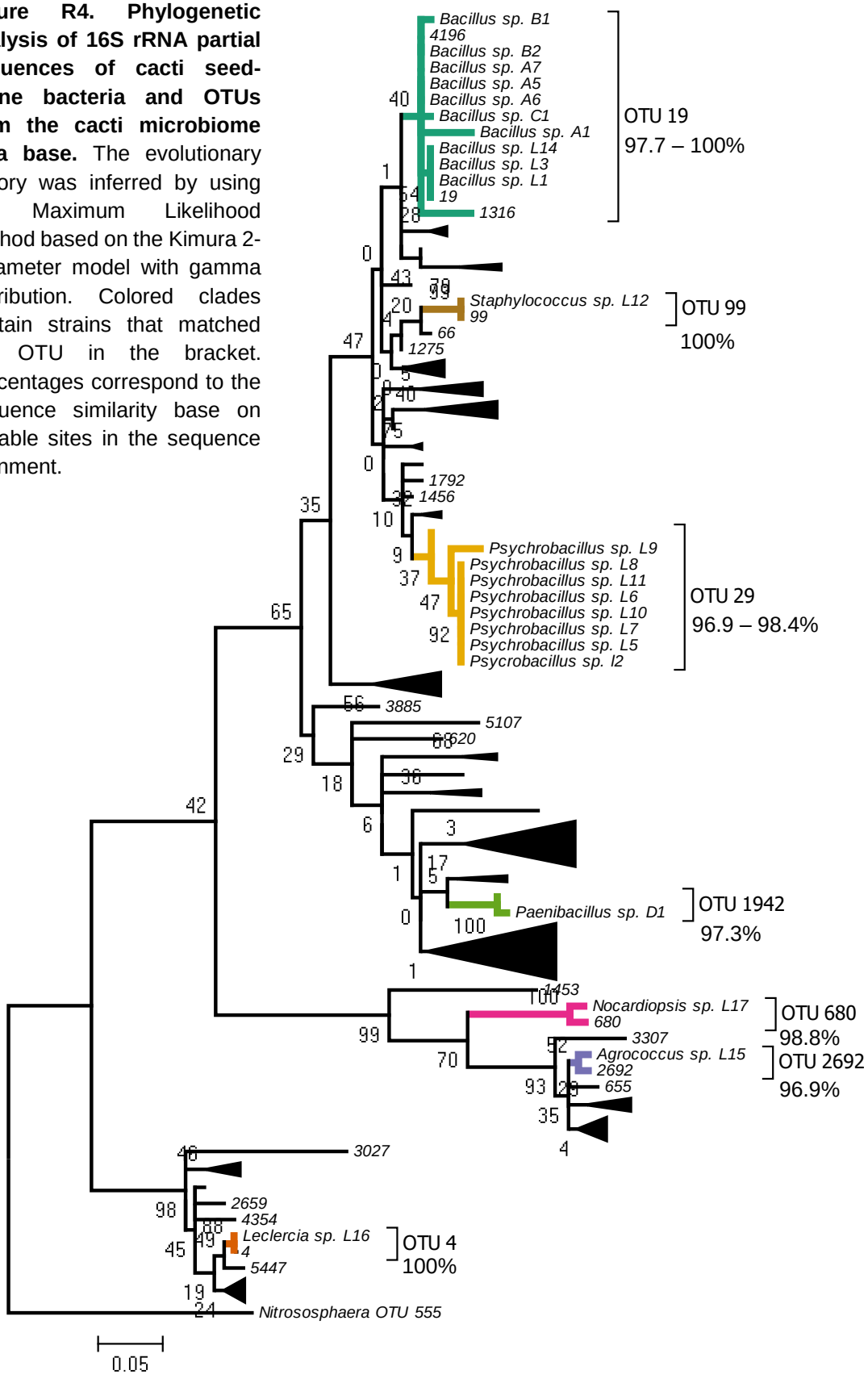


Figure R3. Phylogenetic analysis of 16S rRNA partial sequences of bacteria from cacti seeds and seedlings. The evolutionary history was inferred by using the Maximum Likelihood method based on the Kimura 2-parameter model with gamma distribution.

Figure R4. Phylogenetic analysis of 16S rRNA partial sequences of cacti seed-borne bacteria and OTUs from the cacti microbiome data base. The evolutionary history was inferred by using the Maximum Likelihood method based on the Kimura 2-parameter model with gamma distribution. Colored clades contain strains that matched the OTU in the bracket. Percentages correspond to the sequence similarity base on variable sites in the sequence alignment.



To determine if seed-borne bacterial endophytes are maintained through cacti growth and development, we searched for the abundance and frequency of OTUs from the cacti microbiome that matched 16S rRNA gene sequences of seed-borne strains (Figure R5).

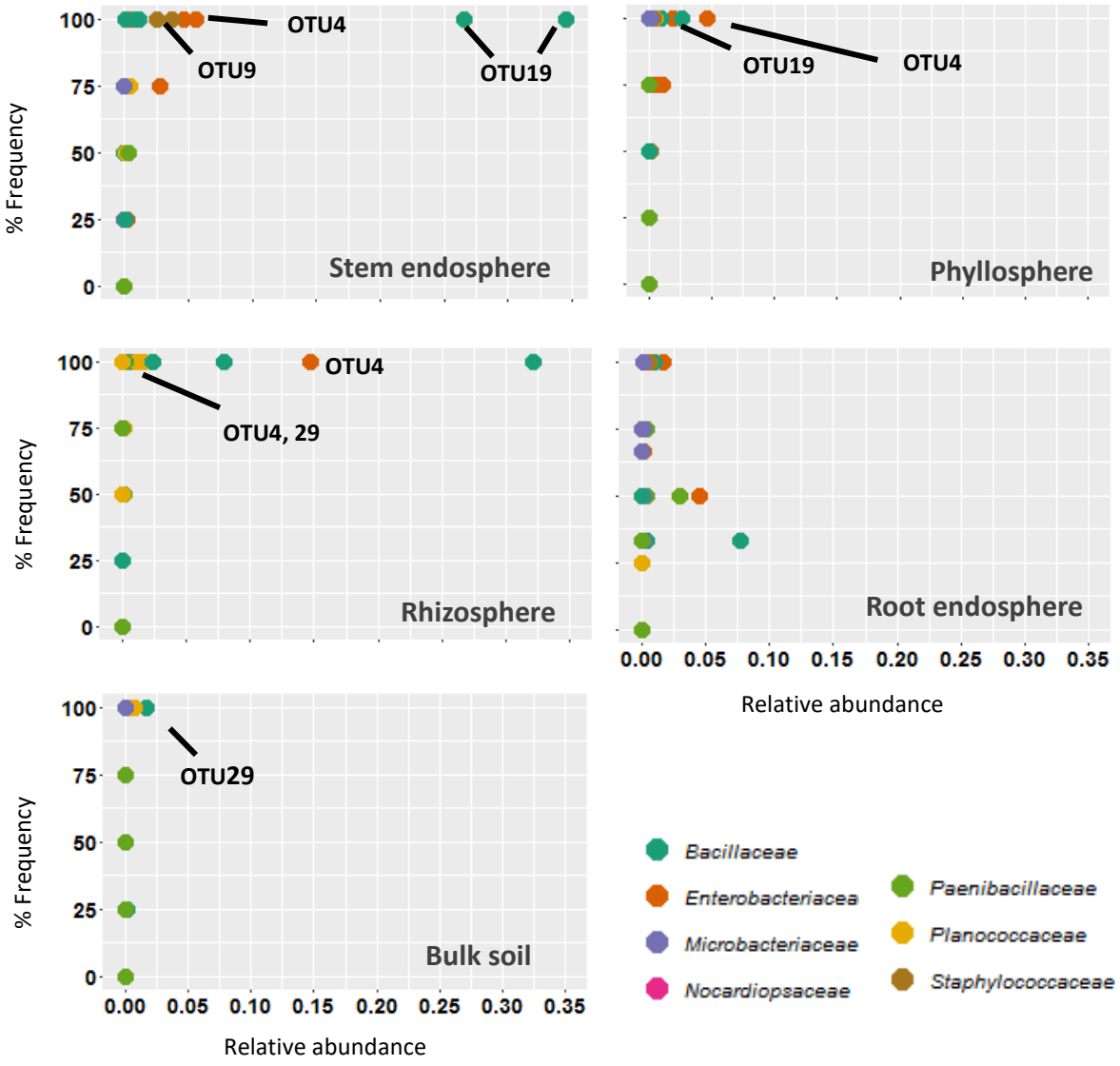


Figure R5. Relative abundance and frequency of OTUs that matched with 16S partial sequences of seed-borne strains. Each dot correspond to an OTU in different plant compartments of *O. robusta* o *M. geometrizans*. Colors correspond to family level classification.

OTU 19 (*Bacillus*) was highly abundant and frequent in the endosphere of *M. geometrizans* and *O. robusta* encompassing 26.6% and 34.5% of relative abundance, respectively. Moreover, OTU 19 was not abundant nor frequent in the rhizosphere, phyllosphere, root

endosphere or bulk soil. OTU 4 (*Leclercia*) and OTU 99 (*Staphylococcus*) were as frequent although less abundant than OTU 19, they comprised between 2.5 and 7.5% of relative abundance in cacti and they were not abundant in other compartments, except OTU 4 that was also abundant in the rhizosphere of *O. robusta* (14.6%). This finding suggests that regardless of whether this OTU are vertically or horizontally transmitted, they cannot proliferate in other compartments.

The other matching OTUs (680, 1942 and 2692) were not frequent nor abundant in any cacti compartment. OTU 29 (*Paenisporosarcina*) was frequently found (100%) in all the compartments but its abundance was very low. This might indicate that some OTUs that are transmitted through the seeds are maintained in very low abundance and others may not remain in the cacti holobiont at all.

EFFECT OF CACTI SEED-BORNE BACTERIA ON *A. thaliana* UNDER LOW WATER ACTIVITY

A value of -0.25 MPa was considered as the standard water potential and -0.50 MPa was considered the low water potential or the hydric stress treatment. No inoculated seeds were used as a control (0 cfu). The effect of cacti seed-borne bacteria on *A. thaliana* is summarized in Figure R6 and R7. Most strains increased lateral root number under standard water potential and only *S. hominis*, *Paenibacillus* sp. and *Bacillus* sp. promoted increased leaf surface. Germination was significantly increased under low water availability by *S. hominis* and *Paenibacillus* sp. None of the effects on the *A. thaliana* phenotype was dependent on the number of cfu of the strain tested.

- Germination

Vernalized *A. thaliana* seeds germinated within two days on standard water potential, but under lower water potential showed a mean decrease of 55% on germination rate on the second day, although after 3 days most treatments reached ~100% of seed germination. Only *S. hominis* L12 and *Paenibacillus* sp. D1 had a significant increase on germination rate on the second day. *N. prasina* L17 and *A. baldri L15* highly and slightly increased this trait, respectively without statistical support (Figure R7, Figure S5). This effect did not depend on the number of cfu of the strain used for inoculation, but the best performance was obtained when using 10 and 1000 cfu/seed. Interestingly, seed germination rate did

not decrease when inoculated with most of the strains, except when inoculated with 100000 UFC/seed *Paenibacillus* sp at D1, regardless of water potential

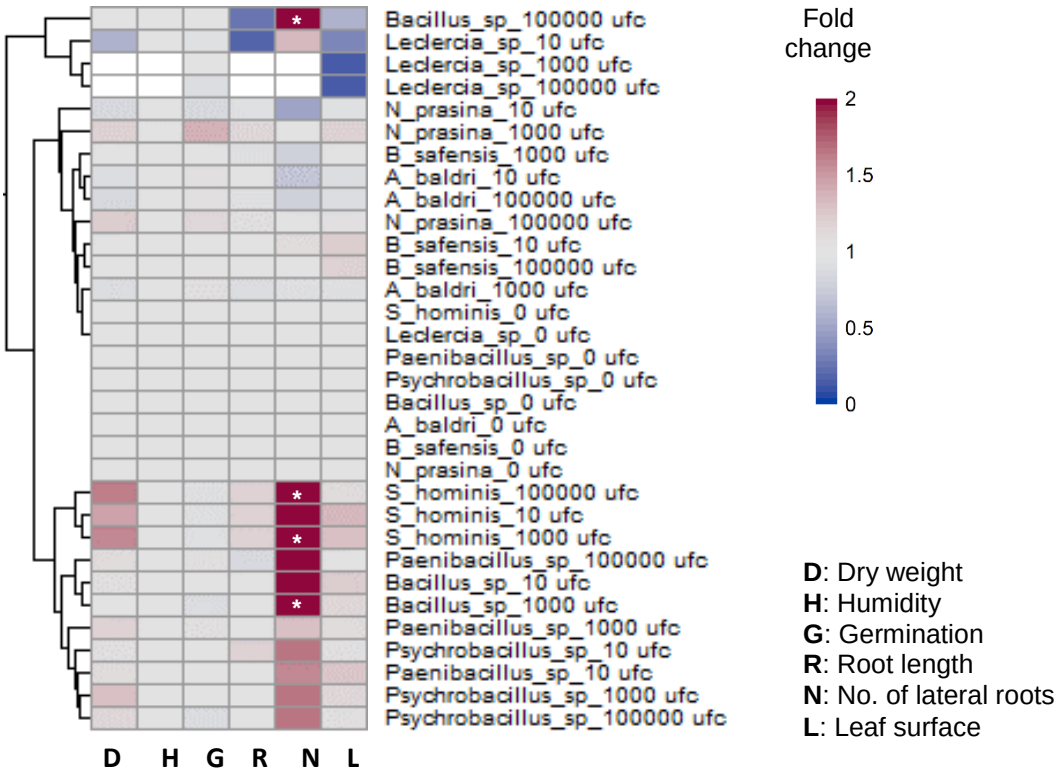


Figure R6. Effect of cacti seed-borne bacteria on *A. thaliana* under standard water potential (-0.25 MPa). Treatments are clustered based on Euclidean distances using UPGMA method. Non inoculated control are labeled as 0 UFC. * indicates treatments which effect is more than two-fold.

- Number of lateral roots.

Lateral root number was strongly influenced by seed-borne strains, especially under standard water potential (Figure R7). *S. hominis* L12, *Bacillus* sp. A1 increased up to two-fold the number of lateral roots under standard water potential regardless the number od cfu inoculated. *Psychrobacillus* sp. L5 and *Paenibacillus* sp. D1 had a moderately positive influence on this trait while *A. baldri* L15 and *N. prasina* L17 had a slightly negative impact under standard water potential. Also, *Leclercia* sp. L16 also showed a negative effect on plant growth. Under low water potential only *Bacillus* sp. A1 (100000 cfu/seed) and *Leclercia* sp. L16 (10 cfu/seed) increased lateral root number, but only on those plants that were not invaded by the bacterial growth (Figure S8).

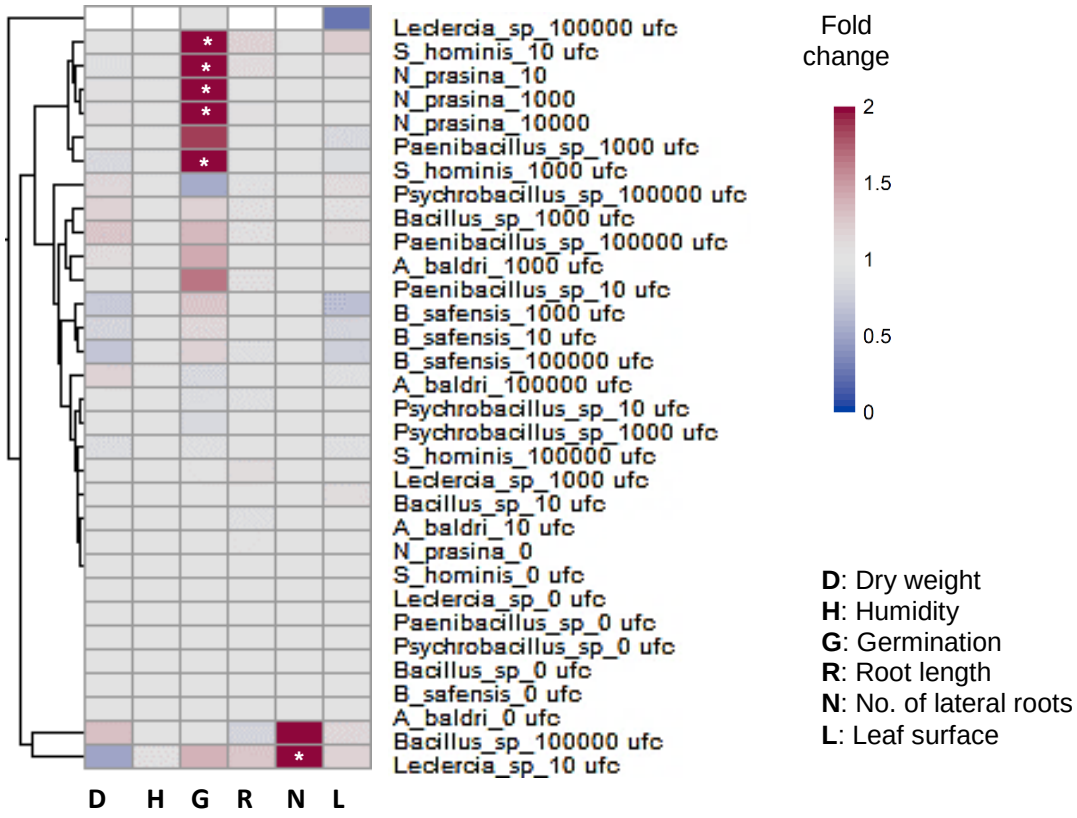


Figure R7. Effect of cacti seed-borne bacteria on *A. thaliana* under low water potential (-0.50 MPa). Treatments are clustered based on Euclidean distances using UPGMA method. Non inoculated control are labeled as 0 UFC. * indicates treatments which effect is more than two-fold.

- Humidity and dry weight

Plant water content was determined as gravimetric humidity of the rosettes. The application of no single seed-borne strain resulted in increased nor decreased water content of the plants under both water potentials. Plants inoculated with *S. hominis* L12 showed a significant increase in dry weight regardless of the cfu applied but it did not correlate with the increase of leaf surface observed in some treatments. Treatment with the rest of the strains did not result in a significant increase or decrease of this trait.

- Root length and leaf surface

Root length and leaf surface was slightly increased under both treatments. Plants inoculated with *Bacillus* sp. A1 (10 and 100000 cfu/seed), *Paenibacillus* sp. D1 (10 UFC/seed) and *Psychrobacillus* sp. L5 (1000 cfu/seed) showed a significant but small increased in leaf

surface while application of *S. hominis* L12 resulted in to two-fold or higher increase in this trait.

Inoculation with *Leclercia* sp. L16 and *Bacillus* sp. A1 (100000 cfu/seed) strongly decreased leaf surface and root length especially under standard water potential. Even though seed inoculated with *Leclercia* sp. managed to germinate within the bacterial biomass they were not able to develop roots nor enough leaf surface.

COMPARATIVE GENOMICS OF CACTI SEED-BORNE BACTERIA.

- Overview

The genomes of only tree cacti seed-borne bacterial strains were available for analysis: *S. hominis* ZBW5 (L12), *A. baldri* IAM 15147 (L15) and *Paenisporosarcina quisquiliarum* SK 55 (L6). *S. hominis* had the best performance in *A. thaliana* assays, it promoted germination under low water availability, and it increased leaf surface and dry weight under standard water potential. Importantly, there were plenty of sequenced genomes of strains of this species from different hosts. *A. baldri* L15 did not have a significant effect on *A. thaliana,* although it was the first plant-associated *Agrococcus* species sequenced, and there were only four *Agrococcus* genomes available of different species from different environmental samples. *P*. *quisquiliarum* L6 was not tested *in planta*, instead *Psychrobacillus* sp. (L5) was used and showed a small effect on leaf surface, and represent a genus closely related to *Paenisporosarcina* genus.

Seed-borne bacteria genomes did not show any evidence of an obligatory symbiotic lifestyle. Genome size, GC content and protein gene number are very similar to plant, human or environmental relatives from the same species and from the same genus (Figure R8). Based on the gene count of pfam categories, strains tended to cluster together regarding their taxonomy rather than their lifestyle. This comparison was only possible with *Staphylococcus* genomes, where seed-borne *S. hominis* SBW5, *S. warneri* SA9, *S. epidermidis* SA5 clustered with their human and plant relatives from the same species.

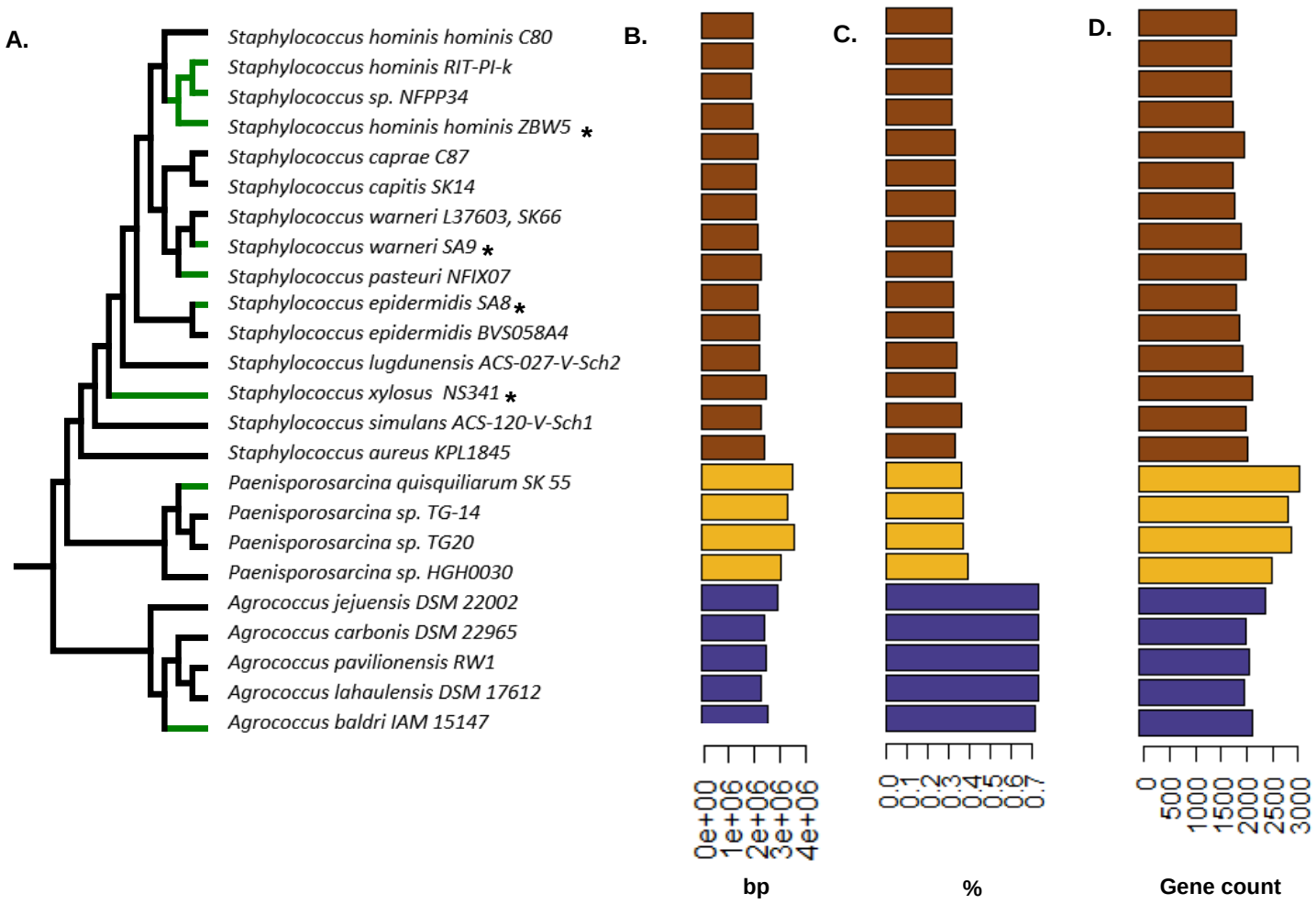


Figure R8. Genomic overview of *Agrococcus*, *Staphylococcus* and *Paenisporosarcina* strains from different lifestyles. A. Hierarchical clustering based on gene counts of pfam categories. B. Genome size. C. GC%. D. Genes with predicted protein product. Green branches indicate plant associated bacteria. * indicates seed-borne bacteria.

- Plant growth promoting traits.

Predicted proteins in seed-borne bacteria genomes showed many of the proteins related to plant growth promotion (Figure R9). We grouped the genomes based on their lifestyle: plant (whether plant tissue or seed), human (mostly Human Microbiome Project) and environmental (many origins). Plant associated *Staphylococcus* tended to possess more butanediol related genes and heat shock proteins than human relatives. Contrary, human associated strains tended to have more genes related to exopolysaccharide excretion and biosynthesis.

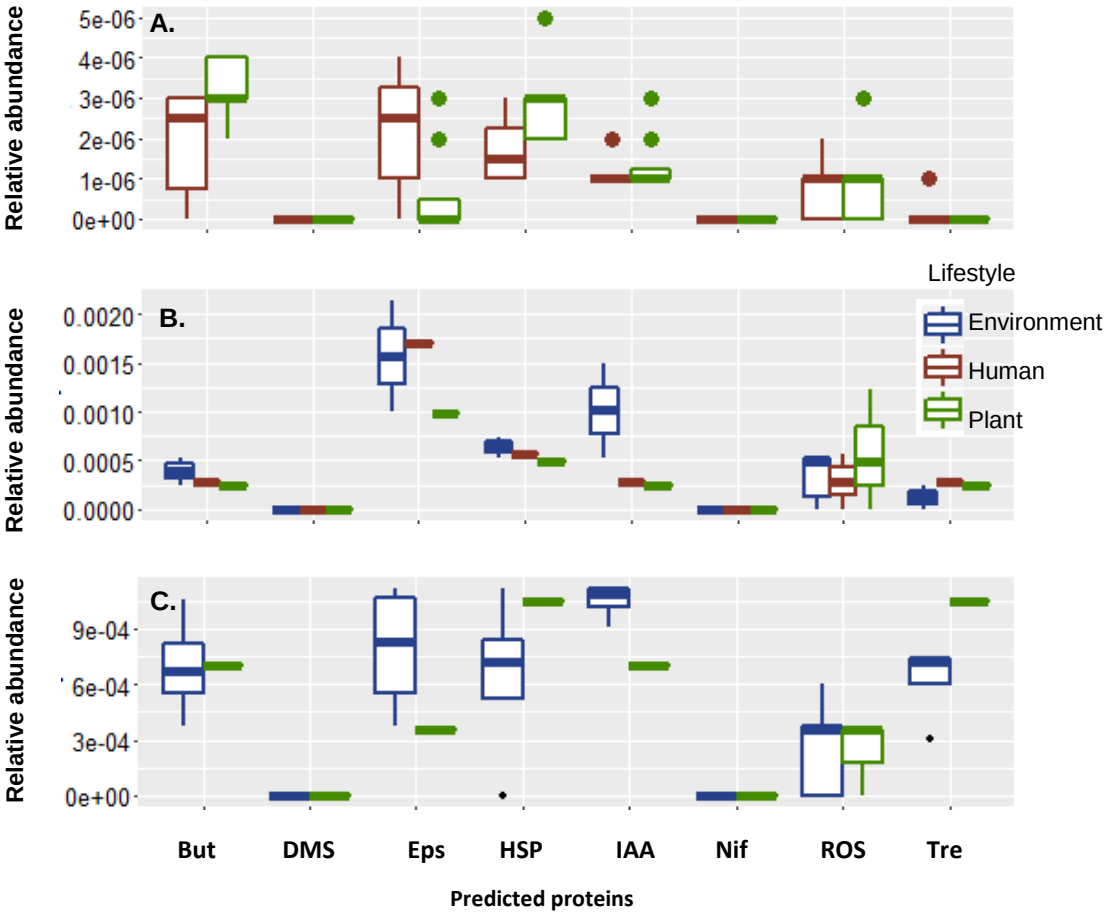


Figure R9. Predicted proteins related to plant growth promotion traits in the genome *Agrococcus*, *Staphylococcus* and *Paenisporosarcina* strains. Strains are clustered based on their lifestyles. But: Butanediol biosynthesis, DMS: dimethyl sulfide, Eps: exopolysaccharide excretion and biosynthesis, HSP: heat shock proteins. IAA: indole acetic acid biosynthesis. Nif: nitrogen fixation. ROS; reactive oxygen species degradation. Tre: trehalose biosynthesis. A. *Staphylococcus*, B. *Paenisporosarcina*, C. *Agrococcus*.

Even though for seed borne *Agrococcus* and *Paenisporosarcina* genomes there was only one genome available, we detected some trends in the protein predictions (Figure R9). Exopolysaccharide biosynthesis and excretion genes tend to be more abundant in environmental and human associated bacteria, indole acetic acid biosynthesis genes are more abundant in environmental bacteria. Seed-borne *A. baldri* also had higher trehalose biosynthesis genes that environmental species. None of the strains possess Nitrogen fixation nor dimethyl sulfide biosynthesis genes.

- Enriched genes in seed borne bacteria.

We used edgeR (Robinson et al. 2009) package of R software described by Bermudez (2016) to unravel significant differences between seed-borne bacterial genomes of other lifestyles. Only the genus *Staphylococcus* was appropriate for this analysis because more than one plant, seed and human associated genomes were available.

This analysis was performed based on gene counts of pfam domains and it showed that *Staphylococcus* genomes were very similar regardless of their lifestyle because no domains were differentially enriched between seed and plant associated bacteria and only few domains were enriched and reduced between plant and human associated bacteria. Only pfam01609 a transposase with DDE domain was enriched in plant strains. Pfam02876 (a staphylococcal toxin with beta-grasp domain) appeared to be enriched in human associated strains, but only *S. aureus* KPL1845 had many copies of it; pfam05031, an iron transport associated domain, was also enriched in these strains (Figure R10, Table R2). This analysis was not performed on *Agrococcus* and *Paenisporosarcina* genomes because of the lack of representability of seed-borne and plant-associated bacterial genomes.

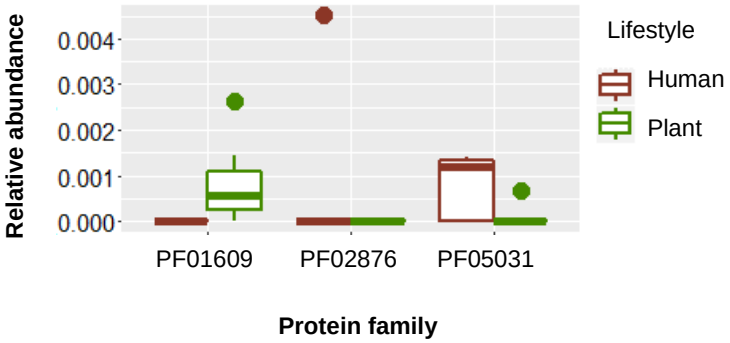


Figure R10. Enriched and reduced protein families between plant and human *Staphylococcus* strains. Strains are clustered based on their lifestyles

Table R2. Enrichment analysis of protein families in *Staphylococcus* strains

Pfam	logFC	logCPM	LR	PValue	FDR	Description
pfam02876	-3.91	10	17.5	2.92E-05	0.0485	Staphylococcal/Streptococcal toxin, beta-grasp domain
pfam05031	-2.99	10.2	15.9	6.58E-05	0.0485	Iron Transport-associated domain
pfam01609	3.91	10	16.4	5.07E-05	0.0485	Transposase DDE domain

- Accessory genome of seed-borne strains.

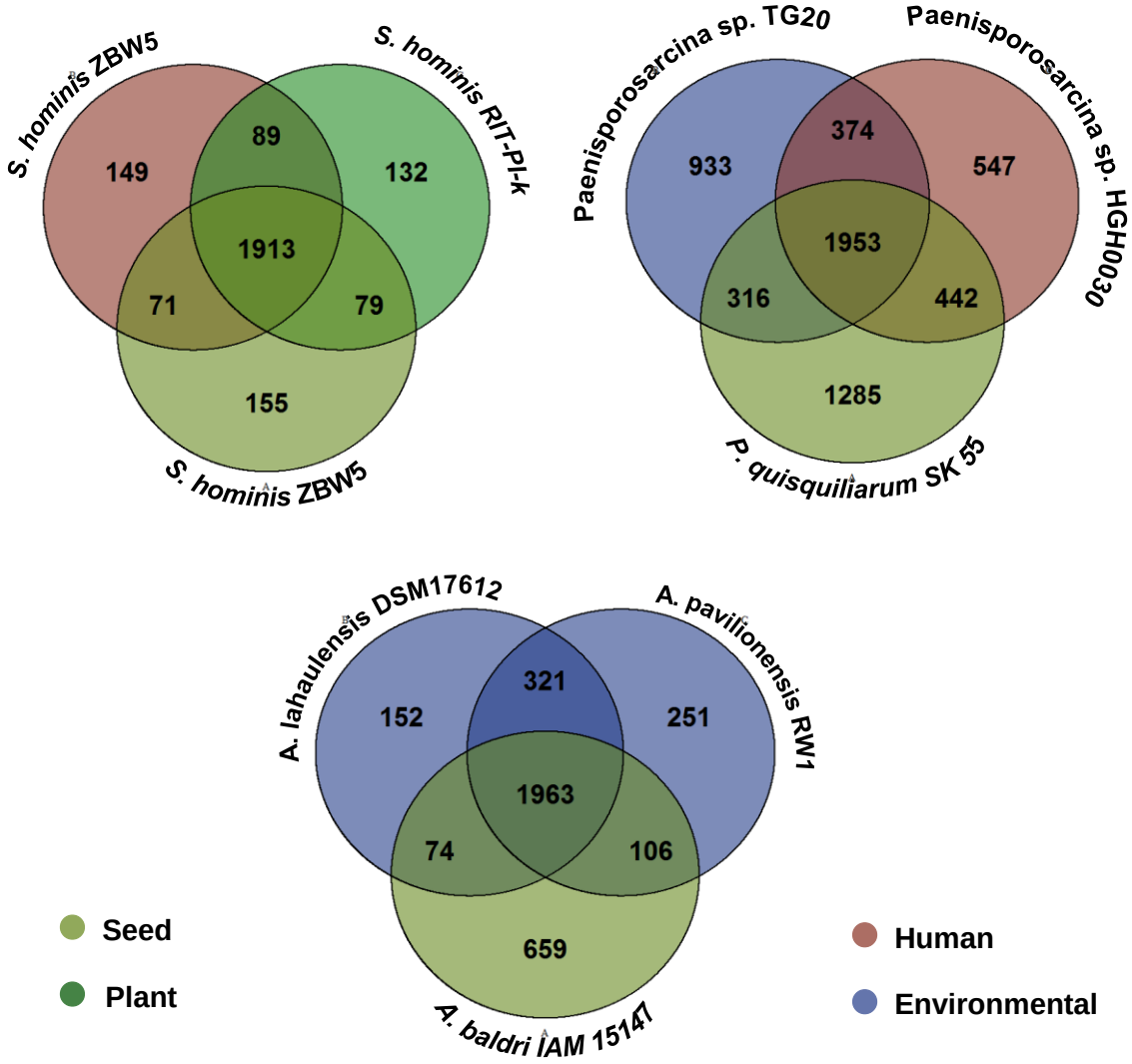


Figure R11. Comparative genomics between cacti seed-borne bacteria and their human, plant and environmental relatives. Venn diagrams represent shared (core) and unique (accessory) genes of each strain.

Core and accessory genome were determine based on the presence and absence of orthologous genes between strains from different lifestyles (Figure R11). Only 6.98% of seed-borne *S. hominis* genes did not have an orthologue in the human and environmental strains and they shared a mean of 86.24% of their genes. It was unclear if *Paenisporosarcina* strains are from the same species, but they only shared an overall of 53.81% of their genes and 32.15% being unique for the seed-borne strain. *Agrococcus* strains are completely different species and they shared a mean of 74.04% of their genes an only 23.51% of them are unique to the seed-borne strain *A. baldri*.

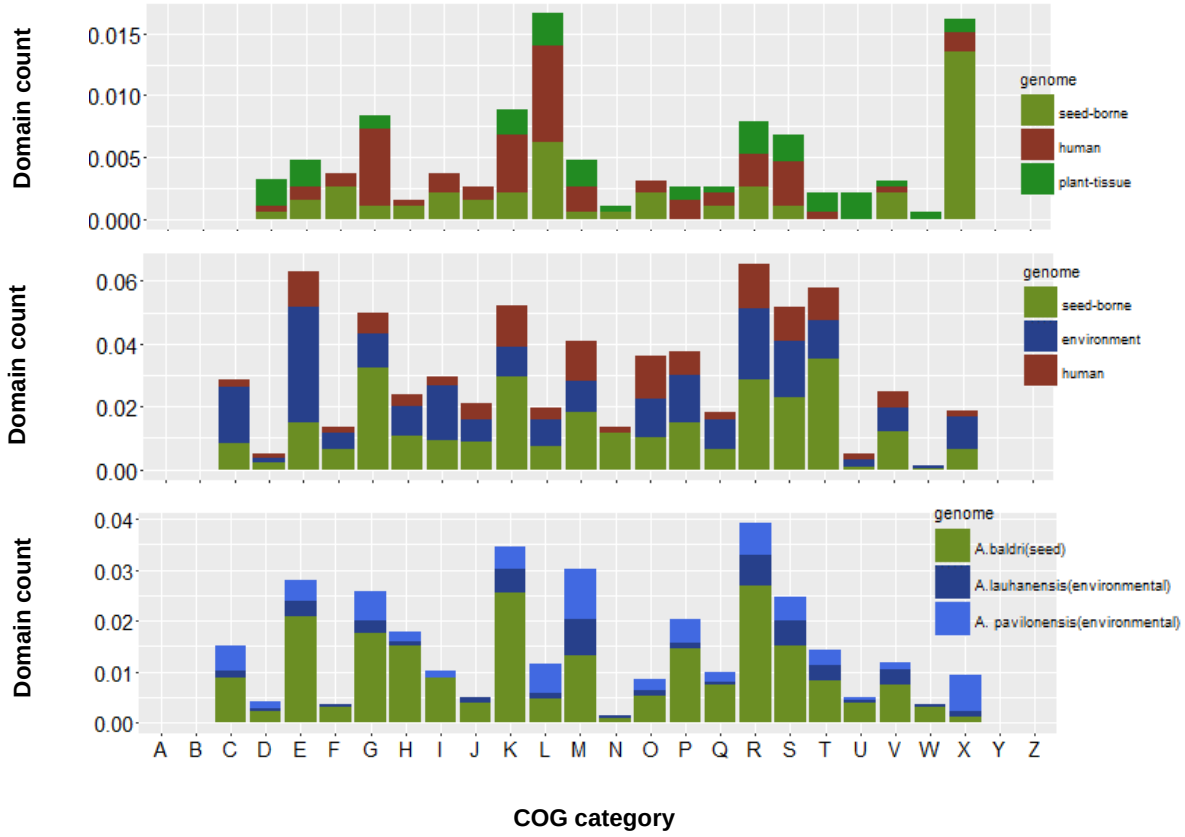


Figure R11. Cog categories of accessory genes in the genome of seed-borne bacteria. Bars represents the abundance of the domain count of each category relative to the total number of domains assigned

In Figure R12, COG domains of the accessory genome of strains were clustered in categories based on their functions. As expected, there were no categories related to RNA modification (A), chromatin (B), nuclei (Y) and cytoskeleton (Z) in bacterial genomes, domains of general function prediction (R) and function unknown (S) will be ignored.

The 3 most abundant categories of unique genes in seed-borne *P. quisquiliarum* are genes related to carbohydrate metabolism (G), signal transduction (T) and translation (K). For *A. baldri* the most abundant categories were also G, K and aminoacid transport and metabolism (E). Abundant categories for seed-borne *S. hominis* were related to replication, recombination and repair (L), nucleotide metabolism (F) and mobilome (X), the last one due to the presence of unique genes with transposase domains that interestingly were also enriched in the edgeR analysis (Table R2).

Table R3. Unique COG domains in accessory genome of seed-borne *S. hominis* ZWB5

COG_ID	Gen name	Product	Cat.	Count
COG0140	HisI2	Phosphoribosyl-ATP pyrophosphohydrolase	E	1
COG0856	PyrE2	Orotate phosphoribosyltransferase homolog	F	3
COG1691	COG1691	NCAIR mutase (PurE)-related protein	F	1
COG1440	CelA	Phosphotransferase system cellobiose-specific component IIB	G	1
COG1455	CelB	Phosphotransferase system cellobiose-specific component IIC	G	1
COG2091	Sfp	Phosphopantetheinyl transferase	H	1
COG4694	RloC	Wobble nucleotide-excising tRNase	J	1
COG0338	Dam	Site-specific DNA-adenine methylase	L	1
COG3392	COG3392	Adenine-specific DNA methylase	L	1
COG3449	SbmC	DNA gyrase inhibitor GyrI	L	1
COG3593	YbjD	Predicted ATP-dependent endonuclease of the OLD family, contains P-loop ATPase and TOPRIM domains	L	1
COG0582	XerC	Integrase	LX	1
COG3875	LarA	Nickel-dependent lactate racemase	M	1
COG1192	BcsQ	Cellulose biosynthesis protein BcsQ	N	1
COG2124	CypX	Cytochrome P450	QV	1
COG1606	COG1606	ATP-utilizing enzyme, PP-loop superfamily	R	1
COG3391	YncE	DNA-binding beta-propeller fold protein YncE	R	1
COG1641	COG1641	Uncharacterized conserved protein, DUF111 family	S	1
COG4815	COG4815	Uncharacterized protein	S	1
COG1401	McrB	5-methylcytosine-specific restriction endonuclease McrBC, GTP-binding regulatory subunit McrB	V	1
COG3440	COG3440	Predicted restriction endonuclease	V	1
COG2801	Tra5	Transposase InsO and inactivated derivatives	X	6
COG2963	InsE	Transposase and inactivated derivatives	X	5
COG3666	COG3666	Transposase	X	12
COG5527	COG5527	Protein involved in initiation of plasmid replication	X	1

Highlighted rows represent COG domains shared with plant *S. hominis* RIT-PI-k

The accessory genome (genes with no orthologues in other genome) of seed-borne *S. hominis* was analyzed in detail due to its great performance on *A. thaliana.* Table R3 shows unique protein domains (not present in human *S. hominis*) in the accessory genome of seed-borne *S. hominis.* Cacti and grass *S. hominis* possessed unique PTS cellobiose-specific component domains involved in cellobiose transport. Seed-borne *S. hominis* possessed domains involved in the biosynthesis of many natural products such as the cytochrome P450, a cellulose biosynthesis protein involved in biofilm biosynthesis, four different genes involved in replication, recombination and repair as well 3 different unique domains of transposases with multiple copies both in plant and seed-borne strains.

The three cacti seed-borne isolates, regardless of their taxonomy, shared accessory genes (without orthologues in other isolates) among them (Figure R13). Seven COG domains were shared between them, these domains are mainly related to DNA, binding, replication and recombination (COG1961, COG3391), transcription (COG0583, COG1396), translation (COG1670, COG0456).

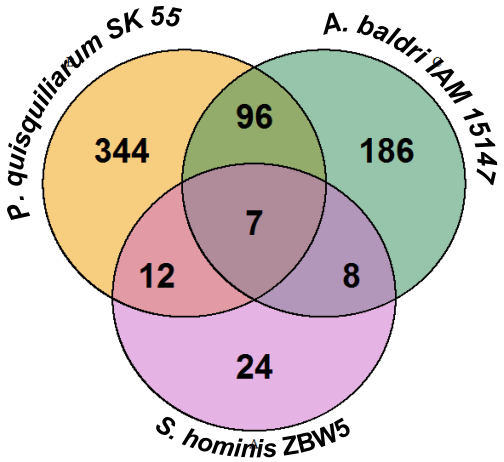


Figure R13. Accessory genes in seed-borne bacterial genomes. Venn diagram represent the accessory genes that are unique and shared between strains.

Table rR4. Shared accessory genes in seed-borne bacterial genomes

COG_ID	Description	Category
COG3391	DNA-binding beta-propeller fold protein YncE	R
COG0583	DNA-binding transcriptional regulator, LysR family	K
COG0561	Hydroxymethylpyrimidine pyrophosphatase and other HAD family phosphatases	H
COG1670	Protein N-acetyltransferase, RimJ/RimL family	JO
COG0456	Ribosomal protein S18 acetylase RimI and related acetyltransferases	J
COG1961	Site-specific DNA recombinase related to the DNA invertase Pin	L
COG1396	Transcriptional regulator, contains XRE-family HTH domain	K

VIII. DISCUSSION

Microbiome research in animals and plants has increased in the past few years due to the recognition of the impact of microbial symbionts in host development, fitness, reproduction and evolution, and due to the improvement of *omic* technologies to study them. Holobionts are highly complex entities and interdisciplinary approaches must be considered to reach an accurate picture of the composition, function and dynamics of their microbial communities. These integrative approaches should account for different levels of regulation, from ecosystems to single genes, using high-throughput sequencing data and functional assays to test hypotheses.

COMMUNITY COMPOSITION OF CULTIVABLE AND NON-CULTIVABLE SEED-BORNE ENDOPHYTES

We aimed to assess the composition of cacti seed-borne endophytes. 16S rRNA gene libraries of cacti seeds were prepared together with maize and sorghum libraries from a co-current project in the laboratory of Microbial Interactions. Unfortunately libraries weren't of the required quality to be sequenced and they must be remade.

Community composition of cacti microbiome was driven by the plant compartment and not by geography, ~90% of OTUs are shared across plant species (Fonseca-García et al. 2016), but assemble rules of seed endophytes might be different because they are subjected to transgenerational transmission. Current studies on *Phaseolus vulgaris* (Klaedtke et al. 2015) and the *Brassicaceae* family (Barret et al. 2015) suggest that geography and temporality are the main factors driving microbial community composition of seeds. Additionally, some studies suggest that seed endophytes are selectively recruited in function of environmental factors for example: Cadmium (Cd)-exposure in *A. thaliana* induces the transmission of Cd resistant and ACC deaminase producing bacteria (Truyens et al. 2016; Truyens et al. 2013). Based on the evidence mentioned above, we would expect the bacterial seed-borne endophytic communities of cacti to be driven by the sampling site in function of edaphic and environmental factors that are very contrasting between Magueyal and San Felipe in organic matter, phosphorous content and iron content (Table M2), and annual precipitation (Fonseca-García et al. 2016). We probably would have also expect fungi not to be present in the seeds of cacti, as we were unable

to culture any fungus from sterilized seeds and seedlings, but analyses of the ITS2 sequencing should help validate this.

Host species and host genotype influence the assemblage of plant microbial communities (Wagner et al. 2016). To test if this was also the case in seeds of cacti, our experimental design included 3 different clades of the cacti phylogeny and 4 different species to determine if phylogenetic relatedness affect vertically transmitted endophytes or if they follow a niche-based assemblage. If cacti species were the main driver of the microbial community composition in cacti seeds, we would expect them to be similar between closely related species (e.g. both *Opuntia* species) regardless of the sample site and even find trends of phylosymbiosis where microbial communities change reflecting the phylogeny of its host (Bordenstein & Theis 2015).

The age of the fruits sampled was not taken into consideration for most of the cacti species (only for *C. imbricata*). We are aware that this might increase the variation between individuals if we consider that other processes can affect the assemblage of microbial communities, for example: the assembly story (e.g. time and order of colonization) can result in the divergence of local communities under similar environments (Nemergut et al. 2013).

Isolation frequency of bacteria and fungi of cacti seeds was very low and only 4 different bacterial morphotypes were recovered. Most of the studies that isolated seed-borne endophytic bacteria recovered a very small number of bacteria of few taxa (reviewed in: Truyens et al. 2015). The isolation of seed-borne endophytes is difficult because they are from a very specific habitat and their abundance is very variable (Truyens et al. 2015). Interestingly, we were not able to count bacteria on plates because we did not have enough seeds and/or because bacterial community was not abundant. Puente et al. (2009) reported that cardon seeds harbor a population of 10^6 cfu/g, but it is unclear if the seeds were surface disinfected or not. Other reports agreed that seeds can harbor a great amount of bacterial endophytes such as rice ($3.5x10^5$ cfu/g) (Hardoim et al. 2012).

Only the genera *Bacillus* and *Paenibacillus* were isolated from cacti seeds and seedlings, many of them from Winogradsky media suggesting that they might be Nitrogen fixing bacteria. Together with the strains isolated by Fonseca-García et al. (2016), the collection comprises three phyla, *Firmicutes*, *Proteobacteria* and *Actinobacteria*, which is consistent with previous works on other plants (reviewed in: Truyens et al. 2015). Most of the seed-

borne strains in our collection are endospore-forming *Firmicutes* (*Bacillus, Paenibacillus, Paenisporosarcina* and *Psychrobacillus*), also *N. prasina* is an actinobacteria that forms spores and this trait might help bacteria survive the seed environment that is characterized by high osmotic pressures because of the loss of water and the accumulation of sugars during seed germination (Mano et al. 2006).

Psychrobacillus and *Paenisporosarcina* strains are known for being psychrophilic and some have been isolated from glaciers (Koh et al. 2012; Reddy et al. 2013; Krishnamurthi et al. 2010). Glaciers can also be considered arid environments as they experience low water availability, and high periodic solar radiation (Bej et al. 2009) suggesting that these bacteria might have traits to survive in both environments.

Leclercia sp. has only been described as an opportunistic human pathogen, *Leclecia adecarboxylata* (Anuradha 2014), while *Agrococcus* have been isolated air, soil, cheese, medieval paintings and the phyllosphere of potato (Parte 2014), being this the first time reported on cacti.

S. hominis represents an interesting strain because *Staphylococcus* genus is a commensal and pathogen of animals including humans (Coates et al. 2014), but it has also been frequently reported as a seed endophyte (Truyens et al. 2015) in *A. thaliana* (Truyens et al. 2013)*, P. pringlei* (cardon cactus) (Puente et al. 2009a) and rice (Midha et al. 2016). We confirmed its presence in *O. robusta* seedlings using DGGE (Figure C2), this means that this strains has been transmitted through the seed and kept during the first stage of cacti development.

We attempted to determine if seed-borne strains were on the cacti microbiome data by matching its 16S rRNA partial sequences with the OTU sequences from the data base as made by Fonseca-García et al. (2016) but using a phylogenetic approach. Only OTU 19 that matched *Bacillus* strains seems able to remain in the cacti stem endosphere and proliferate (~30% of relative abundance in stem endosphere), while others, such as OTU4 and OTU99, remain only in relatively low abundance (~5% of relative abundance). This finding suggests that *Bacillus* could be an important member off the cacti microbiome that is necessary during many stages of cacti development , an observation that is consistent with the idea that microbial amplification is an important source of variation of the hologenome and can be transmitted to the offspring (Rosenberg & Zilber-Rosenberg 2016). As showed by Hardoim et al. (2012) seed-transmitted endophytes can also be

present in other plant compartments like the rhizosphere, these OTUs (e. g. OTU 29) were present in low abundance in rhizosphere, root endosphere, phyllosphere and bulk soil suggesting that they might be also adapted to live inside cacti. Nevertheless, these observations also suggest that seed-transmitted bacteria could be derived from soil or other plant compartments.

Fonseca-García et al. (2016) could not identify a matching OTU for *Psychrobacillus*/*Paenisporosarcina* strains. Using our approach, we found that OTU 29 (*Paenisporosarcina*) match the 16S rRNA sequences of these strains, most of them with more than 97.0% similarity. OTU 2692 (*Curtobacterium*) matched the 16S rRNA sequence of *A. baldri* with 76.9% similarity which indicates that this strains is a different OTU not present in the cacti microbiome. OTU 29 is found in low abundant (median ~0.4% of relative abundance) but present nearly in all the samples, while OTU 2692, 680 and 1492 (*Paenibacillus*) are very infrequent and not abundant, this evidence suggests that some seed strains might be worth to keep only during the first stages of cacti development but not after. Research on cacti microbiome dynamics and transmission through time are necessary to support this evidence, although they are very challenging to perform because they develop very slowly.

CONTRIBUTION OF SEED-BORNE STRAINS TO PLANT FITNESS

Functional assays *in planta* were meant to be performed in seedlings of *O. robusta*. Unfortunately, they had been maintained *in vitro* for 3 years approximately and when we tried to establish them on soil, they could not survive (data not shown). We noticed that *in vitro* shoots have deformations and less and smaller roots that are symptoms of hyperhydratation. This disorder is a consequence of *in vitro* culture conditions such as high humidity, excess carbohydrates and minerals and low light intensity, also it is responsible of its impediment to grow on soil (Pérez-Molphe-Balch et al. 2002)

We decided to perform this assays in the plant model *A. thaliana* that allowed us to assess the contribution of seed-borne strains on plant fitness under low water availability in a controlled environment, fast and with many replicates. We are aware that the outcome of host-microbe interactions depend is context dependent as a function of abiotic and biotic factors occurring during the interaction (Partida-Martínez & Heil 2011; Chamberlain et al. 2014), this means that we are changing the context by using a model plant on in vitro

culture and the seed-borne strains might not have the same effect on *A. thaliana* as they have in cacti. Although if bacterial mechanism of plant growth promotion or stress resistance are conserved across different plant species we would be able study the molecular mechanisms involved in plant bacteria interactions in future research.

van der Weele et al. (2000) developed a reproducible methodology to study hydric stress on *A. thaliana* using infused PEG 8000 to lower the water potential of MS media. *A. thaliana* growing in a moderately low water potential (-0.5 MPa) reduced root length, lateral root number, leaf surface, biomass, germination rate, water content and the presence of root hairs (data not shown). The overall effect was consistent with many other works (van der Weele et al. 2000; Verslues & Bray 2004; Vallejo et al. 2010), although, there are difference regarding intensity because we germinated the seeds directly on PEG infused-plates (Verslues et al. 2006). *A. thaliana* phenotype under hydric stress is the result of the overall decline of photosynthesis, the reduction of water use efficiency, increase of damaging reactive oxygen species (ROS) and ethylene among other consequences that plants experience under drought stress (Farooq et al. 2012)

- Effect of cacti seed-borne bacteria on *A. thaliana* under hydric stress

Eight strains were selected based on their taxonomy and the presence of mechanisms of plant growth promotion and stress resistance to perform in planta assays. Most of the strains were drought tolerant and some produced exopolysaccharides (Fonseca-García et al. 2016) which might improve their capacity to absorb water during desiccation as suggested by Roberson & Firestone (1992).Three seed-borne strains (*Paenibacillus* sp.*, S. hominis and N. prasina*) increased seed germination rate under hydric stress, therefore these strains are recruited from cacti microbiome, they survive the seed environment and might support or regulate seed germination in arid environments. We could not detect any improvement in seed germination under standard growth conditions, probably, because it was very fast or the effect was quite small. Never so, it also suggests that the increase in seed germination rate might be stress-dependent and induced by hydric stress imposed by the PEG-infused media.

The effect of bacteria on seed germination has been poorly studied, for example: *Sphingomonas* sp. and *Mycobacterium* sp. increased germination of *Dendrobium moschatum* orchid (Tsavkelova et al. 2007), while *Azospirillum brasilense* and *Bradyrhizobium japonicum* increased germination of maize and soybean (Cassán et al.

2009). The mechanism underlining this phenotype in both studies was attributed to plant hormones, particularly IAA, but it is still unclear. Germination and dormancy are highly regulated and involve hormones like gibberellic acid (GA), abscisic acid (ABA), ethylene, IAA, brassinosteroids and ethylene. The balance of GA and ABA is very important, ABA induces seed dormancy and inhibits germination, while GA induces this process by inhibiting ABA, inducing hydrolytic enzymes and affecting root growth (reviewed by: Miransari & Smith 2014). Cacti seed-borne *N. prasina* and *S. hominis* can produce IAA in culture, but other assays, such as measuring IAA during germination, using transgenic lines (e. g. GA deficient) or performing a transcriptome analysis should be performed to disentangle the mechanism involved.

Nitrogen compounds like N_2O and NO_3^- can induce seed germination under non stress and stress conditions like salinity (Zheng et al. 2009; Atia et al. 2009), also cytokines can improve germination under salinity, drought and heavy metal stresses (reviewed by: Miransari & Smith 2014). Although, we were able to find in seed-borne *S. hominis a* isopentenyl-diphosphate delta-isomerase and a unique cytochrome P450 (not present in its human nor plants relative) that are known to be involve in cytokine products (Frébort et al. 2011), it possess the respiratory nitrate reductase (might synthetize nitrites) and the assimilatory nitrite reductase (might synthetize ammonia), but it doesn't possess the enzyme require to synthetize nitrites nor nitrous oxide (data not shown).

Some studies have shown that 2,3-butanediol producing bacteria, such as *Bacillus subtilis* GB03 (Zhang et al. 2010) and *Pseudomonas chlororaphis* O6 (Cho et al. 2008; Cho et al. 2012), could improve *A. thaliana* growth and survival under osmotic stress. Contrary, most of our strains did not show any significant effect on *A. thaliana* growth (leaf surface, root length, lateral roots etc.) under low water potential as the studies mentioned above. Even though, seed-borne *S. hominis*, *P. quisquiliarum* and *A. baldri* genomes showed genes related to butanediol production, they might be silent or not expressed in different culture conditions. Inoculation methodology is also a source of variability, Zhang et al. (2010) and Cho et al. (2012) inoculated their strains without contact with the plant on divided petri dishes or directly on the roots of germinated *A. thaliana,* respectively, while we inoculated them on the seeds. Inoculation differences are very clear in *Leclercia* sp. and some *Bacillus* sp. treatments, these bacterial taxa could grow on MS media and invade *A. thaliana* seedlings, but some seedlings remained uninvaded and grew with drastic

morphological effects like the increase of the number of lateral roots (Figure S8) that might be due to the production of volatile or diffusible compounds by the bacteria.

- Growth promotion of *A. thaliana* under standard conditions

Only Firmicutes had a positive effect on *A. thaliana* phenotype under standard water potential, especially in the number of lateral roots. Root architecture and functioning can be modified by plant growth promoting bacteria by releasing hormones and other signals for example microbial IAA can increase the number and length of lateral roots (reviewed by: Vacheron et al. 2013), Interestingly, *S. hominis* possessed genes related to IAA biosynthesis like the indole pyruvate decarboxylase (data not shown) and it indeed produced IAA *in vitro* as many other strains (Fonseca-García et al. 2016), but more detailed analysis should be performed.

S. hominis was the seed-borne strain with the best performance *in planta.* As *Staphylococcus,* many other bacteria have bivalent lifestyles with human and plants, for example *Stenotrophomonas, Pseudomonas, Enterobacter, Ochrobactrum, Burkholderia, Ralstonia* and *Herbaspirillum* (Berg et al. 2005). The genomes of some plant *Staphylococcus* have been sequenced, but there are limited studies demonstrating their performance on plants. Puente et al. (2009) inoculated cardon cactus growing in rocky substrates with a seed-borne *Staphylococcus*, where it increased the volume, dry weight, root length, and height of one year old cacti. Our strain increased lateral root number, dry weight and leaf surface of *A. thaliana*, and it didn't grow on MS media suggesting that it can interact in the surface of the plant or colonize its endosphere because it produced Carboxymethyl cellulase *in vitro* (Fonseca-García et al. 2016)

These findings suggest that seed-borne *S. hominis* - *A .thaliana* interaction could be a good model for testing the mechanism involved in plant growth promotion and germination under stress and as a model for bacterial evolution regarding its lifestyles. Finally, some seed-borne strains might be good candidates to produce biofertilizers for desert-farming because they would accelerate the germination of seeds.

GENOMICS OF A SEED-BORNE LIFESTYLE

-Vertically vs. horizontally seed-borne bacteria.

The mode of transmission of symbionts impacts their dispersal, persistence and genome features (Ebert 2013). The most striking genome changes are those found in obligate and vertically transmitted symbionts such as insect symbionts like *Buchnera* sp. in aphids (Shigenobu et al. 2000) or fungal symbionts like *Burkholderia rhizoxinica* (Lackner et al. 2011). Seed-borne bacteria of cacti are vertically transmitted because they are inherited directly from parents to offspring through the seeds. These strains did not show the characteristic genomic features of obligate symbionts, such as the reduction of genome size and G-C content (Martínez-Cano et al. 2015), suggesting that this symbiosis is not obligated (at least for the bacteria) or that it might be a recent event. The first explanation is consistent with the idea that our strains are present in other plant compartments and soil, thus there might be a horizontally transmitted and/or free-living population that could be recruited to overcome genomic bottlenecks during vertical transmission (Bright & Bulgheresi 2010).

-Differential traits between lifestyles.

Comparative genomics of bacterial symbionts of plants revealed that bacteria with endophytic lifestyles show an enrichment in different functions such as: motility and chemotaxis, signal transduction (e.g. for nitrogen fixation, redox response), transcriptional regulators (e.g. nitrogen assimilation, carbon storage), detoxification and stress-related enzymes (e.g. catalase, glutathione oxidase), transporters (e.g. PTS system), secretion systems and proteins involved in plant growth promotion (e.g. nitrogen fixation, ACC deaminase) (Hardoim et al. 2015). In this context, we aimed to determine the genomic differences between seed (plant)-associated bacteria from their relatives with other lifestyles that might reflect an adaptive trait for the bacteria to its host or for the holobiont to its environment. We focused the data analysis on seed-borne *S. hominis* genome because of the novelty, its great performance in *A. thaliana* and the availability of genomic data from different strains.

We searched for proteins related to plat growth promotion (IAA biosynthesis, ACC deaminase, Nitrogen fixation) and plant stress resistance (Exopolysaccharide biosynthesis reactive oxygen species degrading enzymes etc.) in the genomes of plant (cacti, rice, grass, poison ivy) and human-associated *Staphylococcus* spp. We considered seed-borne strains as plant-associated bacteria for this analysis. Both plant and human *Staphylococcus* possessed plant growth promotion and stress resistance traits implying that these are not specific of a certain lifestyle. The genetic determinants that define a

plant-microbial interaction might be very similar to the ones that define an animal-microbial interaction, for example, human pathogen *Stenotrophomonas maltophilia* shared host-invasion and abiotic resistance with plant-associated *S. maltophilia* and *S. rhizophila* (Alavi et al. 2014) while plant pathogen *Erwinia carotovora* subsp. *atroseptica* shared a core of functions for energy metabolism, motility, cell division and chemotaxis (Toth et al. 2006).

We found tendencies regarding the abundance of some predicted proteins. Genes for exopolysaccharide biosynthesis and secretion tend to be more abundant in human-associated *Staphylococcus* because biofilm production in this genus is important for the establishment of infection (Begun et al. 2007), but they are also relevant for plant-bacteria interactions like *E. carotovora* (Toth et al. 2006). Interestingly HSP seemed to be more abundant in plant-associated strains contrary to what was described Alavi et al. (2014) that did not find heat shock genes in *S. rhizophila* explaining its inability to growth at 37°C. Whether or not this heat shock proteins of plant-*Staphylococcus* are related to its adaptation to seed environment is unclear.

EdgeR analysis revealed very few differentially enriched genes and corroborated that genomes are very similar. Human strains possessed more genes with iron transport domains than plant strains, also the staphylococcal toxin was mainly enriched in the *S. aureus* genome. Only domains of transposases (DDE class) were enriched in plant *Staphylococcus* and surprisingly, seed-borne *S. hominis* had many unique genes (23) with transposase domains without orthologue genes in human *S.hominis.* Transposases are the most abundant genes in nature, they catalyze the insertion of transposable elements (TE) in the genome, also TEs are recognized as major factors for genome evolution (Guérillot et al. 2014; Aziz et al. 2010). Both adaptive and neutral scenarios could explain transposon expansion, the rapid expansion of insertion sequences (TEs that only has a transposase gen) in host-associated bacteria has been related to host restriction, because it decreases the population size and shifts the efficiency of purifying selection. As consequence, more TEs can be fixed by genetic drift (Moran & Plague 2004), this scenario is compatible with the vertical transmission of seed-borne bacteria.

The core and accessory genome of strains were determined based on the orthologue genes they shared (Best Reciprocal Hits). Unfortunately, the genomes of plant (seed), human and environmental *Agrococcus* and *Paenisporosarcina* were of different species. They shared ~70% and ~50% of their genome consistent with the comparison of different enterobacterial plant and animal pathogens that shared three quarters of their CDS (Toth

et al. 2006). This means that the differences we saw between them might be due to their diversification as species so they could overshadow the differences related to their specific lifestyles.

S. hominis strains from different lifestyles (plant, human, seed) shared 85.4% of their genomes (1913 genes) and less than 7% were unique for each. They are very similar in gene content and few genes might reflect specific functions of its unique lifestyle. This is similar to the finding of Chaudhry & Patil (2016), rice seed-borne *Staphylococcus epidermidis* had an average nucleotide identity of 97% against the type strain and a core of 1968 shared genes. They also found a unique genomic region with survival and stress tolerance genes and genes necessary for plant adaptation.

Seed-borne *S. hominis* had only 25 unique and annotated genes (with no orthologous in other genome) and with unique functional domains for plant lifestyle (COG domains not present in the human strain). Regardless of the unique and abundant transposase domains that seed-borne and plant *S. hominis* shared, they also possess PTS type cellobiose transporters that suggest they can hydrolyze cellulose (demonstrated *in vitro* (Fonseca-García et al. 2016)) and incorporate the products of its degradation, this might be an advantage considering the high amount of cellulose in plants. They also contain a unique cellulose biosynthesis domain, but it is surrounded by genes without known function yet (data not shown) and its functionality is doubtful considering that prokaryotic gene groups are composed by functionally related genes (Overbeek et al. 1999). Interestingly a cytochrome P450 gene was found, the first is involved in the biosynthesis of many secondary metabolites that involve oxidation of substrates (Kelly & Kelly 2013). It would be very interesting to study the metabolites that this strain is capable to produce and assess them against seed-borne bacteria and plants.

Interestingly seed-borne *S. hominis*, *P. quisquiliarum*, *A. baldri* didn't share functional domains in their accessory genome that were explicitly related to plant lifestyle, but they shared some categories related to DNA processing, transcription and translation. Most of the other genes were unique to each strain, which suggest that the evolutionary trend is different for bacteria of distant taxa regardless of sharing the same ecological niche.

IX. CONCLUSIONS

Spore-forming and osmotic resistant bacteria are transmitted through the seeds of cacti supporting the idea that these traits are important for their transmission. Cacti seed-borne bacteria can inhabit all plant compartments and soil, but only *S. hominis, Leclercia* sp. and *Bacillus* sp. are highly abundant in the stem endosphere of cacti suggesting their importance to the holobiont.

Seed-borne strains might be helpful for cacti germination and seedling establishment in arid environments as they might affect germination of seeds under drought and increase the number of lateral roots which might increase water and nutrient absorption of plants in arid environments.

There are few genomic differences between seed-borne *S. hominis* bacteria and their plant and human relatives. This strain is enriched in unique transposase domains and had unique genes that might reflect the vertical transmission habit and the adaptation to plant lifestyle. Therefore this strain is a good candidate for the development of biofertilizers for agriculture on arid environments.

X. SIDE PROJECTS

OBTAINING GNOTOBIOTIC CACTI SEEDLINGS

O. robusta seedling were grown on MS medium with and without antibiotics during 1 month (Table C1). Seedlings were aseptically removed from the medium and washed with sterile water. Cacti roots were cut from the aureole with a sterile blade and the length was measured with an electronic caliper. Roots and stems were transferred to sterile Eppendorf and Falcon tubes, respectively, frozen and lyophilized. Root length data was analyzed using R software (R Core Team 2015).

Table C1. Summary of antibiotic treatments

Treatment		Concentration ug/mL				
Ciprofloxacin		5	10	20	40	80
Tetracycline		3	6	12	24	48
Gentamicin		5	10	20	40	80
Mix	**Tetracycline**	1.5	3	6	12	24
	Rifampicin	18.75	37.5	75	150	300
	Carbenicilin	87.5	175	250	500	1000

Genomic DNA was extracted from lyophilized roots (0.0010 ± 0.0005g) and stems (0.03g) using CTAB protocol reported by Edwards (2001). Denaturing gradient gel electrophoresis (DGGE) fingerprinting of 16S rRNA gene fragments (V6-V8) was perform based on the protocol of Desgarennes et al. (2014). 400ng of DNA fragments from shot samples, 18uL of PCR from root samples and a marker of the same amplicon of four different strains isolated from cacti seedlings (*S. hominis* L12*, Bacillus* sp. L8*, Leclercia* sp. L16*, N. prasina* L17) were loaded in four different gels. Band analysis was performed using Bio-Rad Image Lab™ software and analyzed using pheatmap package in R software.

Puente et al. (2009) were the first ones attempting to generate endophyte-free cardon cacti using antibiotics. They showed that cardon seeds treated with antibiotics reduced seed-transmitted endophytes and had an effect in cacti phenotype especially in root length, steam height and dry weight. Although its methodology was particularly questionable because it isn't clear whether the inhibited bacteria came from inside the seed or from its surface since the seeds were not surface disinfected first. Also it is possible that the reduction in the bacterial counts were due to the presence of the antibiotic in the culture media. We weren't able to perform a bacterial count of disinfected seeds nor

aseptic seedlings, so our strategy is based on the fact that fresh cacti cuttings will absorb water from culture media with the antibiotic. If the antibiotics (bacteriostatic and bactericidal) enter cacti tissue they will inhibit bacterial growth and thus reduce bacteria DNA content.

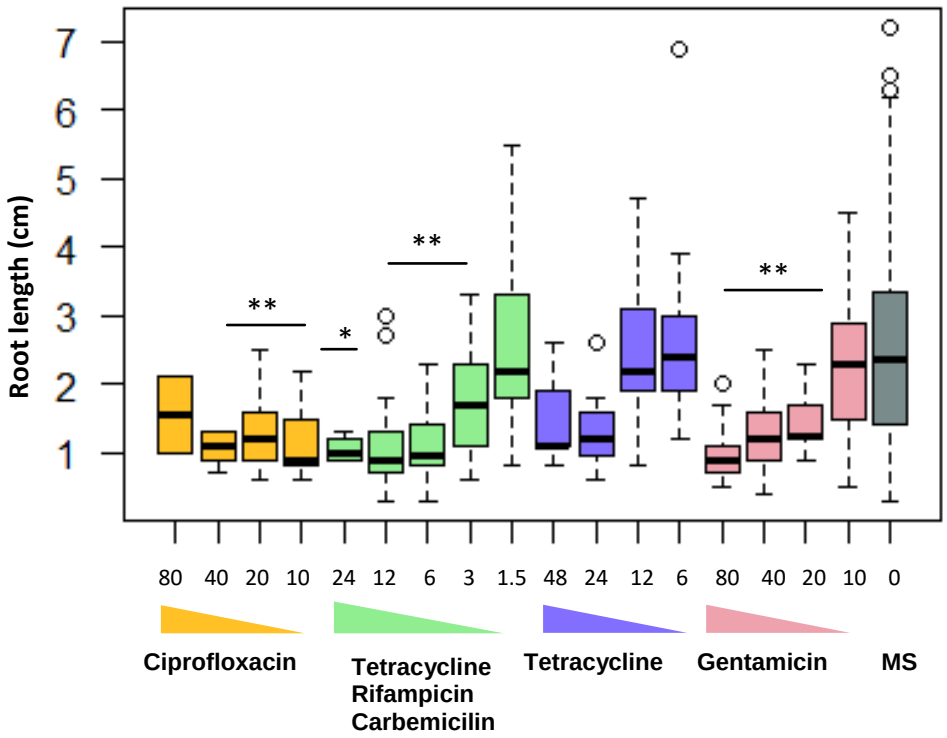


Figure C1. Effect of different concentration of antibiotics on root length of cacti seedlings. Boxes represent the distribution of data, red and blue marks represent the mean and standard deviation. Concentration of Mix treatment is represented as concentration of tetracycline in the mix. P-values are represented with black marks. ** $p<0.001$ * $p<0.01$ against the control in MS (One way ANOVA)

Figure C1 shows that the highest concentration of any antibiotic inhibited rooting, actually many of those roots were aerial that didn't penetrate the agar. Ciprofloxacin treatment was the most toxic for cacti, the other treatments showed a concentration dependent effect in which the lowest concentrations allow a normal rooting similar to the control without antibiotics ($p\approx1$) and the highest concentration inhibited it. Data was highly disperse because cacti roots don't develop at the same time in a single cacti, furthermore some cacti lines (from one single seed) develop longer roots than cacti that are originated from other seed, this is the reason why we choose to randomize cacti in each treatment. In spite of data variability it is clear that antibiotics are toxic for cacti. In fact Antibiotic effects on

plant have been recorded in different species, they can be toxic for them and surprisingly some could promote plant growth (Nickell & Finlay 1954).

DGGE fingerprinting didn't correlate with antibiotic treatments. Some bands that disappeared in certain treatments were present in the highest concentrations of the same antibiotic. It was expected a concentration dependent inhibition, but instead community structure showed a random pattern among treatments (Figure C1 and C2). Cacti cuttings were transferred to MS with antibiotics without roots and it was expected that new roots in antibiotics have its bacterial community inhibited, but DGGE showed the opposite which suggests that antibiotics are not inhibiting root colonization from stems.

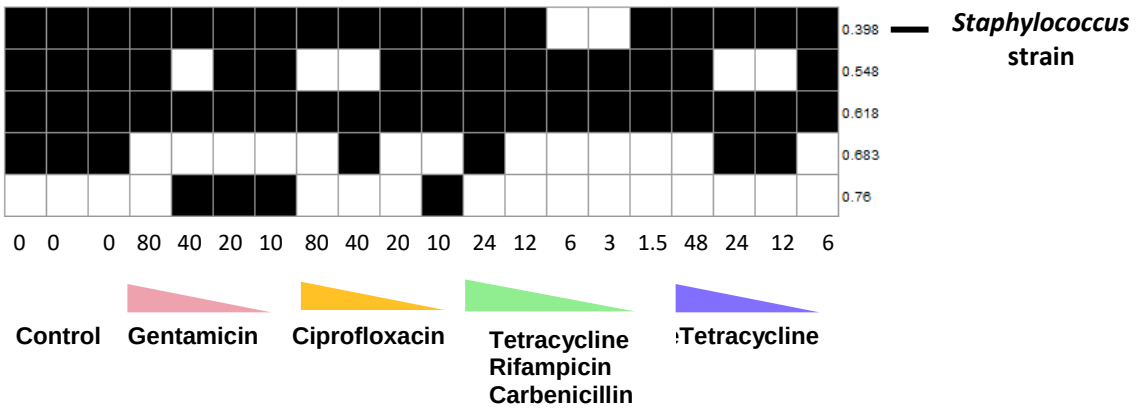


Figure C2. Presence/Absence heatmap of DGGE fingerprinting of the microbial community in cacti stems under antibiotic treatment. The heatmap shows the band fingerprint (relative front in rows) in each treatment (columns). Antibiotic concentrations correspond to the ones in Table S1. MS: control without antibiotic. G: gentamicin. T: tetracycline. M: Mix. C: ciprofloxacin.

Based on the RF of the bands in the DGGE marker we identified 2 different DGGE bands. One band corresponded to seed-borne *S. hominis* L12 and *Leclercia* sp. L16, this finding proved that this strains indeed came from inside the cacti seedling and not from the environment. *Staphylococcus* amplicon was absent or inhibited from two of the Mix treatments, M6 and M3 in shots and M12 in roots (Figure C1 and C2), but *Leclercia* amplicon remain in all the treatments.

In vivo inhibition of seed-borne bacteria was validated with and *in vitro* inhibition assay of seed-borne strains isolated from cacti seedlings. 24h cultures on TSB media were massively streaked on TSA plates. Sterilized filter paper discs were slightly soaked on 2 different concentrations of antibiotics, 1:1 and 1:10 dilution of mixture 1 on Table C1. Three

paper discs for each concentration were placed on two inoculated TSA plates. Cultures were incubated at 28°C during 48 hrs. Finally inhibition halos were measured.

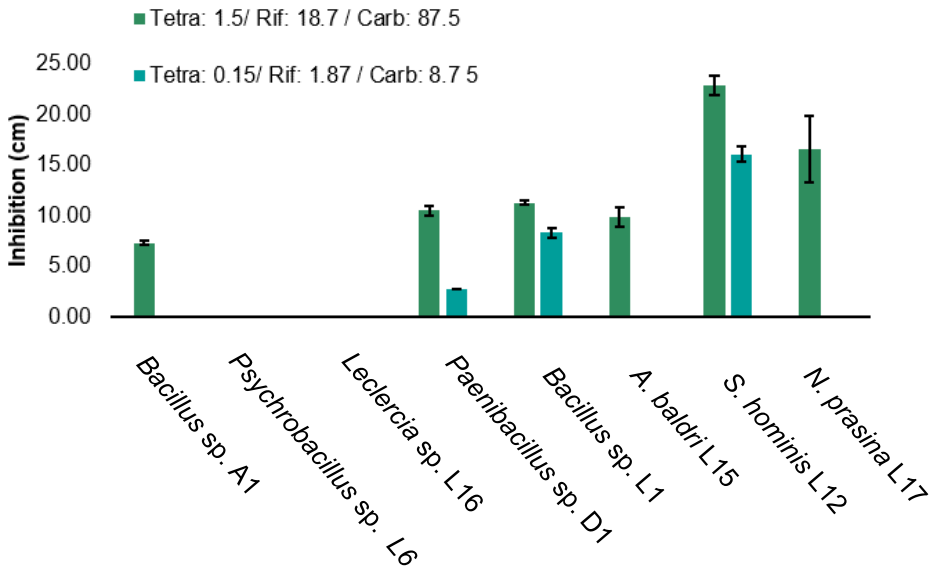


Figure C3. *In vitro* inhibition assay of seed-borne strains with tetracycline , rifampicin and carbenicilin mixture. Bars represent inhibition halo of each strain in two different concentration of the antibiotic mixture.

Interestingly most of the strains were sensible to the antibiotic mixture except for *Leclercia* sp. L16 and *Psychrobacillus* sp. L6 *S. hominis* L12 strain was the most sensitive and was highly inhibited with 10 times less the antibiotic concentration used on the cacti seedlings. This correlated with DGGE of cacti seedlings on antibiotics were the *Staphylococcus* band was inhibited with this mixture of antibiotics, but the *Leclercia* band was always present. Our results suggest that antibiotics inhibit rooting in cacti and cause a dysbiosis that can't be correlated with its phenotype.

FUNCTIONAL ANALYSES OF THE RHIZOSPHERE AND PHYLLOSPHERE OF CAM PLANTS

In the Laboratory of Microbial Interactions, the contribution of microbial symbionts to the plant holobiont's fitness has been evaluated by detecting canonical growth promotion traits in microbial cultures, assessing its effect on plants, studying microbial community composition and assessing genomic features of this symbionts. "Omic" technologies have allowed us to study microbial communities over different levels of regulation (genes, transcripts, proteins, metabolites) which requires not only new computational approaches but interesting study models and creative experimental designs.

To dissect the functional capabilities of cacti and agave microbiome we performed high-throughput metagenome sequencing of the microbial communities from rhizosphere and phyllosphere of *Agave salmiana*, *Agave tequlana*, *M. geometrizans* and *O. robusta* from Magueyal and San Felipe, Gto. Functional profiling allowed us to compare microbial community functionality, while phylogenetic profile allowed us to identify abundant taxa in each community.

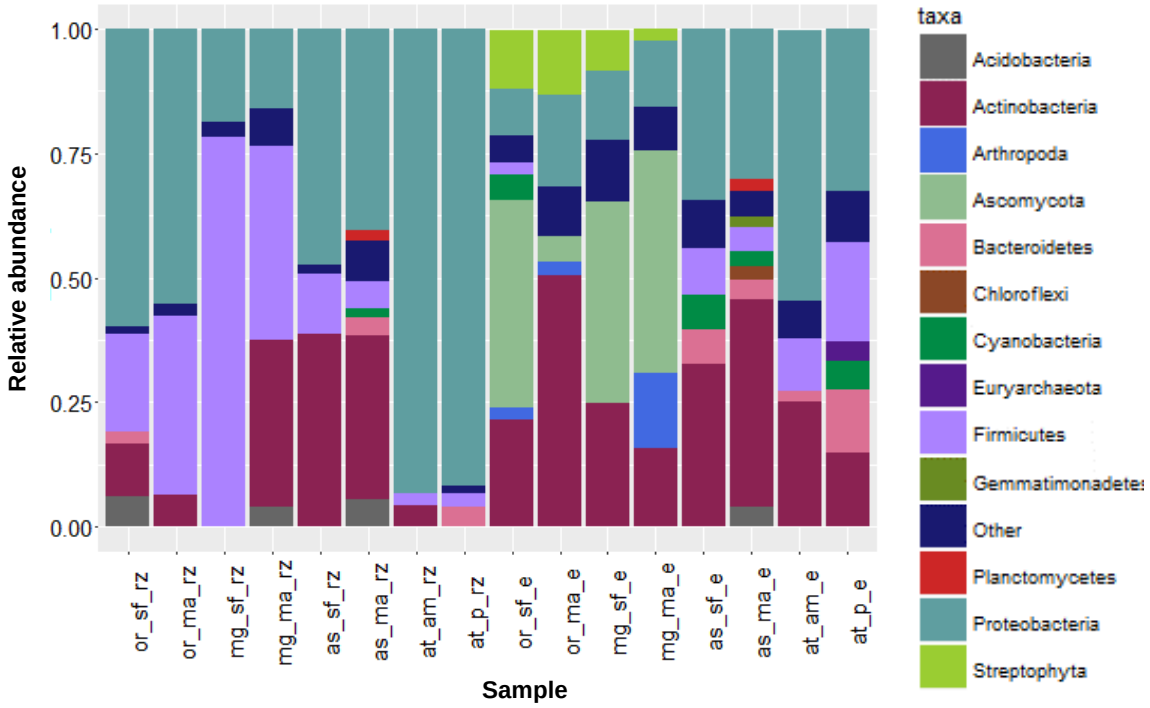


Figure C4. Phylogenetic profile of scaffolds from phyllosphere and rhizosphere sympatric cacti and agave. A threshold of 2% of relative abundance was used. Mg: *M. geometrizans*, Or: *O. robusta*, As: *A. salmiana*, Ma: El Magueyal Gto., Sf. San Felipe Gto., e: phyllosphere, rz: rhizosphere

Interestingly, taxonomic distribution of scaffolds in metagenome data resembles amplicon sequencing data (Figure C4), for example: *Cyanobacteria* scaffolds are more abundant in phyllosphere than in rhizosphere, *Firmicutes* and *Proteobacteria* scaffolds are the most abundant in cacti rhizosphere and *Proteobacteria* was the dominant taxa in both compartments of *A. tequilana.* Although, some other taxa do not correlate with this trend, such as *Actinobacteria* scaffolds that are highly abundant in all phyllosphere samples. Also, eukaryotic sequences were only detected in phyllosphere, especially in cacti were they comprised almost 50% of relative abundance. Most of them where assigned to *Ascomycota* phylum, but other taxonomic groups that are barely considered in metagenomic studies were present such as were detected as *Basidiomycota*, *Arthropoda, etc,*. This findings are important because amplicon sequencing of ssu rRNA introduces bias during the processing steps that masks the true community composition of the samples(Brooks et al. 2015) that might be revealed by metagenomic data.

Each plant compartment is constrained by different environmental conditions and host factors. For example, rhizosphere is pictured as a highly dynamic environment influenced by a complex array of root exudates, in contrast phyllosphere is characterized by severe conditions such as UV radiation desiccation and oligotrophy (Vandenkoornhuyse et al. 2015). In cacti and agave community composition was mostly influenced by the plant compartment, therefor we would expect the functionality in each community to be different as well. Figure C5 shows clustering plots based on gene count of COG group. Most of the samples tend to cluster based on the plant compartment and not based on the sampling site nor the plant species which suggest a functional differentiation between rhizosphere and phyllosphere. Although, the similarity between the 2 main groups is very high (>70%) and some rhizosphere samples cluster within the phyllosphere group. These similarities might be due to a taxonomic and functional overlap between plant compartments, especially the ones that are necessary for essential cellular processes shared among microbial communities. This is consistent with *Arabidopsis thaliana* microbiome where phyllosphere and rhizosphere strains showed limited functional differentiation based on its ecological niche with only few differences among specific groups (Bai et al. 2015). In contrast by using a metaproteogenomic approach of the rice microbiome, Knief et al. (2012). found a clear differentiation between both compartments regarding stress response, Carbone utilization and Nitrogen fixation

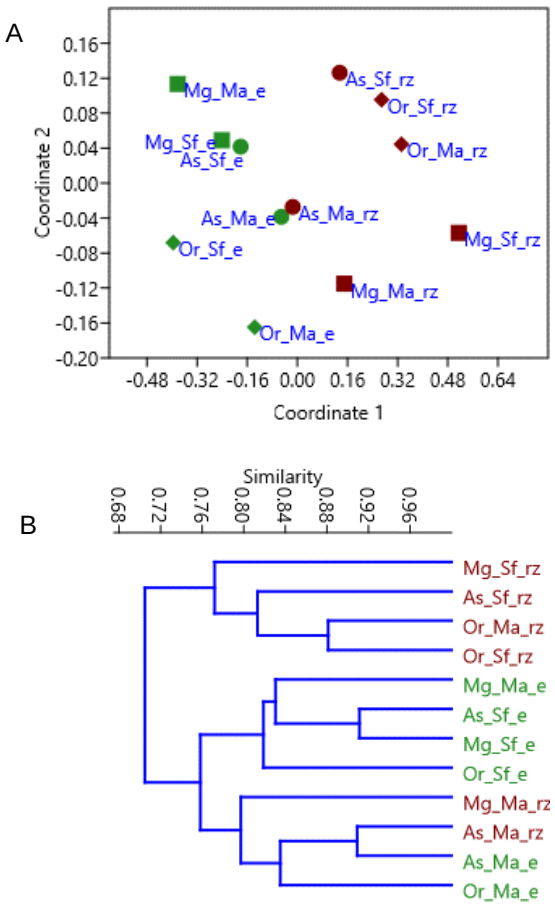


Figure C5. Clustering analysis based on Bray-Curtis dissimilarities of COG gene annotation of metagenomes from phyllosphere and rhizosphere sympatric cacti and agave. A) NMDS clustering. B) UPGMA clustering. Mg: *M. geometrizans*, Or: *O. robusta*, As: *A. salmiana*, Ma: El Magueyal Gto., Sf. San Felipe Gto., e: phyllosphere, rz: rhizosphere

.

To dissect the important functions between phyllosphere and rhizosphere we assessed if there where enrichments of genes that due to their abundance might be important for the functioning of each compartment. We address this problem as (Bermúdez-Barrientos 2016) where he compared protein domains among a set of genomes to assess for reduction or expansion of functions within a statistical framework. By comparing gene counts of COG groups between different groups of samples we found that phyllospheres are very similar between *O. robusta* and *M. geometrizans* with only 13 and 18 expanded COG groups, respectively as well its rhizospheres with 3 and 0 expanded groups, respectively. Contrary, by comparing cacti rhizosphere against phyllosphere we obtained 158 and 113 differentially expanded COG groups, respectively, which supports the idea that both compartments are functionally different.

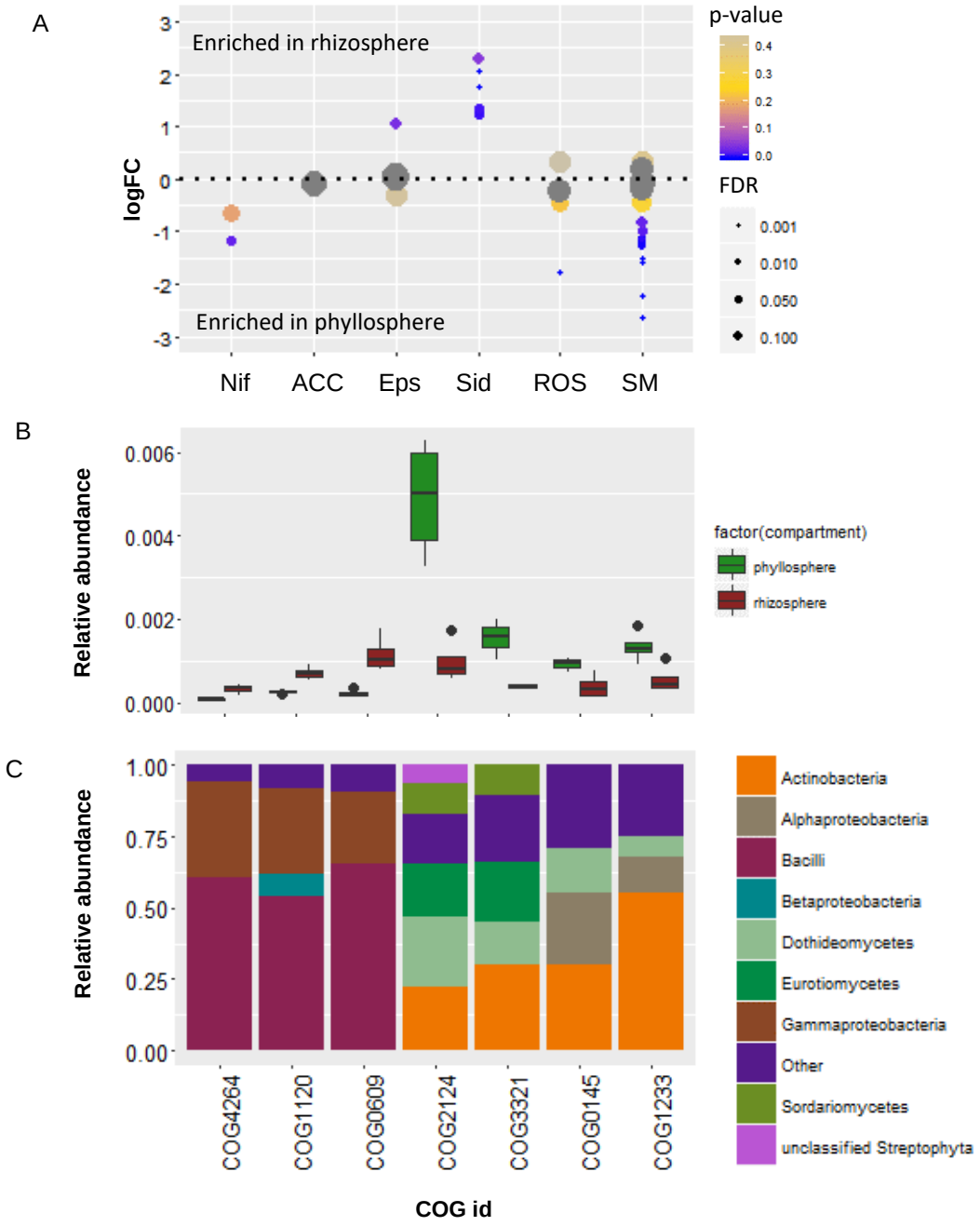


Figure C6. Relevant enriched functions in phyllosphere and rhizosphere of cacti. A) Dot plot of the fold-change, FDR and p-value of COG groups related to: Nif: nitrogen fixation; ACC: ACC deaminase, Eps: exopolysaccharide production, Sid: siderophore production, ROS: ROS degradation, SM: secondary metabolism. B) Relative abundance of genes with enriched COG groups. C) Taxa assignment of scaffold with enriched COG groups that account for more than 3% of relative abundance. **COG4264:** Siderophore synthetase component. **COG1120 and COG0609:** ABC-type cobalamin/Fe3+-siderophores transport system, ATPase and permease component, respectively. **COG2124:** Cytochrome P450. **COG3321:** Acyl transferase domain in polyketide synthase (PKS) enzymes. **COG0145:** methylhydantoinase A/oxoprolinase/acetone carboxylase, beta subunit: **COG1233:** Phytoene dehydrogenase-related protein

We focused on the functions that are known to be important for plant nutrition and health such as nitrogen fixation, siderophore production, ACC deaminase etc. And functions involved in production of secondary metabolites. Interestingly, 3 COG groups related to siderophore production and transport mostly assigned to *Firmicutes* and *Proteobacteria*, were highly enriched in rhizosphere samples which suggest competitive conditions for Iron acquisition were siderophore might facilitate its availability to the plant or scavenge it from pathogens. Contrary COG groups related to secondary metabolism were more enriched in the phyllosphere and mostly assigned to *Ascomycota* and *Actinobacteria* scaffolds, for example the acyl transferase domain of polyketide synthase (PKS) enzymes, that are known to be involved in the biosynthesis of a highly diverse group of natural products (polyketides) with many functions as virulence factors, pigments, info-chemicals etc. (Hertweck 2009). and a phytoene dehydrogenase-related protein that is involved in carotene biosynthesis which have shown to be involved in UV and oxidative protection in bacteria and fungi (Avalos & Carmen Limón 2014; Mohammadi et al. 2012), a phenomenon that is very relevant in the phylosphere of plant that is completely expose to sun radiation

We demonstrated that microbial communities from different compartments of the plant holobiont are not only different in composition but also in functionality and some of the significate differences are related to constrains and factors shaping the rhizosphere and phyllosphere compartments and might contribute to the plant holobiont's fitness.

XI. PERSPECTIVES

We aim to finish the characterization of the community composition of cacti seed-borne endophytes by amplicon sequencing. Our main goal is to determine what factors are driving community composition (species, geography, phylogeny) and how distributed are the OTUs that contributed to plant fitness.

Some of our-seed borne strains are not present in the cacti microbiome data, therefore the microbial communities of flowers and gametes are very interesting to study because they might be responsible for their transmission to the seeds. Although, we are aware that this is a challenge due to the flowering habits of cacti.

We expect to generate a bigger collection of seed-borne strains from the seeds sampled in this work to tests their effect on the germination of *A. thaliana*, cacti and crop seeds under standard and low water availability. Also, it's interesting to shed light in the molecular mechanism involved in the promotion of germination rate by our seed-borne strains and compare them against their closest relatives with other lifestyles.

Finally, we will perform the comparative genomics analysis for the rest of seed-borne strains when their genomes become available and we will analyze phyllosphere and rhizosphere metagenomes between wild and cultivated agaves.

XII. SUPPLEMENTARY MATERIAL

Table S1. Origin of cacti seed-borne strains

Simple	Morphotype	Media	Identification
Seedling Mg.SF.2	A.3	Winogradsky	-
	D.1	TSA	Familia *Paenibacillaceae*
Seedling Or.SF.1	A.2	Winogradsky	-
Seedling Or.SF.3	A.5	Winogradsky	*Bacillus* sp.
Seed Or.SF	A.1	Winogradsky	*Bacillus* sp.
	B.1	Winogradsky	*Bacillus* sp.
Seed Or.Ma	A.4	Winogradsky	-
	B.2	Winogradsky	*Bacillus* sp.
	C1	Winogradsky	*Bacillus* sp.
Seed Mg.SF	A6	Winogradsky	*Bacillus* sp.
Seed Mg.Ma	A7	Winogradsky	*Bacillus* sp.

Mg: *M. geometrizans*. Or: *O. robusta*. Ma: Magueyal. SF: San Francisco

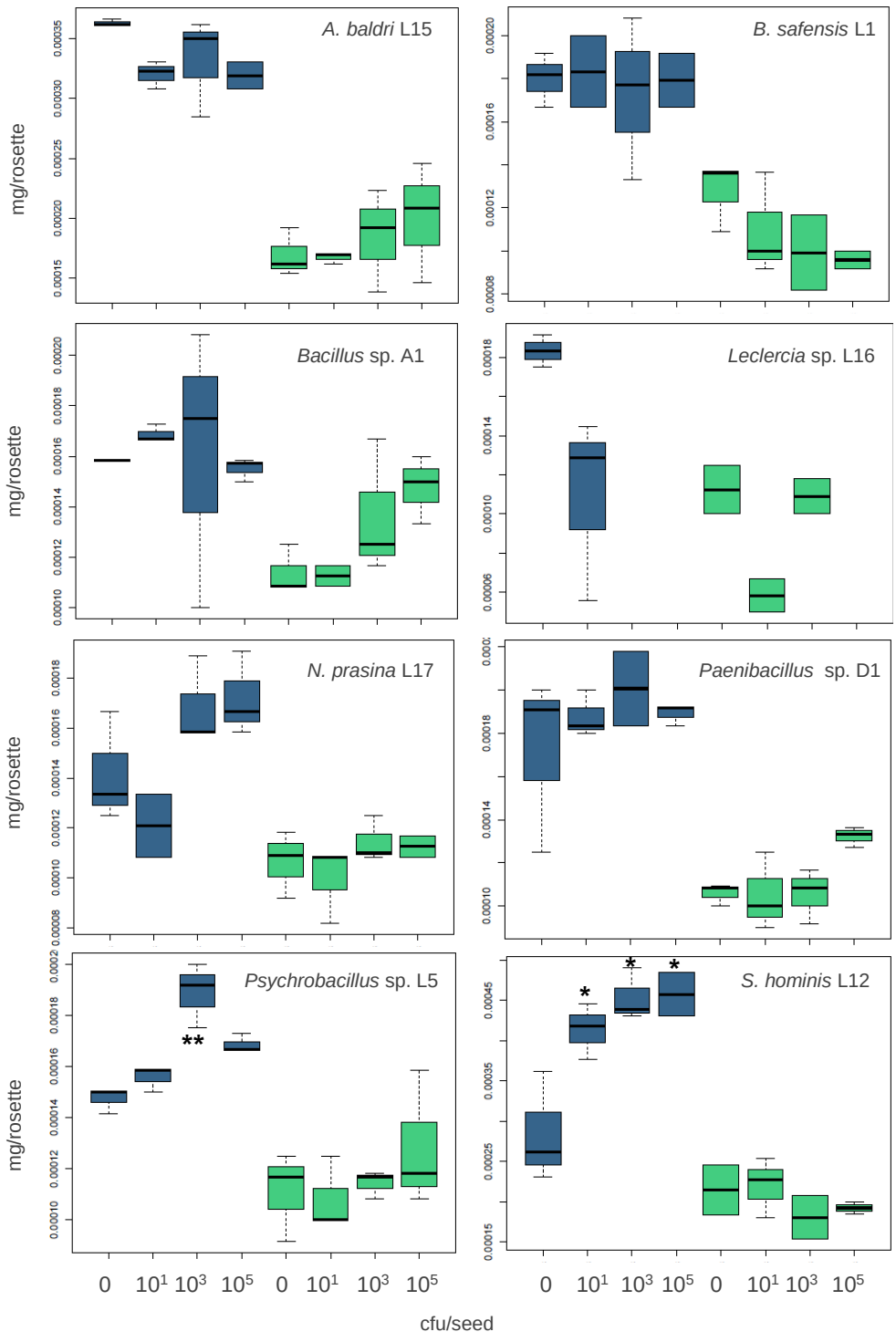


Figure S1. Effect of cacti seed-borne on *A. thaliana* dry weight. Blue and green boxes represent the distribution of the dry weight on 3 replicates under standard and low water potential respectively. * p<0.05 and ** p<0.001against the no inoculated control (One way ANOVA)

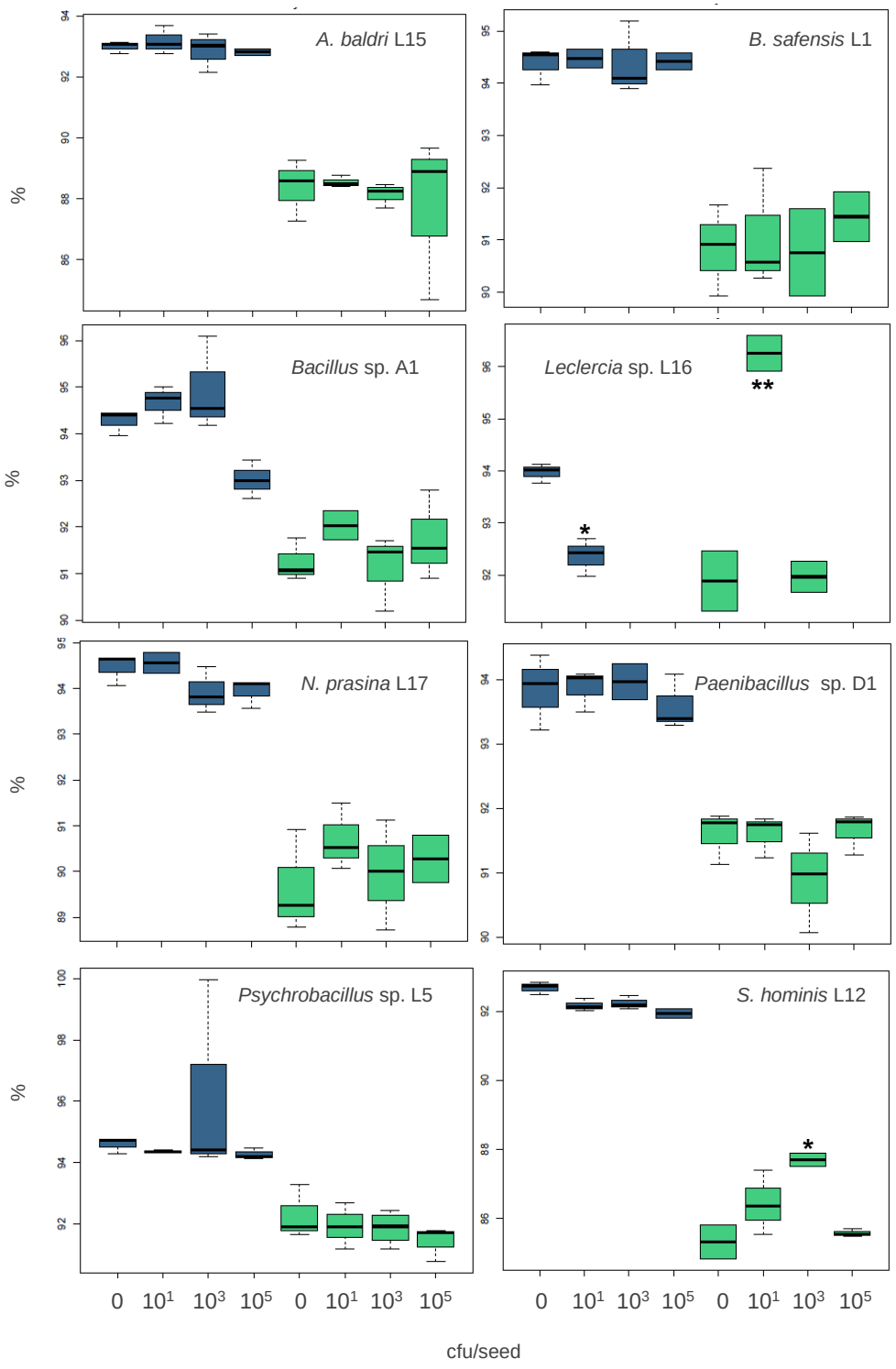


Figure S2. Effect of cacti seed-borne on *A. thaliana* humidity. Blue and green boxes represent the distribution of the humidity on 3 replicates under standard and low water potential respectively. * p<0.05 and ** p<0.001 against the no inoculated control (One way ANOVA)

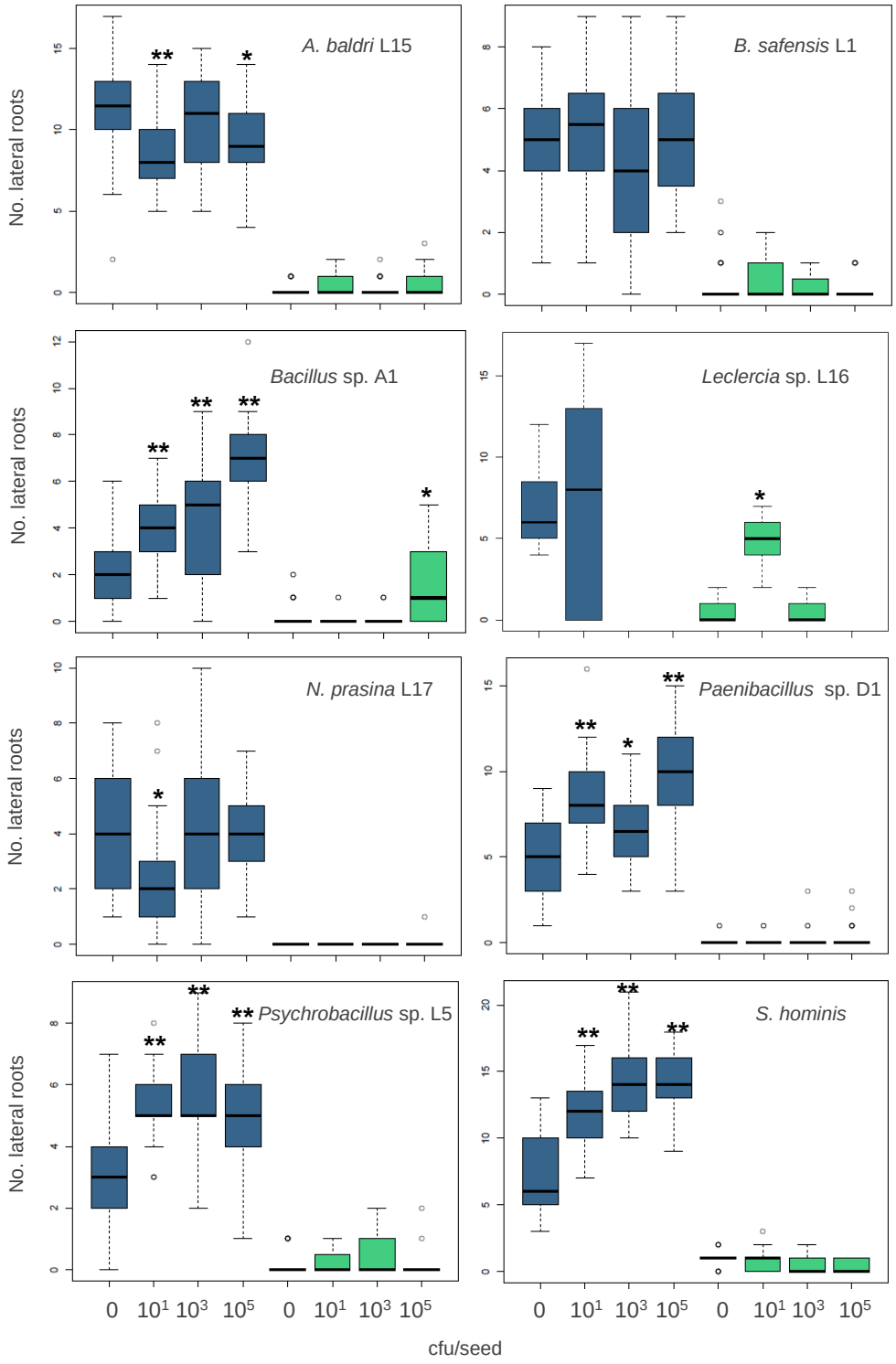


Figure S3. Effect of cacti seed-borne on *A. thaliana* lateral root number. Blue and green boxes represent the distribution of the lateral root on ~36 replicates under standard and low water potential respectively. * p<0.05 and ** p<0.001 the no inoculated control (One way ANOVA)

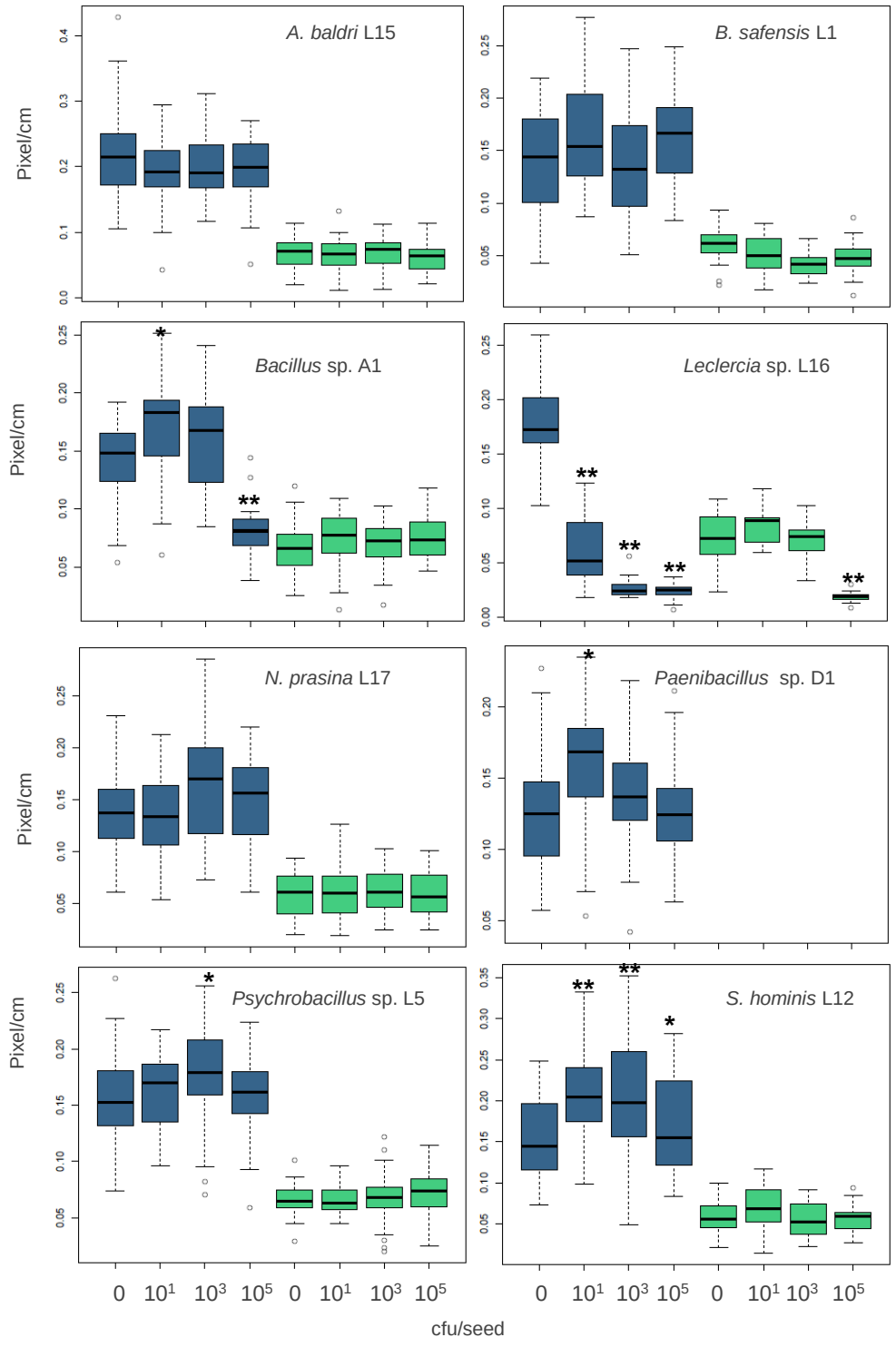


Figure S4. Effect of cacti seed-borne on *A. thaliana* leaf surface number. Blue and green boxes represent the distribution of the leaf surface on ~36 replicates under standard and low water potential respectively. * p<0.05 and ** p<0.001 against the no inoculated control (One way ANOVA)

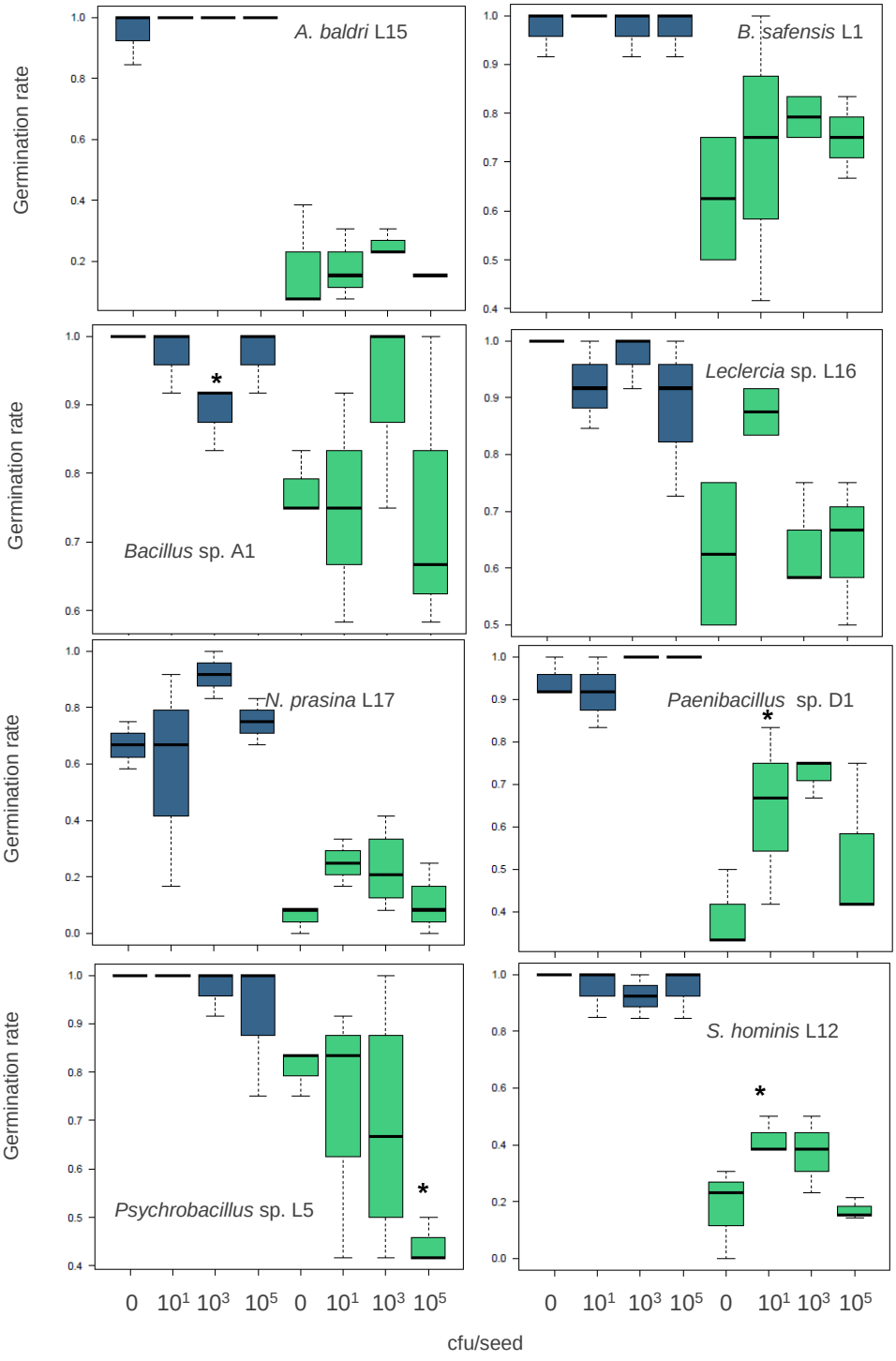


Figure S5. Effect of cacti seed-borne on *A. thaliana* leaf surface number. Blue and green boxes represent the distribution of the leaf surface on ~36 replicates under standard and low water potential respectively. * p<0.05 and ** p<0.001 against the no inoculated control (One way ANOVA)

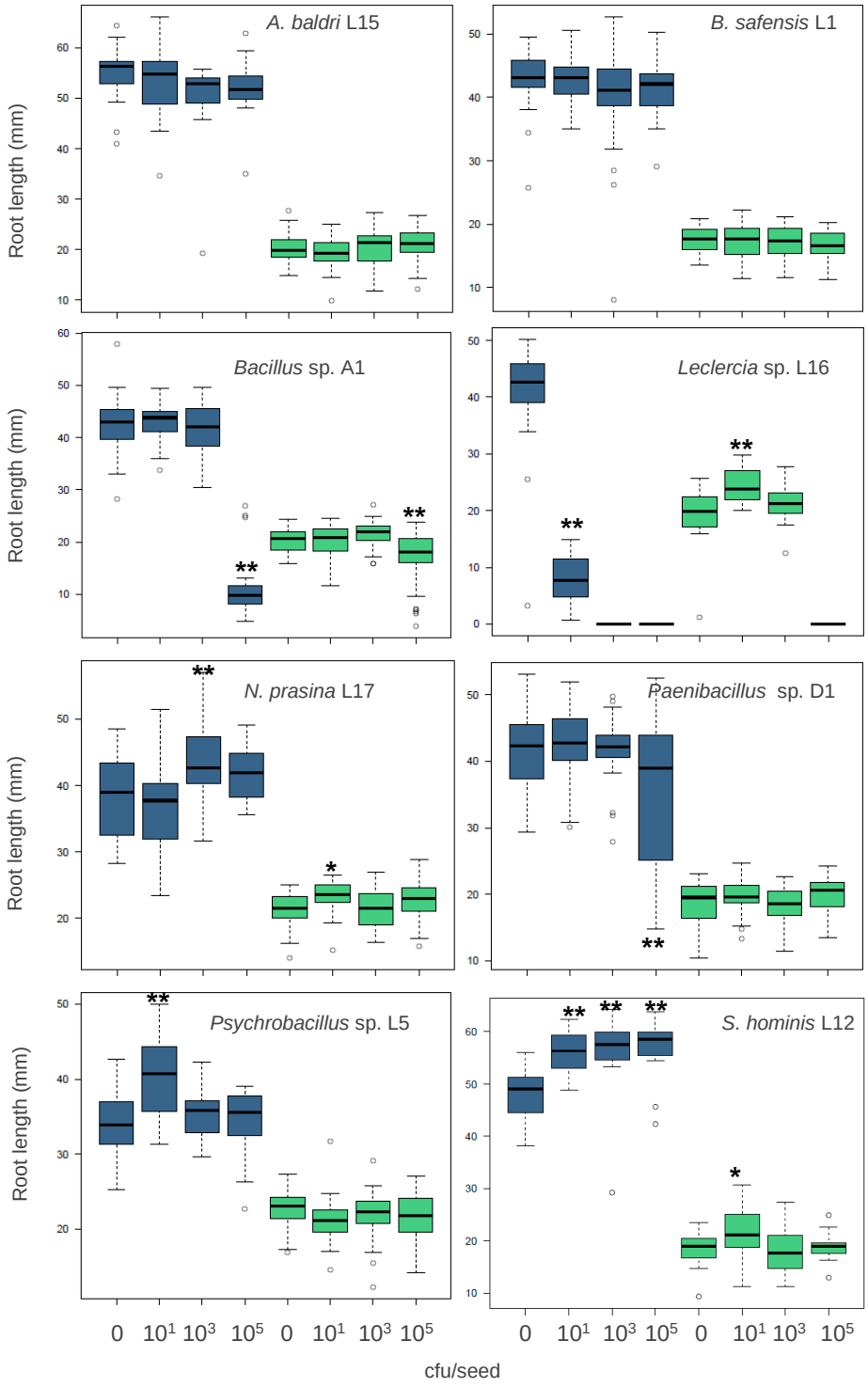


Figure S6. Effect of cacti seed-borne on *A. thaliana* leaf surface number. Blue and green boxes represent the distribution of the leaf surface on ~36 replicates under standard and low water potential respectively. * p<0.05 and ** p<0.001 against the no inoculated control (One way ANOVA)

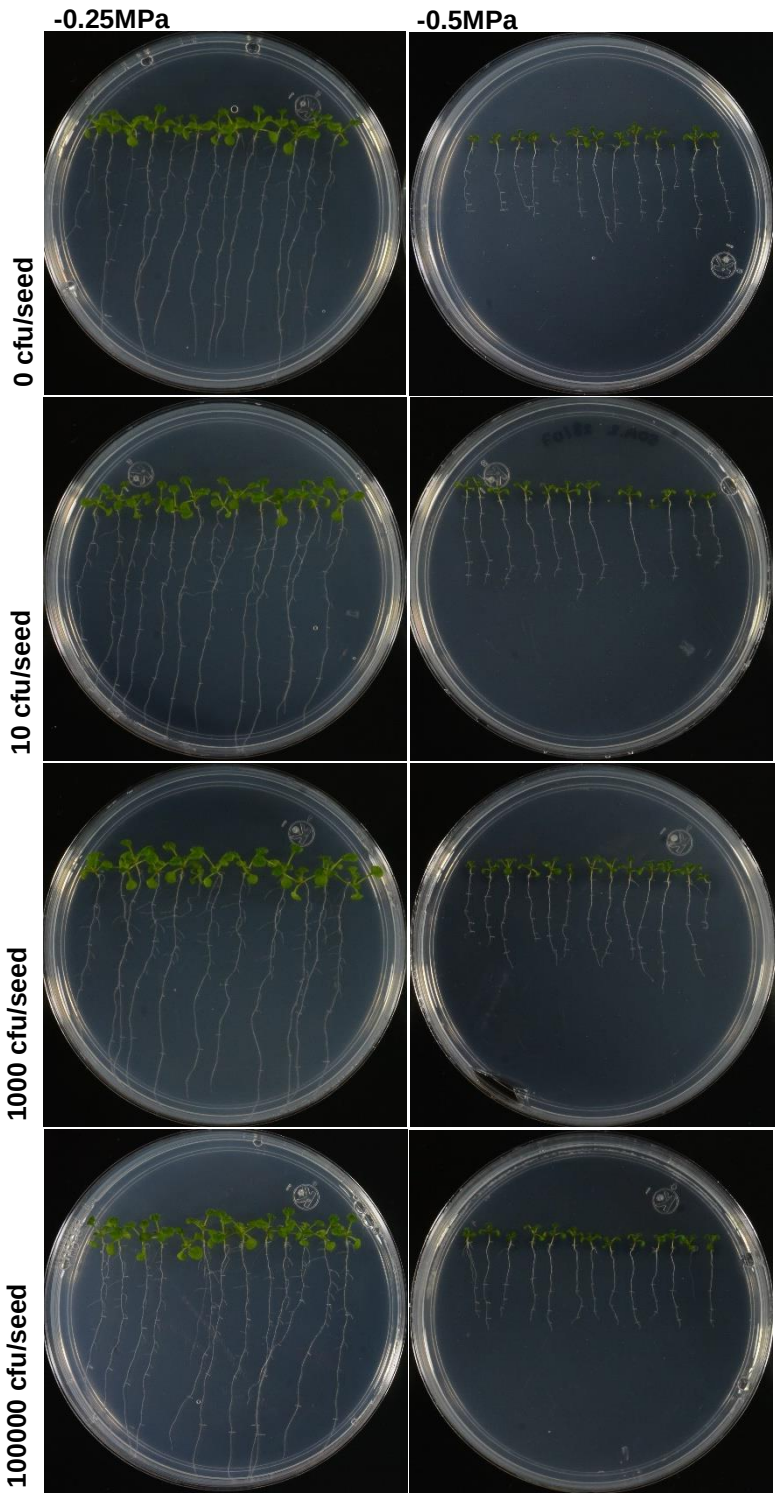


Figure S7. Effect of cacti seed-borne *S. hominis* on *A. thaliana*. -0.25MPa: standard water potential. -0.5MPa: low water potential

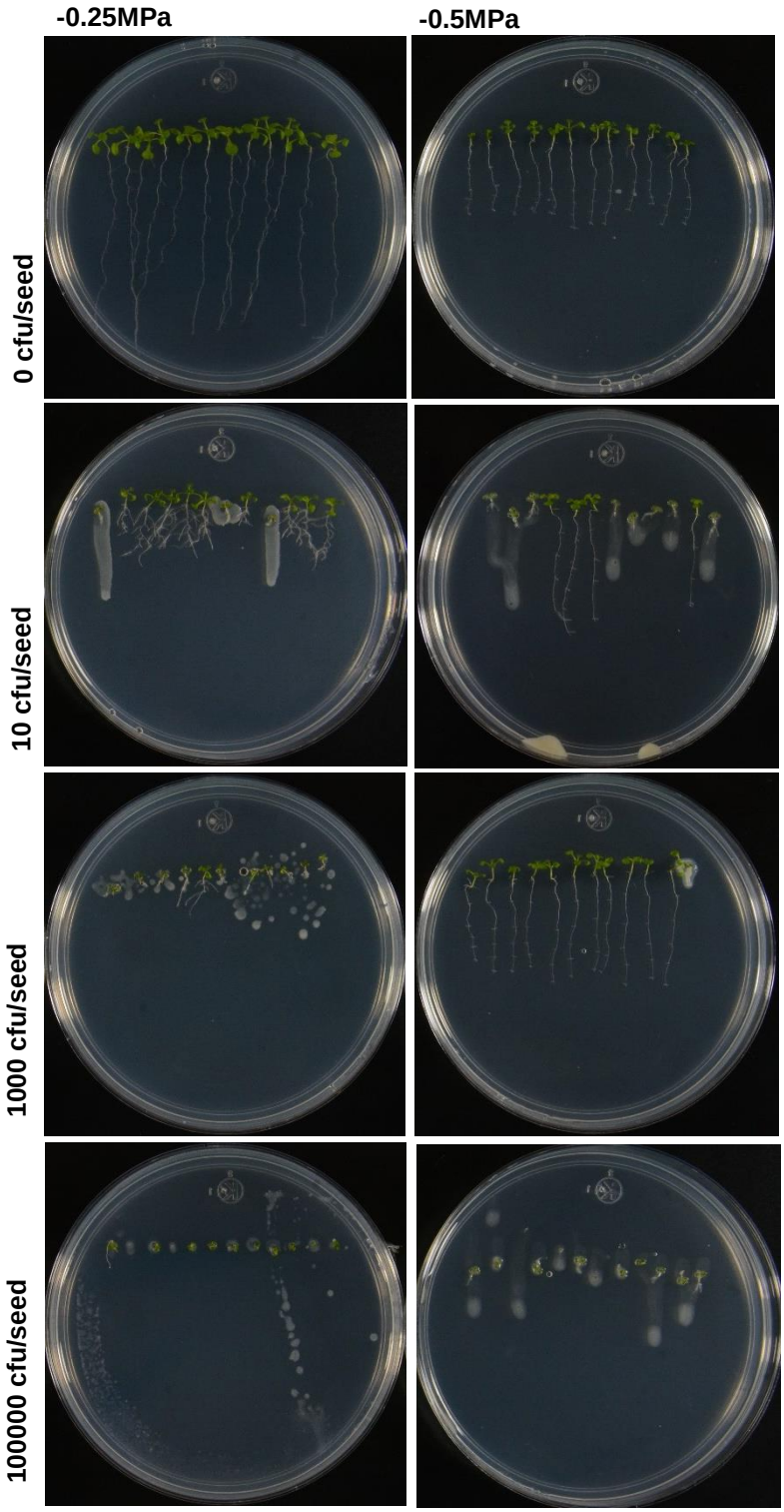


Figure S8. Effect of cacti seed-borne *Leclercia* on *A. thaliana*. -0.25MPa: standard water potential. -0.5MPa: low water potential

Table S2. COG categories description

Category	Description
A	RNA processing and modification
B	Chromatin Structure and dynamics
C	Energy production and conversion
D	Cell cycle control and mitosis
E	Amino Acid metabolism and transport
F	Nucleotide metabolism and transport
G	Carbohydrate metabolism and transport
H	Coenzyme metabolism
I	Lipid metabolism
J	Translation
K	Transcription
L	Replication, recombination and repair
M	Cell wall/membrane/envelop biogenesis
N	Cell motility
O	Post-translational modification, protein turnover, chaperone functions
P	Inorganic ion transport and metabolism
Q	Secondary metabolite biosynthesis/transport/catabolism
R	General Functional Prediction only
S	Function Unknown
T	Signal Transduction
U	Intracellular trafficking and secretion
V	Defense mechanisms
W	Extracellular structures
X	Mobilome: prophages, transposons
Y	Nuclear structure
Z	Cytoskeleton

XIII. REFERENCES

Aguirre-Garrido, J.F. et al., 2012. Bacterial community structure in the rhizosphere of three cactus species from semi-arid highlands in central Mexico. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 101(4), pp.891–904.

Alavi, P. et al., 2014. Stenotrophomonas comparative genomics reveals genes and functions that differentiate beneficial and pathogenic bacteria. *BMC Genomics*, 15(June), p.482: 15 pages.

Anaya-Pérez, M.A. & Bautista-Zane, R., 2008. El nopal forrajero en méxico: del siglo XVI al siglo XX. *Agricultura, Sociedad y Desarollo*, 5(2), pp.167–178.

Anuradha, M., 2014. Leclercia adecarboxylata isolation: Case reports and review. *Journal of Clinical and Diagnostic Research*, 8(12), p.DD03-DD04.

Atia, A. et al., 2009. ABA, GA3, and nitrate may control seed germination of Crithmum maritimum (Apiaceae) under saline conditions. *Comptes Rendus - Biologies*, 332(8), pp.704–710.

Avalos, J. & Carmen Limón, M., 2014. Biological roles of fungal carotenoids. *Current Genetics*, 61(3), pp.309–324.

Aziz, R.K., Breitbart, M. & Edwards, R.A., 2010. Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic Acids Research*, 38(13), pp.4207–4217.

Bai, Y. et al., 2015. Functional overlap of the Arabidopsis leaf and root microbiota. *Nature*, 528(7582), pp.364–369.

Barret, M. et al., 2015. Emergence shapes the structure of the seed microbiota. *Applied and Environmental Microbiology*, 81(4), pp.1257–1266.

Begun, J. et al., 2007. Staphylococcal biofilm exopolysaccharide protects against Caenorhabditis elegans immune defenses. *PLoS Pathogens*, 3(4), pp.526–540.

Bej, A.K., Aislabie, J. & Atlas, R.M., 2009. *Polar microbiology: the ecology, biodiversity and bioremediation potential of microorganisms in extremely cold environments*, CRC Press.

Berg, G., Eberl, L. & Hartmann, A., 2005. The Rhizosphere as a Reservoir for

Oppurtunistic Human Pathogenic Bacteria. *Environmental Microbiology*, 7(11), pp.1672–1685.

Bermúdez-Barrientos, J.R., 2016. *Exploring the molecular mechanisms maintaining the Rhizopus microsporus - Burkholderia rhizoxinica symbiosis*. Centro de Investigación y de Estudios Avanzados del IPN - Unidad Irapuato.

Bordenstein, S.R. & Theis, K.R., 2015. Host biology in light of the microbiome: Ten principles of holobionts and hologenomes. *PLoS Biology*, 13(8), pp.1–23.

Bork, P. et al., 1998. Predicting function: from genes to genomes and back. *J Mol Biol*, 283(4), pp.707–725.

Bragina, A. et al., 2013. Vertical transmission explains the specific Burkholderia pattern in Sphagnum mosses at multi-geographic scale. *Frontiers in Microbiology*, 4(DEC), pp.1–10.

Bright, M. & Bulgheresi, S., 2010. A complex journey: transmission of microbial symbionts. *Nature Reviews Microbiology*, 8(3), pp.218–230.

Brooks, J.P. et al., 2015. The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiology*, 15(1), p.66.

Brucker, R.M. & Bordenstein, S.R., 2013. The capacious hologenome. *Zoology*, 116(5), pp.260–261.

Cankar, K. et al., 2005. Bacterial endophytes from seeds of Norway spruce (Picea abies L. Karst). *FEMS Microbiology Letters*, 244(2), pp.341–345.

Cassán, F. et al., 2009. Azospirillum brasilense Az39 and Bradyrhizobium japonicum E109, inoculated singly or in combination, promote seed germination and early seedling growth in corn (Zea mays L.) and soybean (Glycine max L.). *European Journal of Soil Biology*, 45(1), pp.28–35.

Chamberlain, S.A., Bronstein, J.L. & Rudgers, J.A., 2014. How context dependent are species interactions? *Ecology Letters*, 17(7), pp.881–890.

Chaudhry, V. & Patil, P.B., 2016. Genomic investigation reveals evolution and lifestyle adaptation of endophytic Staphylococcus epidermidis. *Scientific Reports*, 6(October 2015), p.19263.

Chen, I.-M.A. et al., 2016. IMG/M: integrated genome and metagenome comparative data analysis system. *Nucleic acids research*, 45(October 2016), p.gkw929.

Cho, S.M. et al., 2008. 2R,3R-butanediol, a bacterial volatile produced by Pseudomonas chlororaphis O6, is involved in induction of systemic tolerance to drought in Arabidopsis thaliana. *Molecular plant-microbe interactions : MPMI*, 21(8), pp.1067–1075.

Cho, S.M. et al., 2012. Induced systemic drought and salt tolerance by Pseudomonas chlororaphis O6 root colonization is mediated by ABA-independent stomatal closure. *Plant Pathology Journal*, 28(2), pp.202–206.

Cline, W.R., 2008. Global Warming and Agriculture. *Finance and Development*, 1(March), pp.23–27.

Coates, R., Moran, J. & Horsburgh, M.J., 2014. Staphylococci: Colonizers and Pathogens of Human Skin. Available at: http://www.medscape.com/viewarticle/818738_17 [Accessed January 5, 2017].

Cui, M. & Nobel, P.S., 1992. Nutrient status, water uptake and gas exchange for three desert succulents infected with mycorrhizal fungi. *New Phytologist*, 122(4), pp.643–649.

Desgarennes, D. et al., 2014. Diazotrophic potential among bacterial communities associated with wild and cultivated Agave species. *FEMS Microbiology Ecology*, 90(3), pp.844–857.

Ebert, D., 2013. The Epidemiology and Evolution of Symbionts with Mixed-Mode Transmission. *Annual Review of Ecology, Evolution, and Systematics*, 44(1), pp.623–643.

Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), pp.1792–1797.

Farooq, M. et al., 2012. Drought Stress in Plants: An Overview. In R. Aroca, ed. *Plant Responses to Drought Stress*. Berlin Heidelberg: Farooq, M., Hussain, M., Wahid, A., & Siddique, K. H. M. (2012). Drought Stress in Plants: An Overview. In Plant Responses to Drought Stress (pp. 37–61). http://doi.org/10.1007/978-3-642-32653-0, pp. 37–61.

Fonseca-García, C. et al., 2016. The Cacti Microbiome: Interplay between Habitat-Filtering and Host-Specificity. *Frontiers in microbiology*, 7(February), p.150.

Frébort, I. et al., 2011. Evolution of cytokinin biosynthesis and degradation. *Journal of Experimental Botany*, 62(8), pp.2431–2452.

Garcia, J.R. & Gerardo, N.M., 2014. The symbiont side of symbiosis: Do microbes really benefit? *Frontiers in Microbiology*, 5(SEP), pp.1–6.

Glick, B.R. & Glick, B.R., 2012. Plant Growth-Promoting Bacteria: Mechanisms and Applications. *Scientifica*, 2012, pp.1–15.

Griffith, M. & Porter, J.M., 2009. Phylogeny of Opuntioideae (Cactaceae). *International Journal of Plant Sciences*, 170(1), pp.107–116.

Guérillot, R. et al., 2014. The diversity of prokaryotic DDE transposases of the mutator superfamily, insertion specificity, and association with conjugation machineries. *Genome Biology and Evolution*, 6(2), pp.260–272.

Guerrero, R. & Berlanga, M., 2015. From the Cell to the Ecosystem: The Physiological Evolution of Symbiosis. *Evolutionary Biology*, pp.1–10.

Hacquard, S., 2016. Disentangling the factors shaping microbiota composition across the plant holobiont. *New Phytologist*, 209(2), pp.454–457.

Hakeem, K.R., Akhtar, M.S. & Abdullah, S.N.A., 2016. Plant, Soil and Microbes. *Plant, Soil and Microbes*, 45, p.375.

Hardoim, P.R. et al., 2012. Dynamics of seed-borne rice endophytes on early plant growth stages. *PLoS ONE*, 7(2).

Hardoim, P.R. et al., 2015. The Hidden World within Plants: Ecological and Evolutionary Considerations for Defining Functioning of Microbial Endophytes. *Microbiology and Molecular Biology Reviews*, 79(3), pp.293–320.

Hegarty, T.W. & ROSS, H.A., 1978. Differential sensitivity to moisture stress of seed germination and seedling radicle growth in calabrese (Brassica oleracea var. italica) and cress (Lepidium sativum). *Annals of Botany*, 42(180), pp.1003–1005.

Hernández-Hernández, T. et al., 2014. Beyond aridification: Multiple explanations for the elevated diversification of cacti in the New World Succulent Biome. *New Phytologist*,

202(4), pp.1382–1397.

Hernández-Hernández, T. et al., 2011. Phylogenetic relationships and evolution of growth form in Cactaceae (Caryophyllales, Eudicotyledoneae). *American Journal of Botany*, 98(1), pp.44–61.

Hertweck, C., 2009. The biosynthetic logic of polyketide diversity. *Angewandte Chemie - International Edition*, 48(26), pp.4688–4716.

Huang, J. et al., 2015. Accelerated dryland expansion under climate change. *Nature Climate Change*, (October), pp.1–22.

Jefferson, R., 1994. The Hologenome. *Agriculture, Environment and the Developing World: A Future of PCR. Cold Spring Harbor, New York*.

Jukes, T.H. & Cantor, C.R., 1969. Evolution of protein molecules. *Mammalian Protein Metabolism*, pp.21–123.

Kelly, S.L. & Kelly, D.E., 2013. Microbial cytochromes P450: biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 368(1612), p.20120476.

Kimura, M., 1980. Journal of Molecular Evolution A Simple Method for Estimating Evolutionary Rates of Base Substitutions Through Comparative Studies of Nucleotide Sequences. *J. Mol. Evol*, 16(1330), pp.111–120.

Klaedtke, S. et al., 2015. Terroir is a key driver of seed-associated microbial assemblages. *Environmental Microbiology*.

Knief, C. et al., 2012. Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *The ISME Journal*, 6(7), pp.1378–1390.

Koh, H.Y. et al., 2012. Draft genome sequence of Paenisporosarcina sp. Strain TG-14, a psychrophilic bacterium isolated from sediment-laden stratified basal ice from Taylor Glacier, McMurdo Dry Valleys, Antarctica. *Journal of Bacteriology*, 194(23), pp.6656–6657.

Krishnamurthi, S. et al., 2010. Psychrobacillus gen. nov. and proposal for reclassification of Bacillus insolitus Larkin & Stokes, 1967, B. psychrotolerans Abd-El Rahman et al.,

2002 and B. psychrodurans Abd-El Rahman et al., 2002 as Psychrobacillus insolitus comb. nov., Psychrobacillus. *Systematic and Applied Microbiology*, 33(7), pp.367–373.

Kumar, S., Stecher, G. & Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, p.msw054.

De la Barrera, E. & Smith, W.K., 2009. *Perspectives in Biophysical Plant Ecophysiology: A Tribute to Park S. Nobel* m. Mexico, ed., UNAM.

Lackner, G. et al., 2011. Complete genome sequence of burkholderia rhizoxinica, an endosymbiont of Rhizopus microsporus. *Journal of Bacteriology*, 193(3), pp.783–784.

Lima, J.V.L. et al., 2015. Endophytic bacteria in cacti native to a Brazilian semi-arid region. *Plant and Soil*, 389(1–2), pp.25–33.

Lopez, B.R., Bashan, Y. & Bacilio, M., 2011. Endophytic bacteria of Mammillaria fraileana, an endemic rock-colonizing cactus of the southern Sonoran Desert. *Archives of Microbiology*, 193(7), pp.527–541.

Majure, L.C. et al., 2012. Phylogeny of Opuntia s.s. (Cactaceae): Clade delineation, geographic origins, reticulate evolution. *American Journal of Botany*, 99(5), pp.847–864.

Mano, H. et al., 2006. Culturable Surface and Endophytic Bacterial Flora of the Maturing Seeds of Rice Plants (Oryza sativa) Cultivated in a Paddy Field. *Microbes Environ*, 21(2), pp.86–100.

Martínez-Cano, D.J. et al., 2015. Evolution of small prokaryotic genomes. *Frontiers in Microbiology*, 6(JAN), pp.1–23.

Midha, S. et al., 2016. Genomic resource of rice seed associated bacteria. *Frontiers in Microbiology*, 6(JAN), pp.1–8.

Miransari, M. & Smith, D.L., 2014. Plant hormones and seed germination. *Environmental and Experimental Botany*, 99, pp.110–121.

Mohammadi, M., Burbank, L. & Roper, M.C., 2012. Biological role of pigment production for the bacterial phytopathogen Pantoea stewartii subsp. stewartii. *Applied and*

Environmental Microbiology, 78(19), pp.6859–6865.

Moran, N.A. & Plague, G.R., 2004. Genomic changes following host restriction in bacteria. *Current Opinion in Genetics and Development*, 14(6), pp.627–633.

Moreno-Hagelsieb, G. & Latimer, K., 2008. Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics*, 24(3), pp.319–324.

Mukhopadhyay, K. et al., 1996. Identification and characterization of bacterial endophytes of rice. *Mycopathologia*, 134(3), pp.151–159.

Nannipieri, P. et al., 2003. Microbial diversity and soil functions. *European journal of soil science*, 54(4), pp.655–670.

Nemergut, D.R. et al., 2013. Patterns and Processes of Microbial Community Assembly. *Microbiology and Molecular Biology Reviews*, 77(3), pp.342–356.

Nickell, L.G. & Finlay, A.C., 1954. Antibiotics and Their Effects on Plant Growth. *Journal of Agricultural and Food Chemistry*, (27), pp.178–182.

Nobel, P.S., 2002. *Cacti: Biology and Uses*, Los Angeles: University of California Press.

Nobel, P.S., 1988. *Environmental Biology of Agave and Cati*, New York: Cambridge University Press.

Nyffeler, R., 2002. Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from trnK/matK and trnL-trnF sequences. *American Journal of Botany*, 89(2), pp.312–326.

Overbeek, R. et al., 1999. The use of gene clusters to infer functional coupling. *Proceedings of the National Academy of Sciences of the United States of America*, 96(6), pp.2896–2901.

Parte, A.C., 2014. LPSN - List of prokaryotic names with standing in nomenclature. *Nucleic Acids Research*, 42(D1), pp.613–616.

Partida-Martínez, L.P. & Heil, M., 2011. The Microbe-Free Plant: Fact or Artifact? *Frontiers in Plant Science*, 2(December), pp.1–16.

Pérez-Molphe-Balch, E. et al., 2002. In vitro propagation of three species of columnar cacti from the sonoran desert. *HortScience*, 37(4), pp.693–696.

Puente, M.E., Bashan, Y., et al., 2004. Microbial populations and activities in the rhizoplane of rock-weathering desert plants. I. Root colonization and weathering of igneous rocks. *Plant Biology*, 6(5), pp.629–642.

Puente, M.E., Li, C.Y. & Bashan, Y., 2009a. Endophytic bacteria in cacti seeds can improve the development of cactus seedlings. *Environmental and Experimental Botany*, 66(3), pp.402–408.

Puente, M.E., Li, C.Y. & Bashan, Y., 2004. Microbial populations and activities in the rhizoplane of rock-weathering desert plants. II. Growth promotion of cactus seedlings. *Plant Biology*, 6(5), pp.643–650.

Puente, M.E., Li, C.Y. & Bashan, Y., 2009b. Rock-degrading endophytic bacteria in cacti. *Environmental and Experimental Botany*, 66(3), pp.389–401.

R Core Team, 2015. R: A Language and Environment for Statistical Computing.

Reddy, G.S.N. et al., 2013. Paenisporosarcina indica sp. nov., a psychrophilic bacterium from a glacier, and reclassification of Sporosarcina antarctica Yu et al., 2008 as Paenisporosarcina antarctica comb. nov. and emended description of the genus Paenisporosarcina. *International Journal of Systematic and Evolutionary Microbiology*, 63(PART8), pp.2927–2933.

Reynolds, J.F. et al., 2007. Natural and human dimensions of land degradation in drylands: causes and consequences. In *Terrestrial ecosystems in a changing world*. Springer, pp. 247–257.

Roberson, E.B. & Firestone, M.K., 1992. Relationship between desiccation and exopolysaccharide production in a soil *Pseudomonas* sp. *Applied and Environmental Microbiology*, 58(4), pp.1284–1291.

Robinson, M.D., McCarthy, D.J. & Smyth, G.K., 2009. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), pp.139–140.

Rojas-Aréchiga, M. & Vázquez-Yanes, C., 2000. Cactus seed germination: a review. *Journal of Arid Environments*, 44(November), pp.85–104.

Rosenberg, E. et al., 2007. The role of microorganisms in coral health, disease and evolution. *Nature Reviews Microbiology*, 5(5), pp.355–362.

Rosenberg, E. & Zilber-Rosenberg, I., 2016. Microbes drive evolution of animals and plants: The hologenome concept. *mBio*, 7(2), pp.1–8.

Rosenberg, E. & Zilber-Rosenberg, I., 2013. *The Hologenome Concept: Human, Animal and Plant Microbiota*, Switzerland: Springer International Publishing.

Rzedowski, C.G. de & Rzedowski, J., 2005. *Flora fanerogámica del Valle de México*, Pátzcuaro: Instituto de Ecología, A.C. y Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.

Sáenz, C. et al., 2013. *Agro-industrial utilization of cactus pear*,

Saikkonen, K. et al., 2016. Endophytic Epichloë species and their grass hosts: from evolution to applications. *Plant Molecular Biology*, 90(6), pp.665–675.

Shetty, A.A., Rana, M.K. & Preetham, S.P., 2012. Cactus: A medicinal food. *Journal of Food Science and Technology*, 49(5), pp.530–536.

Shigenobu, S. et al., 2000. Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS. *Nature*, 407(6800), pp.81–86.

Soussi, A. et al., 2015. Plant-associated microbiomes in arid lands: diversity, ecology and biotechnological potential. *Plant and Soil*, (Whitford 2002).

de Souza, R., Ambrosini, A. & Passaglia, L.M.P., 2015. Plant growth-promoting bacteria as inoculants in agricultural soils. *Genetics and Molecular Biology*, 38(4), pp.401–419.

Taberlet, P. et al., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, 17(5), pp.1105–1109.

Tatusov, R.L. et al., 1997. A genomic perspective on protein families. *Science (New York, N.Y.)*, 278(5338), pp.631–7.

Toth, I.K., Pritchard, L. & Birch, P.R.J., 2006. Comparative genomics reveals what makes an enterobacterial plant pathogen. *Ann Rev Phytopathol*, 44, pp.305–336.

Truyens, S. et al., 2015. Bacterial seed endophytes: Genera, vertical transmission and interaction with plants. *Environmental Microbiology Reports*, 7(1), pp.40–50.

Truyens, S. et al., 2016. Cadmium-induced and trans-generational changes in the cultivable and total seed endophytic community of *Arabidopsis thaliana*. *Plant*

Biology, 18(3), pp.376–381.

Truyens, S. et al., 2013. Changes in the population of seed bacteria of transgenerationally Cd-exposed Arabidopsis thaliana. *Plant Biology*, 15(6), pp.971–981.

Tsavkelova, E.A. et al., 2007. Orchid-associated bacteria produce indole-3-acetic acid, promote seed germination, and increase their microbial yield in response to exogenous auxin. *Archives of Microbiology*, 188(6), pp.655–664.

UNCCD & Unep, 2011. Global Drylands: A UN system-wide response. , p.132.

Vacheron, J. et al., 2013. Plant growth-promoting rhizobacteria and root system functioning. *Frontiers in plant science*, 4(September), p.356.

Vallejo, A.J., Yanovsky, M.J. & Botto, J.F., 2010. Germination variation in Arabidopsis thaliana accessions under moderate osmotic and salt stresses. *Annals of Botany*, 106(5), pp.833–842.

Vandenkoornhuyse, P. et al., 2015. The importance of the microbiome of the plant holobiont. *New Phytologist*, 206(4), pp.1196–1206.

Verslues, P.E. et al., 2006. Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. *Plant Journal*, 45(4), pp.523–539.

Verslues, P.E. & Bray, E.A., 2004. LWR1 and LWR2 are required for osmoregulation and osmotic adjustment in Arabidopsis. *Plant Physiol*, 136(1), pp.2831–2842.

Wagner, M.R. et al., 2016. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat Commun*, 7, pp.1–15.

van der Weele, C.M. et al., 2000. Growth of Arabidopsis thaliana seedlings under water deficit studied by control of water potential in nutrient-agar media. *Journal of experimental botany*, 51(350), pp.1555–1562.

White, T.. J.. et al., 1990. Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics. In *PCR protocols: a guide to methods and applications*. Academic Press Inc.: Academic Press Inc., pp. 315–322.

Wickham, H., 2009. *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag New York.

Zhang, H. et al., 2010. Choline and osmotic-stress tolerance induced in Arabidopsis by the soil microbe Bacillus subtilis (GB03). *Molecular plant-microbe interactions : MPMI*, 23(8), pp.1097–104.

Zheng, C. et al., 2009. Exogenous nitric oxide improves seed germination in wheat against mitochondrial oxidative damage induced by high salinity. *Environmental and Experimental Botany*, 67(1), pp.222–227.

Zilber-Rosenberg, I. & Rosenberg, E., 2008. Role of microorganisms in the evolution of animals and plants: The hologenome theory of evolution. *FEMS Microbiology Reviews*, 32(5), pp.723–735.

OTRAS ACTIVIDADES

CONGRESOS

Flores-Núñez, V. M., Fonseca-García, C. & Partida-Martínez, L. P. Dissecting microbiome functions in cacti: lessons from seed-transmitted endophytes (póster). XXXI Congreso Nacional de Bioquímica. Aguascalientes, Aguascalientes. Sociedad Mexicana de Bioquímica A.C.

ACTIVIDADES DE DIFUSIÓN

Participación en las Jornadas de Divulgación de Ciencia y Tecnología del Estado de Guanajuato, con el tema *El origen de la vida en la Tierra*. CINVESTAV Unidad Irapuato, Guanajuato. 15 de Octubre de 2016

ESTANCIAS

Estancia en el laboratorio de *Computational conSequences* bajo la asesoría de Gabriel Moreno Hagelsieb. Wilfrid Laurier University, Waterloo, ON, Canadá.

VÍCTOR MANUEL FLORES NÚÑEZ

10/Diciembre/1990 26 años Mexicano

Contacto: vicmfloresnz@gmail.com

+52 55 14 57 83 45

Experiencia

Cinco años de experiencia en métodos de microbiología clásica y molecular para el aislamiento, caracterización, identificación y manejo de microorganismos asociados a plantas. Análisis general de genomas bacterianos y datos metagenómicos. Experiencia intermedia en programación en R.

Formación

M. en C. en Biología Integrativa. (2017). CINVESTAV Irapuato. Titulación pendiente para Marzo de 2017.

Químico Bacteriólogo Parasitólogo (2014). Graduado con honores de la Escuela Nacional de Ciencias Biológicas. Instituto Politécnico Nacional.

Idiomas

Español (Lengua materna)

Inglés

Referencias

Laila P. Partida Martínez. Tutora de tesis de maestría.

laila.partida@cinvestav.mx

Enriqueta F. Amora Lazcano. Tutora de tesis de licenciatura.

eamoral@hotmail.com

DECLARACIÓN DE INDEPENDENCIA

Por este medio declaro que yo he preparado el presente trabajo de tesis de forma independiente y sin ayuda externa. Además, he citado de forma correcta y explícita a los autores en los que esta tesis se respalda, así como las contribuciones de las personas que participaron en su desarrollo.

Lugar: Irapuato, Guanajuato, México

Fecha: 24 de Febrero de 2017 Firma: