



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

Impacto del genotipo del huésped y el tipo de suelo en la diversidad taxonómica y metabólica de comunidades microbianas de raíz

TESIS
QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA

HUGO RAFAEL BARAJAS DE LA TORRE

TUTOR PRINCIPAL

Dr. Luis David Alcaraz Peraza
Facultad de Ciencias, UNAM

MIEMBROS DEL COMITÉ TUTOR

Dr. Luis Servín González
Instituto de Biotecnología, UNAM

Dr. Felipe García Oliva
Instituto de Investigaciones en Ecosistemas y Sustentabilidad, UNAM Morelia

Ciudad de México, abril, 2021



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice

Índice	1
Índice de Figuras	3
Índice de Tablas	4
Abreviaciones	5
Agradecimientos	6
Resumen	7
Abstract	8
Introducción	9
Propiedades generales de los suelos.	9
Índices de diversidad comúnmente empleados en ecología microbiana.....	10
El microbioma del suelo: Diversidad y funcionamiento.	11
El microbioma de plantas y sus estrategias de estudio.....	12
Factores involucrados en la estructuración de las comunidades microbianas de suelos y raíces	14
El modelo de dos pasos de adquisición del microbioma de raíz	16
Diferencias en el microbioma de plantas: Efectos de domesticación y la distancia evolutiva de los huéspedes.	17
Antecedentes.....	21
Planteamiento del problema	23
Hipótesis	23
Objetivo General.....	24
Objetivos Particulares	24
Metodología	25
Muestreo de suelos y raíces de plantas ruderales.....	25
Experimento de jardín común y obtención de muestras	26
Obtención de muestras de rizósfera y endósfera de raíces de <i>S. lycopersicum</i> y plantas ruderales	26

Análisis fisicoquímicos de suelos.....	27
Extracción de DNA metagenómico	27
Bibliotecas metagenómicas y secuenciación de amplicones del gen 16S rRNA y metagenomas totales.....	28
Procesamiento de lecturas de bibliotecas de amplicones.....	29
Asignación taxonómica de plantas ruderales.....	29
Procesamiento de lecturas de metagenomas totales.....	30
Predicción de proteínas y anotación.....	30
Cálculo de abundancias de proteínas.....	31
Análisis de diversidad taxonómica y metabólica	31
Identificación del microbioma y metagenoma núcleo de raíces de <i>S. lycopersicum</i> y plantas ruderales.....	32
Descripción del potencial metabólico de comunidades microbianas de rizósferas	32
Análisis de abundancia diferencial de OTUs y proteínas	33
Resultados.....	33
Descripción biogeoquímica de suelos, diversidad de plantas ruderales y producción de biomasa en plantas de <i>S. lycopersicum</i>	33
Diversidad del microbioma de suelos y raíces de <i>S. lycopersicum</i> y plantas ruderales: El modelo de dos pasos de adquisición del microbioma no se cumple en plantas ruderales	38
Influencia del genotipo del hospedero en la estructura de comunidades bacterianas de raíz: La convergencia del microbioma de raíz de <i>S. lycopersicum</i>	40
Enriquecimiento de grupos taxonómicos en raíces de plantas domesticadas y ruderales: La selección activa del microbioma por parte de los hospederos	43
Descripción del potencial metabólico de comunidades las microbianas de suelos y rizósferas	47
Diversidad metagenómica de suelos y rizósferas de <i>S. lycopersicum</i> y plantas ruderales: El modelo de dos pasos se cumple a nivel metabólico.....	49
Enriquecimiento de proteínas predichas en los metagenomas de rizósferas de <i>S. lycopersicum</i> y plantas ruderales.	52

El metagenoma núcleo de suelos y rizósferas de <i>S. lycopersicum</i> y plantas ruderales.....	56
Discusión	59
Conclusiones	69
Perspectivas	70
Referencias.....	73
Anexo I	86

Índice de Figuras

Figura I. La estructura de las comunidades microbianas de suelo es afectada por factores bióticos y abióticos.....	15
Figura II. Modelo de dos pasos de adquisición del microbioma de raíz.....	16
Figura III. Representación esquemática de las diferencias en el microbioma de plantas silvestres y domesticadas.....	19
Figura 1. Resumen gráfico del trabajo.....	25
Figura 2. Ejemplos de los sitios de colecta.....	34
Figura 3. Diversidad fisicoquímica de los suelos al inicio y final del experimento de jardín común y producción de biomasa de <i>S. lycopersicum</i>	36
Figura 4. Afiliación taxonómica de las plantas ruderales colectadas <i>in situ</i>	37
Figura 5. Riqueza y diversidad de los microbiomas de los suelos al inicio y final del experimento de jardín común, rizósferas y endósferas de <i>S. lycopersicum</i> y plantas ruderales.....	39
Figura 6. Diversidad α de suelos iniciales, rizósferas y endósferas de plantas ruderales y <i>S. lycopersicum</i>	39
Figura 7. Microbioma núcleo de plantas ruderales y <i>S. lycopersicum</i>	40
Figura 8. Influencia del genotipo del huésped y el suelo en la estructura y diversidad de las comunidades bacterianas.....	41
Figura 9. Análisis de diversidad beta de los microbiomas de suelos, rizósferas y endósferas de plantas ruderales y <i>S. lycopersicum</i>	43
Figura 10. Abundancia relativa de phyla bacterianos en suelos y rizósferas de plantas ruderales y <i>S. lycopersicum</i>	44
Figura 11. OTUs enriquecidos en la rizósfera y endósfera de plantas ruderales y de <i>S. lycopersicum</i>	46
Figura 12. Perfiles metabólicos predichos de los metagenomas de suelos y rizósferas.....	48

Figura 13. Riqueza y diversidad α de proteínas predichas en suelos y rizósferas de plantas ruderales y <i>S. lycopersicum</i>	50
Figura 14. Diversidad metagenómica en suelos y rizósferas de plantas ruderales y <i>S. lycopersicum</i>	51
Figura 15. Proteínas diferencialmente abundantes en rizósferas de <i>S. lycopersicum</i> y suelos iniciales.....	53
Figura 16. Proteínas diferencialmente abundantes en rizósferas de <i>S. lycopersicum</i> y de plantas ruderales.....	55
Figura 17. Proteínas diferencialmente abundantes en rizósferas de plantas ruderales y suelos iniciales.....	56
Figura 18. Metagenoma núcleo de <i>S. lycopersicum</i>	58
Figura 19. Esquema representativo de proteínas relevantes en el metagenoma de rizósfera de <i>S. lycopersicum</i>	68

Índice de Tablas

Tabla 1. Propiedades fisicoquímicas de los suelos muestreados, sitios de colecta y producción de biomasa en el experimento de jardín común.....	35
---	----

Las tablas suplementarias se encuentran disponibles en <https://github.com/genomica-fciencias-unam/Barajas-2020>

Tabla S1. Resumen de datos de secuenciación y ensamblado de bibliotecas de amplicones del gen 16S rRNA.	
---	--

Tabla S2. Métricas de diversidad α de los microbiomas

Tabla S3. OTUs únicos y compartidos entre los microbiomas de suelos y raíces de plantas ruderales y *S. lycopersicum*

Tabla S4. Distancias Unifrac entre las muestras de los microbiomas de suelos, rizósferas y endósferas de plantas ruderales y *S. lycopersicum*

Tabla S5. Abundancia relativa a nivel de phylum de los microbiomas de suelos, rizósferas y endósferas de plantas ruderales y *S. lycopersicum*.

Tabla S6. OTUs enriquecidos en los microbiomas de suelos, rizósferas y endósferas de plantas ruderales y *S. lycopersicum*.

Tabla S7. Resumen de secuenciación, ensamblado y anotación de lecturas metagenómicas.

Tabla S8. Clasificación y abundancia de taxonómica de lecturas metagenómicas.

Tabla S9. Métricas de diversidad α de las proteínas predichas en los metagenomas.

Tabla S10. Proteínas únicas y compartidas entre los metagenomas de suelos y raíces de plantas ruderales y *S. lycopersicum*.

Tabla S11. Proteínas enriquecidas en los metagenomas de suelos, rizósferas y endósferas de plantas ruderales y *S. lycopersicum*.

Tabla S12. Proteínas presentes en los metagenomas núcleo de suelos y rizósferas de plantas ruderales y *S. lycopersicum*

Tabla S13. Proteínas únicas y compartidas entre los metagenomas núcleo de suelos y raíces de plantas ruderales y *S. lycopersicum*.

Abreviaciones

SI	Suelo inicial
SF	Suelo Final
US	Suelo sin planta
RT	Rizósfera de <i>S. lycopersicum</i>
RZ	Rizósfera de plantas ruderales
ECT	Endósfera de <i>S. lycopersicum</i>
EC	Endósfera de plantas ruderales
Log2FC	Logaritmo base 2 de las veces de cambio

Palabras clave: Microbioma, metagenoma, rizósfera, selección de microbioma, plantas domesticadas, plantas ruderales, microbioma núcleo.

Agradecimientos

Al Posgrado en Ciencias Bioquímicas de la Universidad Nacional Autónoma de México y al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca otorgada (CVU: 621156).

Al Programa de Apoyo a los Estudios de Posgrado (PAEP) por el apoyo otorgado para la participación en el congreso internacional ISME17 en Leipzig, Alemania.

Este trabajo fue realizado gracias al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) mediante el proyecto número IN221420 y al Proyecto de CONACyT mediante el proyecto número 237387 a cargo del Dr. Luis David Alcaraz Peraza.

Al comité tutorial por sus observaciones puntuales y seguimiento a lo largo de mi formación en el posgrado, Dr. Luis Servín González y Dr. Felipe García Oliva.

Por su colaboración y asesoría experimental, conceptual y revisión meticulosa de la investigación a la Dra. Rocío Cruz Ortega, Dra. Mariana Peimbert Torres, Dr. Luis Servín González y Dr. Felipe García Oliva.

A la M. en C. Shamayim Martínez Sánchez por el trabajo de campo, procesamiento de muestras y los datos de las plantas ruderales. A la M. en C. Jazmín Blaz y a La Q.F.B. Selene Molina por su apoyo en procesamiento de muestras de invernadero y laboratorio. Al Biól. Francisco González por la ayuda en el muestreo de campo, al M. en C. Rodrigo Velázquez por el apoyo en la cuantificación de nutrientes del suelo.

Finalmente, un agradecimiento al doctor Luis David Alcaraz Peraza por la dirección del presente trabajo de tesis.

Resumen

El suelo es considerado uno de los ambientes con mayor diversidad microbiana, la cual es afectada por los diversos factores abióticos que caracterizan a cada tipo de suelo como son su pH, contenido de nutrientes y humedad. De igual manera, factores bióticos como las especies de plantas y otra clase de organismos, como los insectos, también afectan la diversidad del microbioma del suelo. Las raíces de las plantas son colonizadas por bacterias, hongos, artrópodos y anélidos que provienen del suelo en el microambiente de la rizósfera y posteriormente pueden colonizar la endósfera de las raíces a través de una estrecha interacción con la planta. La rizósfera es un ambiente rico en nutrientes, producto de los exudados de raíz de las plantas los cuales varían de acuerdo con su genotipo, encontrándose diferencias entre plantas domesticadas y silvestres. En plantas modelo y cultivables se ha visto una reducción de la diversidad en la rizósfera y aún más acentuado en la endósfera en comparación con el suelo, sugiriendo una especialización del microbioma. Esto ha llevado a la concepción del modelo de dos pasos de adquisición del microbioma de raíces. En este trabajo colectamos 16 suelos con características abióticas contrastantes y raíces de plantas ruderales variadas *in situ*. Los suelos fueron usados como sustrato para el crecimiento de plantas de jitomate (*Solanum lycopersicum*) en un experimento de jardín común para probar la adquisición del microbioma de raíz bajo condiciones controladas en un huésped de genotipo fijo. Secuenciamos 106 bibliotecas de amplicones del gen 16S rRNA de suelos, rizósferas y endósferas; así como 17 metagenomas totales de suelos y rizósferas para evaluar y comparar su diversidad taxonómica y metabólica. Los datos taxonómicos muestran que las plantas domesticadas (*S. lycopersicum*) cumplen con el modelo de dos pasos, mientras que las plantas ruderales presentan un incremento en diversidad α con relación al suelo. A nivel metabólico observamos una mayor diversidad en los metagenomas de suelo en comparación con las plantas, cumpliendo el modelo de dos pasos tanto en las ruderales como las domesticadas. La diversidad β mostró que la composición taxonómica y metabólica de las plantas ruderales y suelos es muy similar, mientras que *S. lycopersicum* tiene un microbioma cohesivo, convergente y diferenciado del suelo y plantas ruderales. Estas diferencias no pudieron ser explicadas por las propiedades abióticas de los suelos. El metagenoma núcleo de *S. lycopersicum* se compone de 51 géneros bacterianos y 2,762 genes que codifican para proteínas, muchas de ellas (8.2%) involucradas en la interacción planta-bacteria. Estos recursos que pueden ser explotados para el diseño de microbiomas específicos de la especie. Las diferencias en la composición y diversidad de los microbiomas puede ser un producto del síndrome de domesticación y el manejo en *S. lycopersicum*, mientras que las plantas ruderales expuestas a la adversidad climática actúan como un oasis de humedad y nutrientes para los microorganismos del suelo.

Abstract

Soil is one of the most diverse microbial habitats. Soil microbial diversity can be affected by its different abiotic properties which are particular for every soil group, like moisture content or pH. Biotic factors, such as plant species and other organisms, such as insects, also alter the microbial diversity of soils. Plants are colonized by bacteria, fungi, arthropod, and annelids in the rhizosphere. Further enhanced interactions between microbes and their plant hosts can be presented in the endosphere of plant roots. The rhizosphere is a nutrient rich environment because of the plants' rhizodeposits which vary according to plant species and between wild and domesticated plants. A loss of microbial diversity in the rhizosphere and endosphere of model and domesticated plants has been accounted for, suggesting the specialization of the microbiome. This fact has led to the two step model of root microbiome acquisition. In this work, we collected 16 soils with contrasting abiotic properties along with in situ growing ruderal plants roots. Soils were used as growing substrates of tomato plants in a common garden experiment to test the microbiome acquisition under controlled and homogeneous conditions using different soil inocula with a single host species. We sequenced 106 16S rRNA gene amplicon libraries from soils, rhizospheres and root endospheres, along with 17 soil and rhizosphere whole shotgun metagenomes. We analyzed and compared the taxonomic and metabolic diversity. Diversity measures of taxonomic data show that domesticated plants (tomato) follow the two step model of root microbiome acquisition, while ruderal plants host larger alpha diversity than soils. At the metabolic level, soils have larger alpha diversity than rhizospheres in both ruderal and domesticated plants, following the two step model. Beta diversity analyses show that the taxonomic and metabolic composition of ruderal plants root and soil microbiomes are very similar, while tomato root microbiomes are cohesive, convergent, and different from soils and ruderal plants. Soil abiotic properties could not explain these differences. Tomato rhizosphere core metagenome is composed of 51 bacterial genera and 2,762 genes coding for proteins, some of them (8.2% core genes) involved in plant-bacteria interactions, resourceful characteristics for species specific microbiome engineering. Finally, differences in the microbiome composition and diversity between ruderals and tomatoes could be a product of the plant domestication syndrome, management in tomatoes and ruderals being a nutrient and moisture oasis under harsh climatic conditions.

Introducción

Propiedades generales de los suelos.

El suelo es un ambiente heterogéneo definido formalmente como: el material mineral u orgánico no consolidado que yace sobre la superficie inmediata de la Tierra y sirve como un medio de crecimiento para plantas terrestres (SSSA, 2016). Este es un producto que surge de las interacciones y alteraciones que suceden en las rocas y sedimentos geológicos por componentes bióticos y abióticos, como la degradación del material debido a la actividad microbiana y la degradación física del material, respectivamente (Sumner, 2000). Las propiedades edáficas son resultado de sus características físicas y químicas. Entre las propiedades físicas se consideran variables como la estructura del suelo, profundidad y disponibilidad de agua, entre otras. Los componentes texturales del suelo (arena, limo, arcilla) afectan cuestiones como la aireación, movimiento de agua y el crecimiento de las raíces de las plantas (Fernández-Illescas *et al*, 2001; Saxton & Rawls, 2006). Entre las propiedades químicas podemos mencionar algunas como el contenido de nutrientes como carbono, nitrógeno y fósforo, el contenido de iones, la salinidad y el pH. Adicionalmente, la ubicación geográfica de los suelos los mantiene sujetos a las condiciones climáticas del sitio que involucran la temperatura y la precipitación de lluvia; variables que son empleadas para la clasificación de cada ambiente mediante el índice de aridez (IA) como el de Lang (Lang, 1915), el cual se calcula como la relación entre la precipitación promedio anual promedio y la temperatura promedio anual en un determinado sitio. Dependiendo de su IA, Lang clasifica a los ambientes en tres categorías: áridos, húmedos e hiperhúmedos.

El INEGI (2007) reporta que México cuenta con 26 de los 30 grupos de suelo reconocidos por el Sistema Internacional de la Base Referencial Mundial del Recurso del Suelo (FAO, 2014). Más del 80% de los suelos en México están clasificados dentro de los grupos de Leptosoles (28.3%), Regosoles (13.7%), Feozems (11.7%), Calcisoles (10.4%), Luvisoles (9%) y Vertisoles (8.6%). Los suelos más fértiles (Feozems, Luvisoles y Vertisoles) ocupan 29.4% del territorio, mientras el 52.4% del territorio se constituye de suelos con bajo aprovechamiento agrícola por ser someros y poco desarrollados como los Litosoles, Regosoles y Calcisoles. A nivel general, los Litosoles se caracterizan por ser suelos poco desarrollados, pedregosos y calcáreos con un pobre potencial agrícola. Los Feozems se forman sobre material no consolidado y se encuentran en climas templados y húmedos, son ricos en materia orgánica y se utilizan para producción de granos y hortalizas. Los Calcisoles

pertenecen a zonas áridas y semiáridas, acumulan calcio e incrementan su potencial agrícola bajo condiciones de riego y fertilización. Los Luvisoles son suelos muy fértiles y su uso agrícola se basa en la producción de grano y caña de azúcar. Los Vertisoles pertenecen a climas semiáridos con estaciones de lluvia y sequía, se encuentran cercanos a lechos lacustres o riberas y tienen un alto contenido de arcilla, que les permite acumular humedad y son muy fértiles. Finalmente, los Regosoles son un grupo de suelos que no entran en ningún grupo clasificado por el Sistema Internacional Base Referencial mundial del Recurso del Suelo (IUSS, 2015).

Índices de diversidad comúnmente empleados en ecología microbiana

En 1972, Whittaker (Whittaker, 1972) elaboró avances sobre los conceptos de diversidad, aplicados y estudiados en comunidades de macro especies, definiendo a la diversidad α , β y γ , haciendo consideraciones sobre la heterogeneidad de especies dentro de una comunidad (α), así como sobre la diferencia en composición entre dos o más comunidades (β) y finalmente, sobre la diversidad de todo un paisaje (γ). La manera más intuitiva de comprender la diversidad en una comunidad es a través del conteo del número de especies (riqueza) en un cierto espacio. Sin embargo, la diversidad no depende exclusivamente de la riqueza de especies, si no que en ocasiones es más relevante comprender la abundancia relativa de cada una de las especies en la comunidad.

El estudio de la diversidad α de las comunidades microbianas se puede llevar a cabo mediante la evaluación de índices de diversidad, los cuales tienen dos componentes: el número total de especies presentes y su distribución o abundancia relativa (Fakhruddin & Bin-Mannan, 2013). Los índices más comúnmente utilizados son el de Shannon (Shannon, 1948) y el de Simpson (Simpson, 1949). El índice de Shannon (H') es una medida de entropía que muestra el nivel de heterogeneidad de una comunidad y se basa en dos parámetros: el número de especies presentes y su abundancia relativa (Magurran, 2004). Este índice representa la equitatividad de la distribución de especies dentro de la comunidad, que es igual a la incertidumbre de la identidad de una especie cuando un individuo es muestreado de forma aleatoria. Por otro lado, el índice de Simpson representa la probabilidad de que dos individuos correspondan a dos especies distintas en un muestreo aleatorio (Magurran, 2004). Por esta razón la selección de estos índices resulta crucial en la interpretación de los datos y pueden ser aplicados para el estudio de la diversidad microbiana. En particular el índice de Shannon resulta de gran utilidad el cual está influenciado por los parámetros de riqueza y

equitatividad, y al mismo tiempo es sensible a la presencia de OTUs en baja abundancia, lo cual brinda más información en sistemas pequeños como micro agregados de suelo o rizósferas (Hill *et al.*, 2003). La cuantificación de la diversidad β se puede realizar mediante análisis multivariados basados en la composición total de especies en una comunidad (Legendre *et al.*, 2005) utilizando coeficientes de disimilitud, como Bray-Curtis o matrices de distancia Unifrac. La disimilitud de Bray- Curtis cuantifica la diferencia en composición de la comunidad entre dos muestras o sitios (Bray & Curtis, 1957), mientras que la métrica de distancia Unifrac se basa en la observación de los integrantes de una comunidad más una medida de la cercanía filogenética de los mismos pudiendo tomar en cuenta la presencia ausencia de éstos o sus abundancias (Lozupone *et al.*, 2007).

El microbioma del suelo: Diversidad y funcionamiento.

El suelo es considerado como el ambiente microbiano más rico y diverso de la tierra, siendo los procariontes los organismos más abundantes, representando el mayor porcentaje de biomasa (Daniel, 2005), encontrándose hasta 10^{11} células bacterianas por gramo de suelo (Roesch *et al.* 2007). Su supervivencia y dinámicas están principalmente influenciadas por la disponibilidad de nutrientes y el contenido de agua del suelo (Kieft *et al.*, 1987). Las bacterias en el suelo son fundamentales para mantener su funcionalidad, ya que éstas llevan a cabo procesos como la degradación de materia orgánica, ciclos de carbono, nitrógeno y fósforo (Garbeva *et al.*, 2004). Además de esto, también están involucradas en el desarrollo de las plantas, ya que dentro del microbioma del suelo encontramos a bacterias promotoras de crecimiento vegetal (Hayat *et al.*, 2010). Muchos microorganismos presentan interacciones de tipo neutro con las plantas; sin embargo, no todos los microorganismos del suelo resultan benéficos para las plantas, muchos de ellos son patógenos fúngicos como *Fusarium oxysporum* que ocasiona marchitez vascular en más de 100 especies de plantas (Ma *et al.*, 2013) o patógenos bacterianos, como *Ralstonia solanacearum* que produce marchitez bacteriana en tomate, papa y otros cultivos (Elphinstone *et al.*, 2005).

La literatura muestra que la diversidad taxonómica bacteriana en suelos a nivel de phylum se ve dominada por los phyla Actinobacteria, Proteobacteria, Firmicutes, Bacteroidetes y Verrucomicrobia (Kuramae *et al.*, 2012). Estas tendencias son similares entre distintos tipos de suelos, pero con variabilidad en la diversidad que se puede encontrar (Lauber *et al.*, 2009), sobre todo a nivel de la estructura de las comunidades al comparar diferentes tipos de biomas como desiertos, bosques templados y pastizales, entre otros

(Fierer *et al.*, 2012). En términos funcionales o metabólicos se sabe que las comunidades microbianas de suelo también presentan diferencias en la composición de los genes codificantes en función del tipo de suelo o las concentraciones de nutrientes como el nitrógeno (Fierer *et al.*, 2012).

Dependiendo del manejo del suelo, el microbioma de éste puede diferir en su composición. Algunos estudios han demostrado que los phyla Proteobacteria y Firmicutes son más abundantes en suelos utilizados para la agricultura bajo un manejo orgánico, encontrándose proteobacterias reportadas como promotoras de crecimiento vegetal pertenecientes al orden Rhizobiales y géneros como *Burkholderia*, *Stenotrophomonas*, *Pseudomonas* y *Sphingomonas* (Upchurch *et al.*, 2008). Por otro lado, los phyla Actinobacteria y Acidobacteria suelen ser más dominantes en suelos de pastizales o sin manejo (Shange *et al.*, 2012). Por otro lado, se puede encontrar un incremento del phylum Acidobacteria a medida que el pH del suelo incrementa su acidez.

El microbioma de plantas y sus estrategias de estudio

Las comunidades bacterianas que habitan en el suelo e interactúan directamente con las plantas y en particular con sus raíces han sido evaluadas desde principios del siglo XX por Lorenz Hiltner, quien en 1904 acuñó el término rizósfera, definido como la capa superficial de suelo (1 mm) unida estrechamente a las raíces (Hartmann *et al.*, 2008). En sus trabajos, remarcaba la importancia de los microorganismos en estos ambientes, ya que observó que conferían resistencia contra patógenos que causan pudrición de la raíz en chícharos y demostrando que pueden ser benéficos para las plantas (Hiltner, 1904). Aunado a esto, se ha tomado también un interés especial por el estudio de los compartimentos endofíticos de las raíces, definidos como los espacios inter o intracelulares de las plantas. Desde entonces distintas estrategias han sido utilizadas para describir y cuantificar la diversidad bacteriana en el sistema suelo-planta, tanto dependientes como independientes de cultivo.

Los huéspedes y su microbioma son sistemas ecológicos conformados por ensamblajes de múltiples especies (Bohannan *et al.*, 2018). En las plantas, el ensamblaje del microbioma está mediado por la selección del huésped y distintos factores ambientales. Sin embargo, la combinación de interacciones ecológicas dentro de las poblaciones de microorganismos, así como la deriva y dispersión de tales poblaciones propicia el establecimiento o desaparición de distintas especies (Fitzpatrick *et al.*, 2020). El entendimiento de estas interacciones es necesario para poder explicar la variación entre el

microbioma de los huéspedes basándonos en la teoría de metacomunidades, la cual considera factores como la escala, la heterogeneidad de las comunidades y la posibilidad de dispersión de los individuos de cada parche o comunidad (Miller *et al.*, 2018). Aunado a esto, se deben de considerar también al banco de especies microbianas que prevalecen y quedan depositadas en el ambiente de los huéspedes (Zobel *et al.*, 1998), en este caso el suelo y agua, principalmente y que funcionan como un reservorio de microorganismos disponibles para recolonizar al huésped vía retroalimentación con el ambiente (Bohannan *et al.*, 2018).

Históricamente se han utilizado distintas estrategias para la caracterización del microbioma del suelo y plantas; empezando por técnicas clásicas dependientes de cultivo de los microorganismos (Barea *et al.*, 1976), posteriormente usando técnicas moleculares como el análisis a través de perfiles de ácidos grasos (Griffiths *et al.*, 2008), análisis de fragmentos de restricción y determinación de ribotipos (Fulthorpe *et al.*, 1998). Diferentes estudios sobre diversidad bacteriana han sido realizados en rizósferas de plantas como *Arabidopsis* (Lundberg *et al.*, 2012; Bulgarelli *et al.*, 2012), cebada (Bulgarelli *et al.*, 2015), arroz (Edwards *et al.*, 2015), maíz (Peiffer *et al.*, 2015), remolacha (Tsurumar *et al.*, 2015) y tomate (Ottesen *et al.*, 2013; Tian *et al.*, 2015), agaves (Coleman-Derr *et al.*, 2015) y cactáceas (Fonseca-García *et al.*, 2016). En general estos estudios han sido desarrollados en plantas modelo o de interés agronómico, encontrando una dominancia de los phyla Proteobacteria, Firmicutes, Acidobacteria y Bacteroidetes, mientras que el compartimento endofítico lo domina el phylum Actinobacteria.

El desarrollo y avance en las tecnologías de secuenciación masiva ha permitido incrementar la cantidad de información disponible de organismos no cultivables en estos ambientes. Entre los enfoques independientes de cultivo se encuentra la metagenómica, la cual consiste en el análisis genómico de las comunidades microbianas totales no cultivadas a partir del DNA aislado directamente del ambiente (Wooley *et al.*, 2010). Este tipo de análisis tiene dos vertientes de forma general. El primero, que se basa en la construcción y secuenciación de bibliotecas de genes marcadores para realizar la descripción taxonómica de la comunidad microbiana total. El uso de amplicones del gen 16S rRNA ha sido el enfoque más utilizado hasta el día de hoy para realizar estudios de microbiología ambiental, ya que esta molécula ha sido indicada como el estándar de oro para la identificación de especies bacterianas (Stackerbrandt & Goebel, 1994). Esta estrategia se basa en la amplificación del gen a partir del DNA metagenómico, lo cual permite la identificación y descripción de las bacterias presentes en el ambiente a nivel de Unidades Taxonómicas Operativas (OTUs) lo cual nos permite obtener una mayor resolución en la información taxonómica obtenida (Kennedy *et al.*, 2014). El segundo enfoque consiste en la secuenciación del DNA total

metagenómico, a partir del cual es posible obtener tanto información taxonómica como funcional, a nivel de predicción de los genes codificantes. En este caso, la información del potencial metabólico y funcional en los metagenomas totales es dependiente de la anotación de las secuencias codificantes predichas contra bases de datos. La base de datos M5nr (Wilke *et al.*, 2012) es una base de datos relacional que compila la información de múltiples bases de datos como NR del NCBI y otras más pequeñas como COG (Tatusov *et al.*, 2000) y subsistemas de SEED (Overbeek *et al.*, 2014) que contienen información con distintos niveles de organización en forma jerárquica o Refseq (Pruitt *et al.*, 2007), la cual proporciona la anotación directa de las proteínas más su taxonomía.

Factores involucrados en la estructuración de las comunidades microbianas de suelos y raíces

El suelo representa la mayor fuente de microorganismos para la colonización de las plantas residiendo en ellos y su vida, desarrollo, estructura y diversidad como una comunidad compleja pueden ser alterados por diferentes factores (Figura I). Estos factores pueden distinguirse en dos categorías: **I)** Factores abióticos, los cuales pueden tener un origen climático, como la aridez de un sitio, o ser características propias de los suelos, como su textura, cantidad de materia orgánica, contenido de nutrientes o su pH, y **II)** Factores bióticos, como son el genotipo o especie de la planta (Marques *et al.*, 2014) o la edad de ésta (Inceoglu *et al.*, 2010).

Entre las propiedades abióticas consideramos los factores edáficos propios de cada tipo de suelo, los cuales juegan un papel importante en la estructuración y diversidad de sus comunidades microbianas, por ejemplo, la textura y el horizonte del suelo determinan micronichos estables (Garbeva *et al.*, 2004), la cantidad y tipo de nutrientes favorecen el desarrollo ciertos grupos bacterianos dependiendo de sus capacidades para metabolizarlos (Kuramae *et al.*, 2012). Así mismo, las concentraciones de nutrientes (C, N, P) y sus relaciones estequiométricas explican diferencias entre suelos con diferentes tipos de manejo (Kuramae *et al.*, 2012; Peralta *et al.*, 2013). El pH, el cual ha sido propuesto como uno de los mejores descriptores de la diversidad de bacterias en suelos (Lauber *et al.*, 2009) y que es producto de la combinación de diferentes propiedades como la salinidad y abundancia de compuestos orgánicos (Al-Busaidi & Cookson 2003; Rukshana *et al.*, 2010). De igual manera, los factores climáticos también tienen una participación relevante en las dinámicas de las comunidades microbianas de los suelos, como se ha visto en el caso de los cambios de

abundancias y actividad de distintos grupos microbianos al aplicar riego (Banerjee *et al.*, 2016) o al comparar suelos de climas áridos y húmedos, donde existe una diferencia significativa entre las proporciones de Actinobacteria y Proteobacteria (Neilson *et al.*, 2017), siendo así la cantidad de humedad un factor determinante de la estructura de las comunidades (Zhang *et al.*, 2012).

Por otro lado, entre los principales factores bióticos que podemos tomar en cuenta, están los compuestos de los exudados radiculares que las plantas depositan en el suelo. Estos son producto de la actividad fotosintética de la planta y pueden representar entre el 5% y 21% del carbono fijado por la planta, constatando finalmente la mayor fuente de carbono y energía para los microorganismos (Marschner *et al.*, 1995). Los exudados de raíz se clasifican en dos clases de metabolitos: exudados de alto peso molecular como los polisacáridos y proteínas, y los exudados de bajo peso molecular, como aminoácidos, ácidos orgánicos, azúcares, compuestos fenólicos y algunos metabolitos secundarios (Badri & Vivanco, 2009) y su composición está relacionada con propiedades intrínsecas de las plantas como el estado de desarrollo de los individuos y la especie de planta (Rovira, 1956).

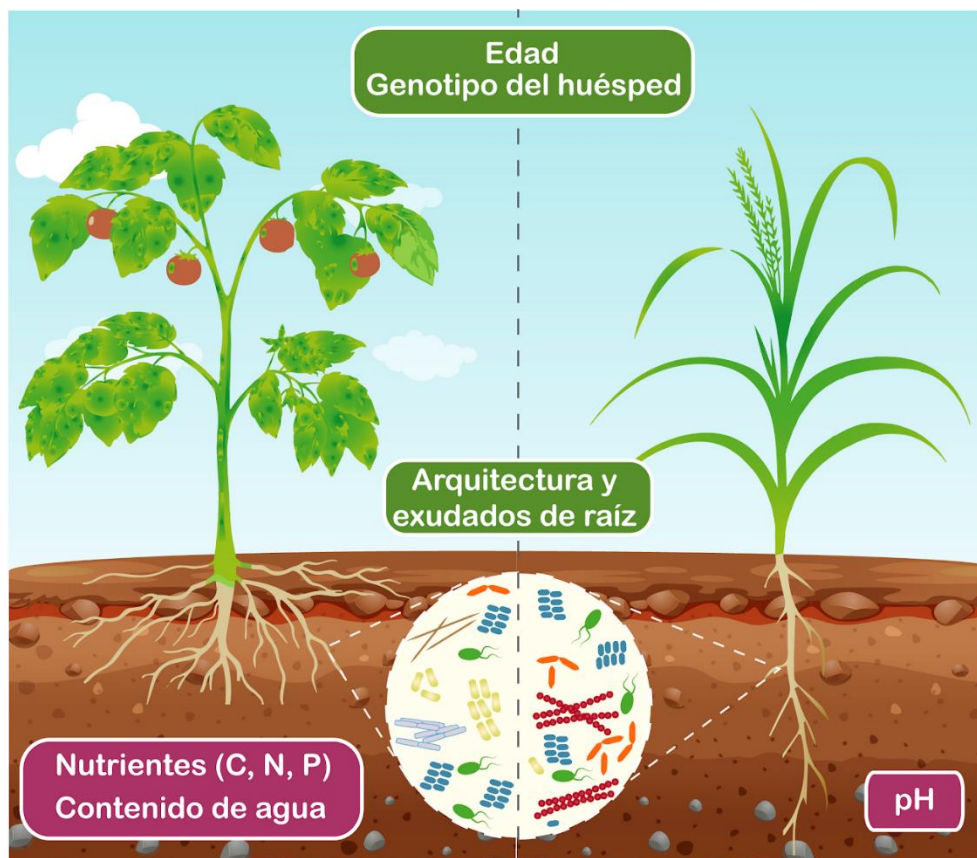


Figura I. La estructura de las comunidades microbianas de suelo es afectada por factores bióticos y abióticos. Entre los factores bióticos (cuadros en verde) se encuentran el genotipo, especie y edad de planta y sus exudados radiculares asociados. Los factores abióticos (cuadros en morado) son propiedades edáficas y climáticas como el pH, contenido de nutrientes y humedad el suelo. La combinación de factores genera diferencias en la estructura de las comunidades microbianas del suelo y raíces. Modificada de Pérez-Jaramillo *et al.*, 2017.

El modelo de dos pasos de adquisición del microbioma de raíz

El microbioma de las raíces de las plantas se conforma principalmente por un subconjunto de microorganismos presentes en el suelo (Garbeva *et al.*, 2004), los cuales son capaces de colonizar en primera instancia la rizósfera y en segundo lugar la endósfera, la cual se define como los espacios inter o intracelulares de las raíces de las plantas. Además de esto, también se deben considerar a otras fuentes de microorganismos que tienen una menor contribución en la composición del microbioma de raíz, como son los que puedan provenir en aerosoles y agua (Graham *et al.*, 2018), así como los integrantes del microbioma endófito de semillas, el cual parece ser transmitido verticalmente (Truyens *et al.*, 2014).

Se ha visto en plantas modelo y plantas domesticadas, como *Arabidopsis thaliana* y arroz (*Oriza sativa*), que la diversidad microbiana disminuye en la zona de la rizósfera en comparación con el suelo. Esta pérdida de diversidad es aún más acentuada en los compartimentos endofíticos de la raíz de las plantas (Bulgarelli *et al.*, 2012, Lundberg *et al.*, 2012; Edwards *et al.*, 2015). En particular, la domesticación de las plantas ha dejado como producto diferencias en la composición general de los microbiomas de raíces en las especies cultivables como frijol y cebada (Pérez-Jaramillo *et al.*, 2018), así como una diversidad disminuida en relación con sus contrapartes silvestres (Martínez *et al.*, 2020). Estas diferencias también se han observado al comparar la diversidad del microbioma rizosférico de *Agave tequilana*, una especie domesticada, contra *A. salmiana* y *A. desertica*, especies silvestres de agave, cuyo componente procariótico es más diverso (Coleman-Derr *et al.*, 2016). Estos patrones han llevado a proponer el modelo de dos pasos de adquisición del microbioma de raíz (Bulgarelli *et al.*, 2015) el cual considera a los factores edáficos y posteriormente a los factores bióticos impuestos por las plantas como filtros para la diversidad del microbioma de raíz (Figura II).

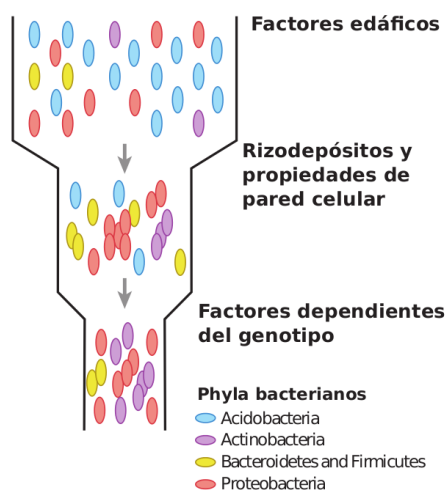


Figura II. Modelo de dos pasos de adquisición del microbioma de raíz. El microbioma de raíz proviene principalmente del suelo. El suelo define a las comunidades microbianas a través de sus propiedades edáficas, principalmente abióticas. Posteriormente el primer paso de selección del microbioma ocurre en la rizósfera, el suelo influenciado por rizodepósitos de las plantas. En un segundo paso, las plantas, de manera activa seleccionan a microorganismos capaces de la colonización e interacción con la planta. Modificado de Bulgarelli *et al.*, 2015.

Los factores edáficos del suelo en conjunción con factores climáticos como la temperatura y precipitación de lluvia son los componentes abióticos que sirven como el primer filtro para la estructuración de sus comunidades microbianas. Por ejemplo, se ha visto que, dependiendo de variables como tipo de suelo, existen ligeras diferencias en la composición del microbioma residente del suelo (Girvan *et al.*, 2003), suelos áridos tienen proporciones más elevadas de Actinobacteria (Naylor *et al.*, 2017) y que el pH influye fuertemente la estructura de las comunidades bacterianas, incrementando a los miembros de Acidobacteria en suelos ácidos (Lauber *et al.*, 2009).

Las raíces de las plantas dentro del suelo generan el microambiente de la rizósfera, el cual se diferencia del suelo adyacente debido a ligeros cambios fisicoquímicos producidos por las células de la rizodermis de las plantas, las cuales secretan una variedad de compuestos químicos como ácidos orgánicos, fitosideróforos, azúcares, vitaminas, aminoácidos y mucílago, entre otros (Dakora *et al.*, 2002). Además de esto, también existen en la rizósfera las células descamadas de la raíz y que son consideradas también como parte de los rizodepositos (Dennis *et al.*, 2010) y que pueden mantenerse con vida hasta por 30 días en el caso del maíz (*Zea mays*) (Vermeer *et al.*, 1982; Somasundaram *et al.*, 2008b). Estas propiedades dependientes de las plantas funcionan como un segundo filtro que selecciona un subconjunto de microorganismos provenientes del suelo circundante como se ha visto en *Arabidopsis*, cuyas rizósferas tuvieron una pérdida de diversidad bacteriana en comparación con el suelo (Lundberg *et al.*, 2012). Finalmente, como tercer filtro están las propiedades intrínsecas de las plantas que tienen que ver con la interacción planta-bacteria y que son reguladas finamente por las plantas para permitir su colonización (Yu *et al.*, 2019) y que genera un decaimiento mucho mayor en la diversidad.

Diferencias en el microbioma de plantas: Efectos de domesticación y la distancia evolutiva de los huéspedes.

La domesticación de las plantas es una actividad humana que se ha llevado a cabo desde el neolítico, hace aproximadamente 10,000, años con el objetivo de obtener plantas que cumplan los requisitos alimenticios de las poblaciones humanas (Purugganan *et al.*, 2009). La selección artificial a través de la domesticación nos ha llevado a producir plantas con fenotipos y genotipos sumamente homogéneos, reduciendo así la diversidad natural de sus poblaciones (Doebley *et al.*, 2006). Alrededor del 75% de la variabilidad genética de plantas se ha perdido debido al abandono de variedades locales y razas de plantas por los

agricultores, los cuales buscan una productividad homogénea en sus cultivos (FAO, 1999b). La domesticación de plantas ha generado repercusiones importantes en su fenotipo que van más allá del síndrome de domesticación (selección de características deseadas y antropocéntricamente seleccionadas) que involucran cambios en perfiles transcriptómicos y metabólicos (Bellucci *et al.*, 2014; Belleggia *et al.*, 2016). Este tipo de modificaciones han resultado en cambios en los perfiles de exudados radiculares en distintos genotipos de trigo (Iannucci *et al.*, 2017).

Debido a que las interacciones planta-bacteria tienen un componente genético microbiano y por parte de la planta, es lógico pensar que la composición de los microbiomas de plantas domesticadas y silvestres puedan contener diferencias contrastantes en cuanto a su composición y funcionamiento (Wissuwa *et al.*, 2009) (Figura III). La composición del microbioma de plántulas de cinco diferentes variedades de maíz, producto de la diversificación de esta planta, mostraron tener una ligera señal marcando las diferencias dentro de cada grupo; en particular, diferencias en las abundancias de algunos miembros de Proteobacteria como *Azospirillum*, *Sphingomonadaceae*, *Rhodopseudomonas*, *Acideocella*, *Variovorax* y *Pseudomonas* (Bouffaud *et al.*, 2014). Otros estudios más actuales corroboran estas tendencias comparando 10 genotipos de maíz y teocintle, concluyendo que la estructura de las comunidades difiere en las variedades cultivables modernas, las cuales se enriquecen con géneros como *Solirubrobacter*, *Novosphingobium*, *Sphingobium*, *Pseudoxantomonas* y *Gemmatimonas*, lo cual puede ser un producto del fenómeno de domesticación (Brisson *et al.*, 2019). En particular, el fenómeno de domesticación de las plantas ha propiciado la pérdida de diversidad genética, lo cual puede repercutir en los perfiles de exudados radiculares, la arquitectura de la raíz y mecanismos de defensa de las plantas, finalmente, propiciando diferencias en la composición del microbioma (Martínez *et al.*, 2020). Análisis realizados con otras plantas muestran que existe una diferencia en la composición del microbioma de plantas domesticadas y sus variedades silvestres, como en el caso del frijol (*Phaseolus vulgaris*), en donde existe una mayor abundancia de Bacteroidetes en las variedades salvajes, mientras que las domesticadas albergan una mayor proporción de Proteobacteria y Actinobacteria (Perez-Jaramillo *et al.*, 2019). Este tipo de diferencias también se han visto en plantas modelo como *A. thaliana* y su pariente lejano *Cardamine hirsuta* (Schlaeppli *et al.*, 2014) y en otros estudios con variedades modernas (*Hordeum vulgare* spp. *vulgare*) y silvestres (*Hordeum vulgare* spp. *spontaneum*) de cebada, donde se reporta que la abundancia de Bacteroidetes también es mayor en las contrapartes no domesticadas (Perez-Jaramillo *et al.*, 2018).

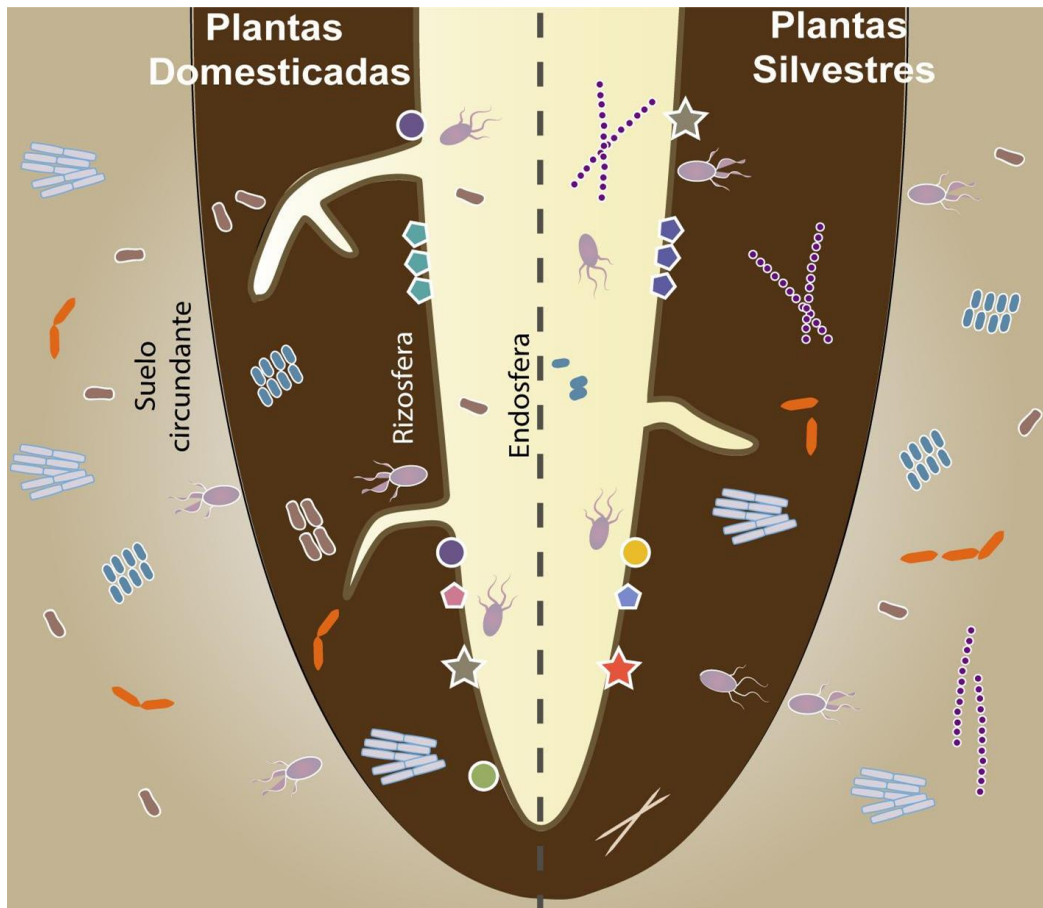


Figura III. Representación esquemática de las diferencias en el microbioma de plantas silvestres y domesticadas. La diferencia en la arquitectura de las raíces y los perfiles de sus exudados en plantas silvestres y domesticadas favorecen la colonización por distintos tipos de microorganismos en la rizósfera. La selección activa en la endósfera genera un microbioma más específico en donde se contemplan también microorganismos heredados verticalmente que no provienen del suelo. Las figuras geométricas representan los distintos rizodepositos de las plantas.

Un aspecto importante por considerar es que, históricamente, las plantas silvestres han estado presentes en condiciones marginales al crecer generalmente en suelos pobres y carentes de un manejo antropogénico, lo cual puede tener implicaciones también en la composición del microbioma de las plantas dado que los factores edáficos influyen en ésta (Chen *et al.*, 2017). Por otro lado, las plantas, al estar seleccionando activamente a los microorganismos que colonizan sus raíces sugiere que los integrantes de su microbiota contengan un repertorio particular de estrategias metabólicas que les permitan la interacción con las plantas. Los microorganismos y las plantas tienen una historia co-evolutiva que se remonta a 400 millones de años (Brundrett, 2002) y al menos 60 millones de años atrás, como es el caso entre plantas que forman nódulos con sus simbiontes microbianos (Sprent *et al.*, 2008). La caracterización de las interacciones planta bacteria se ha hecho en estudios altamente detallados pero limitados que contemplan mecanismos como la nodulación (Peters *et al.*, 1986) o la patogénesis involucrada en los sistemas de secreción tipo III (Hueck *et al.*, 1998). Sin embargo, se han hecho esfuerzos para capturar la diversidad contemplada en la

información genómica de microorganismos asociados a plantas para la búsqueda de otros mecanismos de interacción, encontrando algunos involucrados en la competencia entre microorganismos, la movilidad y quimiotaxis (Levy *et al.*, 2017).

Al basarnos en las diferencias que existen en la composición del microbioma de plantas domesticadas y silvestres, podemos considerar catalogar a los microbiomas de acuerdo con un tipo o especie de planta. Este concepto ha sido utilizado en estudios de microbiomas intestinales humanos para definir poblaciones dependiendo de la composición del microbioma en enterotipos (Lozupone *et al.*, 2012; Arumugam *et al.*, 2011). Una forma práctica de realizar esta tarea es usando el concepto de microbioma núcleo, el cual se puede definir como el conjunto de microorganismos detectados en un conjunto determinado de huéspedes (Lundberg *et al.*, 2012) y que puedan ser utilizados para la optimización de funciones microbianas a nivel individual o en un ecosistema. Los microorganismos en este conjunto no presentan relaciones forzadas de mutualismo, sino que también pueden ser comensales (Toju *et al.*, 2018).

Antecedentes

Dentro de la literatura existen múltiples trabajos en los cuales se ha caracterizado y estudiado la composición de las comunidades microbianas residentes en distintos tipos de suelos provenientes de diferentes regiones geográficas (Fierer *et al.*, 2006; Lauber *et al.*, 2009; Crowther *et al.*, 2019; Fierer *et al.*, 2017), así como los microbiomas de rizósferas de una gran variedad de especies de plantas tanto domesticadas como silvestres (Lundberg *et al.*, 2014; Bulgarelli *et al.*, 2012; Bulgarelli *et al.*, 2015; Coleman-Derr *et al.*, 2016; Perez-Jaramillo *et al.*, 2018). La mayor parte de las investigaciones del microbioma de rizósfera se han centrado en la comparación entre genotipos de una misma especie de planta como en el caso de *Arabidopsis*, arroz y tomate (Lundberg *et al.*, 2012; Schlaeppi *et al.*, 2014, Edwards *et al.*, 2015) o en distintas especies y genotipos de pastos y plantas herbáceas (Bouffaud *et al.*, 2014; Naylor *et al.*, 2017; Fitzpatrick *et al.*, 2018). Sin embargo, todas estas comparaciones han sido realizadas utilizando como máximo dos o tres suelos distintos sobre un par de genotipos de plantas de *Arabidopsis* (Lundberg *et al.*, 2012). Por esta razón, resulta de interés explorar microbiomas rizosféricos de distintas especies de plantas bajo una amplia variedad de tipos de suelos.

Solanum lycopersicum L., el tomate, es una planta de interés agronómico que se cultiva mundialmente, con una producción en 2017 que suma arriba de los 182 millones de toneladas (FAO, 2019). Además de esto, el tomate es una planta cuyo genoma ha sido secuenciado (The 100 Tomato Genome Sequencing Consortium, 2014) y presenta tiempos generacionales cortos, lo cual ha propiciado su uso como organismo modelo para el estudio de su microbioma. El microbioma del tomate se ha caracterizado en hojas, tallos, flores, frutos y raíces, siendo estas últimas dominadas por miembros del phylum Proteobacteria, particularmente las familias Pseudomonadaceae, Rhizobiaceae y Sphingomonadaceae (Ottesen *et al.* 2013, Dong *et al.*, 2019). Otros estudios se han centrado en el análisis del microbioma de raíz dentro del contexto de la presencia de patógenos eucariontes como especies del nematodo *Meloidogyne* (Lamelas *et al.*, 2020) y el oomiceto *Phytophthora parasitica* (Larousse *et al.*, 2017) los cuales fueron probados en un solo tipo de suelo. En el trabajo de Cheng y colaboradores (Cheng *et al.*, 2020), se describieron las comunidades microbianas de rizósferas de un mismo genotipo de *S. lycopersicum* y caracterizaron el microbioma núcleo bajo siete suelos distintos, encontrando que este nicho se encuentra dominado por *Sphingobacteriales*, *Rhizobiales*, *Xanthomonadales*, *Burkholderiales*. De manera similar, otro grupo de trabajo (Lee *et al.*, 2019) caracterizó el microbioma de esta planta dentro de entornos agrícolas, en 23 invernaderos, encontrando un microbioma núcleo

compuesto de 334 OTUs pertenecientes a los géneros *Sphingobium*, *Sphingomonas*, *Streptomyces*, *Arthrobacter* y *Rhizobium*.

Por otro lado, los microbiomas de raíces de plantas ruderales, pastos, también han sido explorados. En general, los sistemas de raíces de los miembros de la familia Poaceae son colonizados preferentemente por miembros de Proteobacteria y Actinobacteria (Fitzpatrick *et al.*, 2018; Marasco *et al.*, 2018). En particular, los géneros *Saccharopolyspora*, *Actinopolymorpha*, *Glycomyces* y *Streptomyces* se han encontrado sobre representados bajo condiciones de sequía en distintos miembros de esta familia de plantas (Naylor *et al.*, 2017). En otra especie de la familia Poaceae (*Panicum virgatum* y *P. hallii*) se han encontrado sobrerrepresentados a los géneros *Bradyrhizobium*, *Streptomyces* y *Enterobacter*, formando parte del microbioma núcleo de la especie en dos tipos de suelo distintos (Singer *et al.*, 2019). Al comparar microbiomas de pastos y plantas herbáceas, Vieira y colaboradores (Vieira *et al.*, 2019) también reportaron un incremento de los géneros *Streptomyces*, *Pseudomonas*, *Mesorhizobium*, y *Rhizobium*, pertenecientes a los phyla Actinobacteria y Proteobacteria.

En última instancia, las comunidades microbianas tanto de suelos como de las raíces de plantas también han sido descritas a nivel funcional a través de la secuenciación del DNA metagenómico total. Este tipo de estudios en suelos se han enfocado en el descubrimiento de productos naturales con aplicaciones biotecnológicas (Daniel, 2004; Reddy *et al.*, 2012; Katz *et al.*, 2015). También se han realizado comparaciones entre suelos desérticos y no desérticos, en donde se observaron diferencias en la proporción de genes microbianos relacionados a la tolerancia a estrés, metabolismo de aminoácidos y producción de esporas en los suelos desérticos, mientras que los suelos no desérticos tienen mayor abundancia de genes de degradación de compuestos orgánicos complejos (Fierer *et al.*, 2012). De igual forma, se han descrito las características metabólicas y funcionales del metagenoma de *Agave tequilana* dentro de este tipo de ambientes, encontrándose que las rizósferas de estas plantas se caracterizan por tener una alta abundancia del gen *nifH* y genes de producción de biofilm (Flores-Núñez *et al.*, 2020). Por otro lado, las funciones metabólicas codificadas en los metagenomas de rizósferas de plantas como *Vaccinium angustifolium* (arándanos) y diferentes especies del género *Citrus* muestran mayores proporciones de genes relacionados a la degradación de carbohidratos complejos, metabolismo de aminoácidos, formación de biofilm y genes involucrados en respuestas de quimiotaxis y movilidad (Yurgel *et al.*, 2019; Xu *et al.*, 2018). En especies de plantas invasoras como *Acacia dealbata* se encontraron genes microbianos sobrerrepresentados que tienen que ver con el metabolismo de nitrógeno, carbohidratos y vitaminas, así como sistemas de transporte membranal (Kamutando *et al.*, 2018). Otros trabajos se han centrado en la selección del microbioma de soya a nivel

metagenómico (Mendes *et al.*, 2014) y los efectos del monocultivo de plantas medicinales como *Rhemanía glutinosa* donde observaron un incremento de genes implicados en la resistencia a antibióticos por bombas de eflujo y transportadores ABC, así como degradación de compuestos xenobióticos (Wu *et al.*, 2018).

Planteamiento del problema

La composición y diversidad del microbioma de la raíz de las plantas puede variar dependiendo de distintos factores bióticos y abióticos. Las propiedades abióticas del suelo, como la concentración de nutrientes o el pH, varían en cada tipo de suelo y tienen un papel en la estructuración de las comunidades microbianas del mismo, resultando en distintos conjuntos de especies microbianas asociadas a cada suelo. Las comunidades microbianas del suelo representan el principal aporte de microorganismos que colonizan las raíces de las plantas, en las cuales generalmente disminuye la diversidad, lo cual llevó a la concepción del modelo de dos pasos de adquisición del microbioma de raíces que considera la reducción de diversidad microbiana: suelo > rizósfera > endósfera. Por otro lado, cada especie de planta difiere en la composición de exudados radiculares que genera, actuando como una presión selectiva que afecta la composición del microbioma de raíz. Sin embargo, al usar dos tipos de suelo diferentes y un mismo genotipo de planta, se han observado diferencias en la composición del microbioma de raíz. Por estas razones resulta de interés comprender el rol de los factores bióticos y abióticos en la estructura y diversidad del microbioma de raíz. Específicamente, múltiples suelos con propiedades fisicoquímicas contrastantes (pH y contenido de C, N, P) como propiedades abióticas y como propiedades bióticas la comparación entre especies de plantas ruderales en campo y cultivables (*S. lycopersicum*) en un experimento de jardín común homogeneizando las condiciones para los suelos.

Hipótesis

Los factores bióticos (naturaleza del hospedero) tendrán preponderancia sobre los factores abióticos (concentración de nutrientes y pH) en la estructura de las comunidades bacterianas de raíz.

El modelo de dos pasos de adquisición del microbioma se cumplirá a nivel taxonómico y metabólico tanto en plantas ruderales como en las domesticadas.

El microbioma de raíz presentará un comportamiento convergente en plantas de un mismo genotipo (*S. lycopersicum*) a comparación de las plantas de diversas especies.

Objetivo General

Evaluar las diferencias en la composición, diversidad y estructura taxonómica y metabólica de comunidades microbianas de distintos tipos de suelos, raíces de plantas ruderales colectadas *in situ* y raíces de plantas domesticadas (*Solanum lycopersicum*) en un experimento de jardín común.

Objetivos Particulares

Realizar una colecta de distintos tipos de suelos y raíces de plantas ruderales creciendo *in situ* con el objetivo de capturar una amplia diversidad de suelos y plantas.

Realizar un experimento de jardín común para cultivar plantas de *S. lycopersicum* en los suelos colectados para evaluar la composición del microbioma de raíz en una sola especie de planta bajo distintos inóculos microbianos.

Describir las diferencias entre la diversidad taxonómica y funcional predicha de los microbiomas de suelos, rizósferas y endósferas de raíces de las plantas ruderales y plantas de *S. lycopersicum*.

Evaluar la selección preferencial del microbioma dependiente de la especie del hospedero a través del análisis de enriquecimiento de OTUs y genes codificantes en los microbiomas de suelos y raíces de plantas ruderales y *S. lycopersicum*.

Caracterizar del microbioma y metagenoma núcleo de raíz de *S. lycopersicum* para conocer las características compartidas dentro del microbioma de una especie cultivable de plantas.

Probar la influencia de los factores abióticos del suelo en la composición del microbioma de raíces.

Metodología

Muestreo de suelos y raíces de plantas ruderales

Se utilizó la carta edafológica de la República Mexicana, (INEGI, 2014) para planear puntos de muestreo en la zona centro-oriente del país en los cuales se colectaron 16 muestras de distintos suelos correspondientes a 7 distintos tipos de suelo de acuerdo con la clasificación de la FAO (IUSS, 2015) y un jal minero (desechos de producción minera) en el cual existían comunidades de plantas establecidas. Se tomaron los siguientes datos en cada sitio: Fecha de colecta, coordenadas, altitud, observaciones generales del sitio de muestreo, fotografías del sitio, hora de colecta y el tipo de suelo fue determinado usando las coordenadas de la carta edafológica (Figura 1, Barajas *et al.*, 2020). Adicionalmente, se calculó el índice de aridez de Lang (Lang, 1920) para cada sitio, basado en la temperatura y precipitación anual promedio, datos consultados en el centro de ciencias de la atmósfera UNAM (<http://uniatmos.atmosfera.unam.mx/ACDM/>).

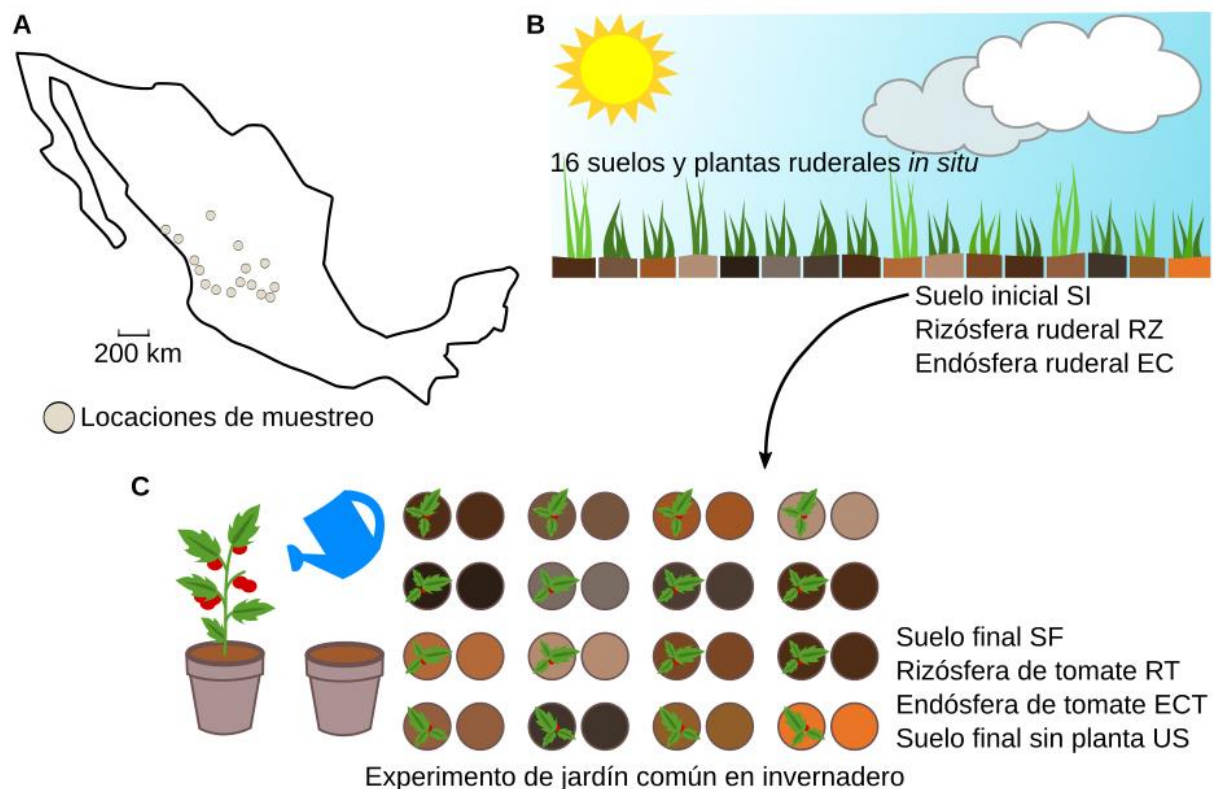


Figura 1. Resumen gráfico del trabajo. A) Sitios de muestreo de suelos y plantas ruderales **B)** Colecta de distintos tipos de suelos utilizados para el experimento de jardín común (SI) y raíces de plantas locales para la obtención de DNA metagenómico de la rizósfera (RZ) y endósfera (EC). **C)** Experimento de jardín común con *S. lycopersicum* bajo condiciones controladas en invernadero (riego y temperatura homogéneos) para la obtención de DNA metagenómico de rizósferas (RT) y endósferas (ECT) de *S. lycopersicum*, suelos al final del experimento (SF) y suelos control sin presencia de planta (US).

En cada sitio se colectaron 2 kg de suelo a una profundidad máxima de 10 cm. La toma de muestras de suelos y raíces de plantas ruderales fueron hechas utilizando guantes y las palas utilizadas se lavaron y desinfectaron con etanol al 70% al iniciar y finalizar cada colecta. Los suelos se guardaron en bolsas de plástico y fueron conservados a 4°C en oscuridad. Un kilo de suelo se utilizó para realizar el análisis de la composición fisicoquímica en el IIES, UNAM, campus Morelia. El segundo kilo fue utilizado para llevar a cabo la siembra de plantas de *S. lycopersicum* en condiciones de invernadero quince días después de la recolección del suelo (Figura 2). Adicionalmente, se tomaron muestras de suelo por duplicado en tubos cónicos de centrifuga estériles de 50 mL, los cuales fueron congelados en nitrógeno líquido para la posterior extracción de DNA.

Experimento de jardín común y obtención de muestras

Se realizó un experimento de jardín común, en el cual se sembraron plantas de *S. lycopersicum* en cada uno de los suelos colectados (SI). Se utilizaron semillas de *S. lycopersicum* L. Cv. *Río grande* (Sun Seeds, Parma, ID, USA), las cuales fueron superficialmente desinfectadas con detergente y NaOCl al 2% y germinadas en cajas de Petri con agar al 1% en oscuridad durante 96 horas, de acuerdo con el protocolo descrito en Barajas (2016, tesis de maestría). En el invernadero se colocaron de manera aleatoria macetas por duplicado conteniendo un volumen aproximado de 300 mm³ de cada uno de los suelos colectados. Cinco plántulas fueron transferidas asépticamente a las macetas y tras su establecimiento (25 días), se dejaron a 2 individuos por maceta, sumando 4 réplicas biológicas en cada uno de los suelos. Adicionalmente, se colocaron macetas sin plantas. Todas las macetas recibieron riego cada dos días durante los 60 días de crecimiento, tiempo tras el cual se cosecharon las plantas y se obtuvieron las muestras de suelos finales (SF) y suelos control, sin plantas (US) en tubos Falcon estériles de 50 mL. Las raíces de las plantas fueron sacudidas para dejar el suelo rizosférico, separadas de las partes aéreas y almacenadas en tubos Falcon estériles de 50 mL para su posterior procesamiento.

Obtención de muestras de rizósfera y endósfera de raíces de *S. lycopersicum* y plantas ruderales

Para la obtención de rizósferas y endósferas se siguió un protocolo publicado previamente (Lundberg *et al.*, 2012), en el cual, las raíces fueron lavadas con 30 mL de buffer

de fosfatos 1x (PBS) estéril para desprender el suelo rizosférico (RT). Posteriormente, las raíces fueron transferidas a nuevos tubos Falcon estériles con 30 mL de PBS 1x, los tubos fueron sonicados en un baño sonicador para obtener las endósferas (ECT) y finalmente, se apartaron las raíces para juntarlas con sus partes aéreas correspondientes. Las plantas fueron secadas a 70 °C para la determinación de biomasa producida en términos de peso seco.

Los pellets de rizósfera y endósfera fueron recuperados mediante la centrifugación de los tubos de 50 mL a 1,300 g durante 10 minutos. En el caso de las plantas ruderales, el protocolo fue modificado debido a la baja eficiencia en la extracción de DNA. Las raíces de plantas ruderales fueron segmentadas y colocadas en 10 tubos Eppendorf de 1.5 mL, los cuales recibieron el mismo tratamiento de lavado para obtener la rizósfera (RZ) y sonicación para obtener la endósfera (EC). Todos los pellets fueron congelados a -80 °C hasta la extracción de DNA.

Análisis fisicoquímicos de suelos

Los suelos iniciales y finales fueron sujetos a la caracterización de los siguientes parámetros fisicoquímicos: pH, nitrógeno total (NT), fósforo total (PT), carbono total (CT).

Las muestras de suelos se secaron en un horno a 75 °C durante 24 horas, se eliminó la materia orgánica (raíces y hojas) y arenas grandes, y finalmente, se molió cada muestra en un mortero de ágata. El pH fue medido en agua desionizada (1:4 w:v) con un medidor digital de pH Corning. El carbono total fue determinado por el método coulométrico de combustión directa (Huffman, 1977) en un coulómetro UIC CM5012, el nitrógeno total fue cuantificado por medio del método semi Kjeldahl y el fósforo total fue cuantificado por el método colorimétrico de molibdato después de una reducción con ácido ascórbico (Murphy & Riley, 1962). El NT y PT se cuantificaron en un autoanalizador Bran-Luebbe III (Norderstedt, Germany). Las propiedades fisicoquímicas de SI y SF fueron representadas a través de un análisis de escalamiento no métrico multidimensional (NMDS).

Extracción de DNA metagenómico

El DNA de rizósferas, endósferas, suelos iniciales, finales y suelos control sin plantas fue extraído con el *kit PowerSoil DNA Isolate Kit* (MoBio Laboratories, Solana Beach, CA,

USA) siguiendo el protocolo provisto por el fabricante, modificando el paso final de elución, en el cual se calentó la solución de elución C6 a 60 °C y se realizaron dos eluciones de 30 uL. Para la extracción de DNA metagenómico de suelos, se pesaron aproximadamente 0.25 g de suelo. Para la extracción de DNA de rizósferas y endósferas, los pellets se resuspendieron en 200 uL de PBS estéril y fueron transferidos a los tubos de extracción del kit *PowerSoil*. El mismo DNA fue utilizado para la preparación de bibliotecas de amplicones del gen ribosomal 16S rRNA y la secuenciación de metagenomas totales. La integridad del DNA fue evaluada mediante geles de agarosa al 0.8% en buffer TBE y posteriormente cuantificados con un equipo Nanodrop lite (Nanodrop Technologies, Wilmington, DE).

Bibliotecas metagenómicas y secuenciación de amplicones del gen 16S rRNA y metagenomas totales

La amplificación por PCR del gen ribosomal 16S rRNA se realizó por duplicado a partir de las extracciones de DNA metagenómico de suelos, rizósferas y endósferas utilizando el par de primers 341F/805R (341F: 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACCTACGGGNGGCWGCAG-3'; 805R: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACTACHVGGGTATCTAATCC 3'), el cual amplifica las regiones hipervariables V3-V4 del gen ribosomal 16S (Muyzer *et al.*, 1993). Las reacciones de PCR se llevaron a cabo en volúmenes de 20 uL. por duplicado, usando el siguiente programa para la amplificación: un ciclo a 95 °C por 3 min, seguido de cinco ciclos a 94 °C por 30 s, 55 °C por 30 s y 68 °C por 30 s, finalizando con 25 ciclos a 94 °C por 5s y 68 °C por 30 s. Posteriormente, los amplicones se purificaron por medio de columna con el kit Wizard SV Gel and PCR Clean-Up System (Promega).

Las bibliotecas de amplicones del gen ribosomal 16S rRNA fueron secuenciadas en la plataforma Illumina® Miseq™ en una configuración 2 x 300 *paired-end* en la Unidad de Secuenciación Masiva y Bioinformática del Instituto de Biotecnología de la UNAM, México. Las bibliotecas de secuenciación para metagenomas totales suelos y rizósferas de *S. lycopersicum* y plantas ruderales fueron preparadas con el kit de preparación Truseq PCR free y secuenciadas en un equipo Illumina HiSeq 2000 (100 pb) en las instalaciones de MacroGen, Korea (<https://www.macrogen.com>). Los datos crudos de secuenciación de las bibliotecas se encuentran disponibles en la base de datos del *Sequence Read Archive* (SRA) del Centro Nacional de Información Biotecnológica (NCBI) bajo los números de Bioproyectos: PRJNA603586, PRJNA603590 para las bibliotecas de amplicones y el PRJNA603603 para

los metagenomas totales. Todos los protocolos detallados de laboratorio e invernadero pueden ser consultados en la tesis de maestría (Barajas, 2016) en la siguiente liga http://oreon.dgbiblio.unam.mx/F/X7IMYKQF8QQU3RERVMXETIRKRN18D5YBIRX4RYXI5H42CBQJM1-51240?func=full-set-set&set_number=017121&set_entry=000001&format=999

Procesamiento de lecturas de bibliotecas de amplicones

La calidad de las lecturas de secuenciación se inspeccionó con FastxToolkit (http://hannonlab.cshl.edu/fastx_toolkit/) y se recortaron a una longitud de 250 pb debido a la deficiencia de calidad en el extremo 3'. Las secuencias recortadas se ensamblaron con Pandaseq (Masella *et al.*, 2012) usando un empalme mínimo de 15 pb, un mínimo de longitud de secuencia de 250 pb y máximo de 470 pb para las secuencias ensambladas y un valor de certidumbre del 95%. Las secuencias ribosomales ensambladas fueron concatenadas en un solo archivo y agrupadas en unidades taxonómicas operativas (OTU) al 97% de identidad usando *cd-hit-est* (Li *et al.*, 2006), obteniendo una secuencia representativa de cada OTU asignado. La taxonomía de las secuencias representativas fue asignada con la base de datos de Greengenes (De Santis *et al.*, 2006) usando el script de QIIME (Caporaso *et al.*, 2010) *parallel_assign_taxonomy_blast.py* con los parámetros default y una tabla de abundancia de OTUs se generó con el script *make_otu_table.py*. Las secuencias que representaban *singletons* y las que fueron asignadas como mitocondrias, cloroplastos o que no tuvieron aciertos contra la base de datos fueron descartadas. Las secuencias filtradas fueron alineadas con SSU-align (Nawrocki, 2009) y un árbol filogenético fue generado con el programa Fasttree (Price *et al.*, 2009) para posteriormente calcular distancias Unifrac ponderadas.

Asignación taxonómica de plantas ruderales

La identidad de las especies de plantas ruderales fue recuperada desde las secuencias de amplicones mediante el alineamiento contra genes ribosomales de plantas los cuales fueron extraídos y posteriormente se realizaron búsquedas en la base de datos NR del NCBI. La asignación taxonómica se hizo hasta el nivel de familia (Martínez, tesis de maestría 2017).

Procesamiento de lecturas de metagenomas totales

El control de calidad y remoción de secuencias de adaptadores de las lecturas de secuenciación de los metagenomas totales se hizo con el software Trimmomatic (Bolger *et al.*, 2014) utilizando los parámetros LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:36, obteniendo sólo las lecturas pareadas que se utilizaron para los procesos y análisis subsecuentes. Para eliminar secuencias provenientes de los hospederos, las lecturas metagenómicas de las bibliotecas de rizósfera de *S. lycopersicum* fueron alineadas contra el genoma de *S. lycopersicum* y las lecturas de bibliotecas de rizósferas de plantas ruderales y suelos fueron alineadas contra el genoma de *Oryza sativa* como representante de la familia Poaceae con el programa Bowtie2 (Langmead & Salzberg, 2012). Las secuencias resultantes fueron ensambladas con el programa metaSPADES (Nurk *et al.*, 2017). Posteriormente, las lecturas fueron mapeadas contra los contigs resultantes del ensamblado con el software BMap (Bushnell, 2014). Las lecturas que no lograron ser alineadas contra los contigs fueron sujetas a un segundo ensamble con el programa Velvet (Zerbino *et al.*, 2008).

Predicción de proteínas y anotación

Los contigs resultantes de ambos ensamblados fueron unidos y utilizados para realizar la predicción de marcos de lectura abiertos y proteínas codificantes con Prodigal (Hyatt *et al.*, 2010). La base de datos M5NR (Wilke *et al.*, 2012) fue utilizada para realizar la anotación de las proteínas predichas mediante alineamientos con el programa DIAMOND (Buchfink *et al.*, 2015) solicitando reportar alineamientos con un valor de *e-value* mínimo de $1e^{-10}$. Posteriormente, se obtuvieron los mejores aciertos contra la base de datos M5NR filtrando los mejores valores de *bitscore* y eliminando los valores duplicados en las tablas de resultados de DIAMOND. Los identificadores md5sum del M5NR correspondientes a los mejores aciertos fueron utilizados para extraer las anotaciones de las bases de datos de subsistemas de SEED (Overbeek *et al.*, 2014) y Refseq (Pruitt *et al.*, 2007), obteniéndose así dos archivos de sistemas distintos de anotación distintos para la descripción de la diversidad metagenómica. Adicionalmente, se realizó la anotación taxonómica de las lecturas metagenómicas mediante binning con el programa Kaiju (Menzel *et al.*, 2016).

Cálculo de abundancias de proteínas

Para calcular la abundancia de los genes que codifican para las proteínas predichas se utilizaron los scripts *hitter.py* y *hitter_table.py*, disponibles en https://github.com/genomica-fciencias-unam/tabla_genes. Las lecturas metagenómicas de calidad fueron mapeadas por medio de alineamientos contra los ORFs predichos utilizando Bowtie2, filtrando únicamente los alineamientos de calidad, obteniendo tablas de conteos de lecturas alineadas contra las secuencias de referencia. Posteriormente, se compararon las tablas de los mejores aciertos contra las tablas de conteos de proteínas anotadas para obtener tablas de abundancia de cada una de las muestras. Para recuperar el total de información metagenómica no incluida en los conteos de proteínas anotadas, las secuencias de proteínas predichas sin aciertos contra la base de datos M5NR fueron extraídas y concatenadas en un archivo, a partir del cual se realizó un agrupamiento en familias de proteínas al 70% de identidad con el programa cd-hit (Li *et al.*, 2006) con los parámetros tamaño de palabra 4, y cobertura del 70% del alineamiento. Se obtuvo una tabla de secuencias representativas de cada familia de proteínas, la cual se comparó contra los archivos de mapeo de lecturas para poder estimar la abundancia de las proteínas hipotéticas sin homólogos en las bases de datos.

Análisis de diversidad taxonómica y metabólica

Las métricas de diversidad α y β fueron calculadas utilizando la paquetería phyloseq (McMurdie *et al.*, 2013) y vegan (Oksanen *et al.*, 2015) en R (R Core Team). Como métricas de diversidad α taxonómica se calcularon el número de OTUs observados y el índice de diversidad de Shannon para cada muestra de rizósferas y endósferas de plantas locales o jitomates (RZ, RF, EC y ECT, respectivamente), así como para los suelos inóculo (SI), suelos finales (SF) y suelos control sin plantas (US). La composición de los microbiomas de se describió a nivel de phylum (Proteobacteria se describió a nivel de clase) en cada muestra y se hicieron cálculos de ANOVA para determinar grupos taxonómicos que fueran estadísticamente más abundantes entre los siguientes grupos de comparación: I) Raíces de *S. lycopersicum* (RF + ECT), II) raíces de plantas ruderales (RZ + EC) y III) suelos (SI + SF + US).

Para explorar la diversidad taxonómica β se calculó la distancia Unifrac ponderada entre las muestras de SI y SF, RZ, RFT, EC y ECT partiendo de las tablas de conteos de OTUs. Las muestras fueron agrupadas de manera jerárquica con la función *hclust* (*method*:

complete) de acuerdo con la matriz de distancia y se graficó como un dendrograma. La validez de los agrupamientos fue evaluada a través de la función ANOSIM que analiza la similitud entre las muestras basado en la matriz de distancias. Además de esto, se calcularon las distancias entre las muestras a partir del dendrograma con la función *cophenetic* de paquete *vegan* (Oksanen *et al.*, 2015).

A partir de las tablas de taxonomía, los OTUs fueron conglomerados a nivel de género y se compararon los perfiles taxonómicos de los grupos de raíces de *S. lycopersicum*, raíces de plantas ruderales y suelos utilizando las listas de géneros para generar diagramas de Venn a través del software en línea (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). Los OTUs fueron considerados como presentes dentro de uno de los grupos con estar presente en una sola muestra.

La diversidad α (riqueza de proteínas e índice de diversidad de Shannon) de proteínas predichas fue calculada tomando en cuenta el total de proteínas anotadas (aciertos contra Refseq) y no anotadas. La diversidad metabólica β fue evaluada a partir de una matriz de disimilitudes con la distancia Bray-Curtis, calculada a partir de las abundancias de proteínas predichas anotadas con la base de datos Refseq y proteínas no anotadas mediante un Análisis de Coordenadas Principales (CAP) constreñido por las variables bióticas y abióticas de concentraciones de nutrientes (C, N y P), pH y el tipo de muestra (SI, RF, RZ).

Identificación del microbioma y metagenoma núcleo de raíces de *S. lycopersicum* y plantas ruderales

La composición del microbioma y metagenoma núcleo de las rizósferas de plantas ruderales y plantas de *S. lycopersicum* se obtuvo comparando las listas de OTUs colapsados a nivel de género en el caso del microbioma y con los identificadores de las proteínas anotadas y no anotadas en el total de muestras de cada grupo (SI, RF, RZ) mediante la paquetería UpSet (Conway *et al.*, 2017) en R (R Core Team, 2017).

Descripción del potencial metabólico de comunidades microbianas de rizósferas

Se realizó la descripción metabólica de los metagenomas de suelos y rizósferas de plantas ruderales y *S. lycopersicum* utilizando las proteínas anotadas bajo el sistema

jerárquico de anotación de Subsistemas del SEED de acuerdo con el nivel 1 de la clasificación. Se realizaron pruebas de análisis de varianza (ANOVA) de cada categoría del nivel 1 en todas las muestras y se aplicó una prueba de Tukey (TukeyHSD) para corroborar las diferencias estadísticas entre cada una de las categorías analizadas.

Análisis de abundancia diferencial de OTUs y proteínas

Se realizaron comparaciones pareadas entre cada grupo de microbiomas de suelos (SI, SF, US) contra cada grupo de rizósferas (RZ y RF) y endósferas (EC y ECT) basadas en los conteos de OTUS para los microbiomas y de proteínas anotadas por Refseq para los metagenomas utilizando el paquete DESeq2 (Love *et al.*, 2014) en R, aplicando la prueba estadística de Wald. En el caso de los OTUs se consideró una $p < 0.01$ para considerar a un OTU, los cuales fueron descritos a nivel de género. En el caso de las proteínas anotadas, se utilizó un valor de corte de $p < 0.001$ y fueron descritos al nivel de función. Todos los protocolos bioinformáticos descritos en los métodos pueden ser consultados en el sitio <https://github.com/genomica-fciencias-unam/Barajas-2020>.

Resultados

Descripción biogeoquímica de suelos, diversidad de plantas ruderales y producción de biomasa en plantas de *S. lycopersicum*

Para obtener los inóculos que serían utilizados como fuente de microorganismos para las raíces de plantas de *S. lycopersicum* en el experimento de jardín común se realizó la colecta de un total de 16 muestras de suelos que corresponden a 7 tipos de suelos y un jal minero, el cual es un substrato producto de desechos de minería y no es estrictamente un suelo (Figuras 1 y 2; Tabla 1). Para conocer la diversidad y composición fisicoquímica de las muestras de suelos inóculo colectados en campo y los cambios en dichas propiedades posterior al cultivo de plantas de *S. lycopersicum* en el invernadero, el cual involucró un tratamiento de riego uniforme, se cuantificaron las concentraciones de nutrientes totales (C, N, P) y se midió el pH.



Figura 2. Diversidad ambiental y geográfica de los sitios de colecta. A) Cada sitio de colecta se fotodocumentó y se colectaron distintos metadatos como la temperatura y tipo de vegetación. El sitio SIN1 no se presenta debido a que era de noche. **B)** Carta edafológica del territorio mexicano donde se ubican los sitios de colecta (INEGI, 2020).

Los suelos colectados se clasificaron en tres grupos dependiendo de sus características de acidez. Siete suelos son alcalinos ($\text{pH} > 8.0$), seis fueron clasificados como neutros ($7.5 < \text{pH} < 6.5$) y 3 como ácidos ($\text{pH} < 6.5$). Después del experimento de jardín común, sólo dos suelos cambiaron sus características de acidez, ZAC1 volviéndose neutro y JAL4 volviéndose ácido. Las concentraciones de nutrientes en los suelos también sufrieron cambios al comparar las muestras del inicio y final del experimento de jardín común. Las concentraciones de N tuvieron incrementos en 11/16 de las muestras, siendo el suelo de la locación SIN2 el que presentó la mayor concentración de N y al mismo tiempo, la mayor disminución del nutriente al final del experimento. El C total presentó incrementos en 10/16 de las muestras, de las cuales, la mayor concentración e incremento fue en el suelo de la locación SLP1. Finalmente, las concentraciones de P total tuvieron incrementos en 9/16 de las muestras. La información de las propiedades fisicoquímicas se resume en la Tabla 1 (Barajas *et al.*, 2020).

Sitio	Tipo de Suelo	pH (SI)	pH (SF)	N total (mg/g) (SI)	N total (mg/g) (SF)	C total (mg/g) (SI)	C total (mg/g) (SF)	P total (mg/g) (SI)	P total (mg/g) (SF)	Índice de Aridez	Altitud	Latitud	Longitud	Fecha de colecta	Biomasa de tomate (peso seco en g)
JAL4	Phaozem	7.2	6.0	0.623	11.721	16.503	31.003	0.009	0.16	48.8	1490	20.849	-103.855	11/8/2014	0.221 ± 0.14
NAY2	Cambisol	5.7	6.1	6.83	16.667	19.347	22.086	0.453	0.284	56.5	38	21.836	-105.083	11/8/2014	0.083 ± 0.04
JAL3	Luvisol	6.5	6.8	3.99	17.799	32.863	25.683	0.475	0.263	46.1	1862	20.821	-102.798	11/6/2014	0.090 ± 0.07
NAY3	Cambisol	6.9	6.9	2.027	11.584	9.924	13.569	0.157	0.197	47.1	4	22.142	-105.261	11/8/2014	0.204 ± 0.01
SLP1	Kastanozem	6.9	7.8	9.702	5.898	79.348	146.37	0.447	0.287	22.1	2130	21.988	-101.277	11/10/2014	0.767 ± 0.14
ZAC1	Cambisol	5.0	7.2	6.433	12.961	4.921	5.117	0.324	0.347	25	2221	22.818	-102.696	11/10/2014	0.017 ± 0.02
SIN1	Phaozem	7.3	7.4	7.523	2.095	16.862	24.452	0.377	0.42	29.8	23	22.988	-105.880	11/8/2014	0.223 ± 0.13
JAL5	Lithosol	7.4	7.9	2.158	24.011	5.067	5.427	0.235	0.192	44	1328	21.179	-104.540	11/8/2014	0.199 ± 0.04
JAL1	Planosol	7.5	7.6	1.892	1.339	9.482	14.425	0.118	0.164	32.9	1810	21.162	-101.853	11/6/2014	0.313 ± 0.04
AGS1	Planosol	8.0	8.3	5.671	12.46	19.616	22.01	0.268	0.28	29.2	2059	21.828	-102.121	11/10/2014	0.269 ± 0.06
GTO1	Vertisol	8.3	8.6	12.298	13.387	22.733	18.546	0.378	0.323	32.9	1708	20.580	-100.947	11/11/2014	0.259 ± 0.08
GTO3	Vertisol	8.4	8.4	2.624	6.008	10.559	10.378	0.271	0.323	29.2	2012	20.897	-100.674	11/11/2014	0.180 ± 0.12
JAL2	Phaozem	8.6	9.0	2.075	8.646	14.929	13.285	0.01	0.336	38.2	1791	21.253	-102.298	11/6/2014	0.219 ± 0.09
GTO2	Vertisol	8.6	8.7	7.089	13.948	18.038	17.722	0.4	0.417	33.8	1710	20.721	-101.329	11/11/2014	0.156 ± 0.07
DGO1	Planosol	8.9	8.9	5.018	10.838	23.36	23.773	0.251	0.248	33.8	1906	24.006	-104.385	9/11/2014	0.156 ± 0.05
SIN2	Regosol	9.1	8.8	18.578	4.91	49.131	50.98	1.103	0.157	29	10	23.293	-106.479	9/11/2014	0.314 ± 0.14

Tabla 1. Clasificación edafológica de los suelos, propiedades abióticas medidas en suelos iniciales y finales, metadatos y producción promedio de biomasa de *S. lycopersicum*. Los suelos se clasificaron de acuerdo con las cartas edafológicas nacionales utilizando las coordenadas medidas en cada sitio. Los metadatos y propiedades medidas en los suelos incluyen: pH, concentración de nitrógeno, fósforo y carbono totales, índice de aridez de Lang, altitud, coordenadas del sitio y fecha de colecta. SI: Suelos iniciales; SF: Suelos al final del experimento de jardín común (60 días).

Para describir la diversidad fisicoquímica de los suelos y encontrar las diferencias globales entre los suelos iniciales y suelos finales producto del experimento de jardín común, se realizó un análisis de ordenamiento con el método NMDS utilizando los valores de las concentraciones totales de nutrientes, el pH y el índice de aridez de Lang. El análisis de ordenamiento muestra que existen diferencias en las características globales de los suelos finales, comparados contra los suelos iniciales, encontrándose como dos grupos separados (Figura 3A). Mientras más cercanos se encuentren dos puntos (muestras), más similar es su composición fisicoquímica.

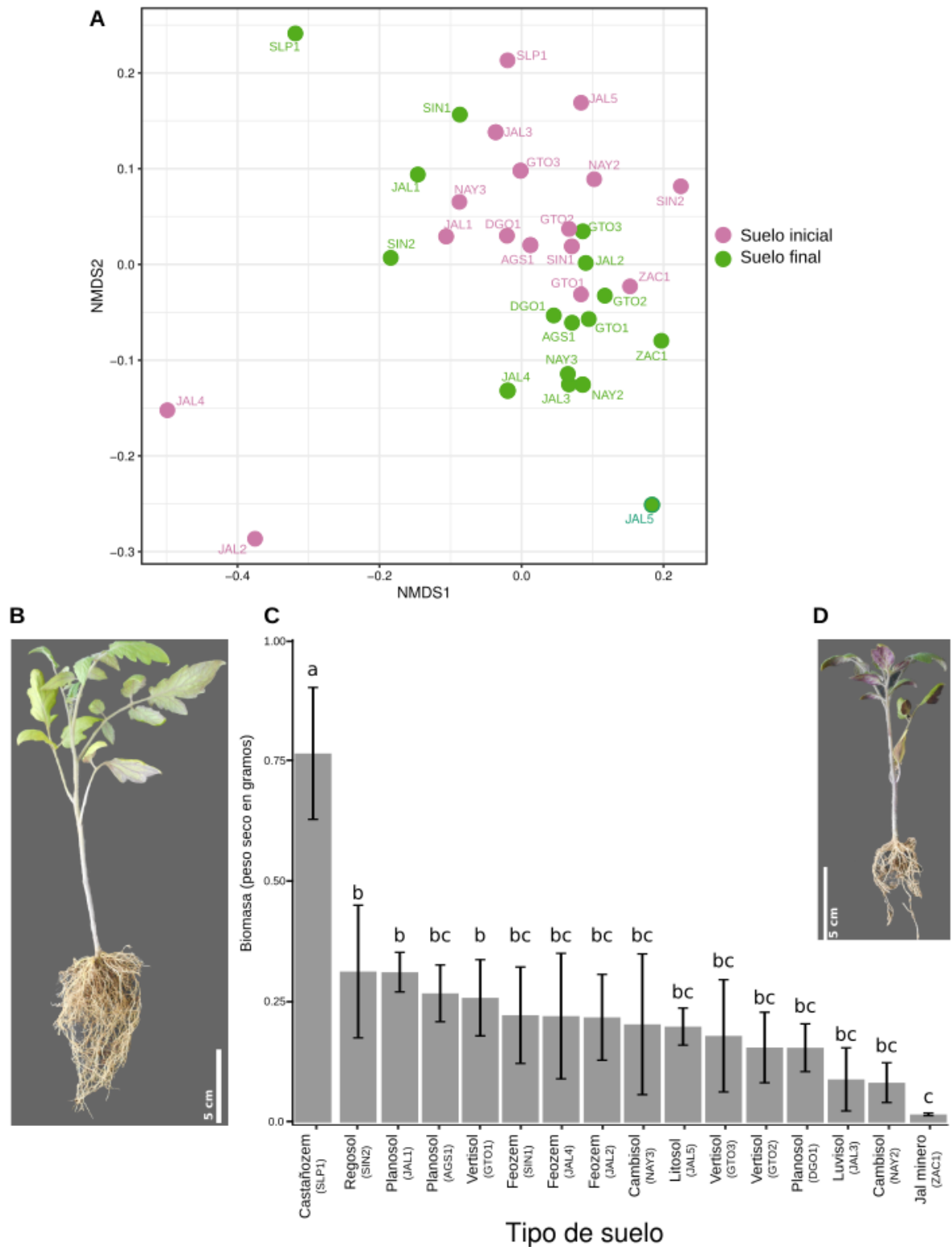


Figura 3. Diversidad fisicoquímica de los suelos al inicio y final del experimento de jardín común y producción de biomasa de *S. lycopersicum*. **A)** Análisis de ordenamiento NMDS (Stress: 0.1158) utilizando las variables: Índice de aridez, pH, concentraciones de N, C y P totales de los suelos al inicio del experimento (SI) y después del crecimiento de las plantas (SF). **B)** Ejemplo de un individuo de *S. lycopersicum* en el suelo SLP1. **C)** Gráfico de barras que muestra la biomasa producida por *S. lycopersicum* en cada suelo expresada como peso seco después de 60 días de crecimiento. Las diferencias significativas fueron determinadas a través de una ANOVA seguida de una prueba de Tukey ($p < 0.05$). **D)** Ejemplo de un individuo de *S. lycopersicum* en el suelo NAY2.

En el experimento de jardín común encontramos diferencias importantes en el desempeño de las plantas. El peso seco de las plantas de *S. lycopersicum* (n=4) fue utilizado como un índice de producción de biomasa en cada uno de los suelos colectados (n=16) (Figura 3B, 3C y 3D). La mayor producción de biomasa fue en el suelo SLP1 (0.77 ± 0.14 g), seguido de SIN2 (0.31 ± 0.14 g), JAL1 (0.31 ± 0.04 g), AGS1 (0.27 ± 0.06 g), GTO1 (0.26 ± 0.08 g), SIN1 (0.22 ± 0.14 g), JAL4 (0.22 ± 0.13 g), JAL2 (0.22 ± 0.09 g), NAY3 (0.20 ± 0.015 g), JAL5 (0.20 ± 0.04 g), GTO3 (0.18 ± 0.12 g), GTO2 (0.16 ± 0.07 g), DGO1 (0.16 ± 0.05 g), JAL3 (0.09 ± 0.07 g), NAY2 (0.08 ± 0.04 g) y el menor en ZAC1 (0.017 ± 0.02 g) (Tabla 1), el cual no es un suelo en sentido estricto, sino un jal minero, producto de desechos de minería, el cual había sido colonizado por plantas del género *Lupinus*.

Para identificar la diversidad de las plantas ruderales colectadas *in situ* se realizaron alineamientos de las secuencias de rRNA 16S extraídas de las secuencias de bibliotecas de amplicones y se asignaron hasta nivel de familia contra la base de datos NR de NCBI de acuerdo con el acierto con mayor porcentaje de identidad, encontrándose que la mayor parte de los huéspedes corresponden a la familia *Poaceae* (n=10), *Asteraceae* (n=3), *Lamiaceae* (n=1), *Fabaceae* (n=1) y *Fagales* (n=1) (Figura 4, Barajas *et al.*, 2020).

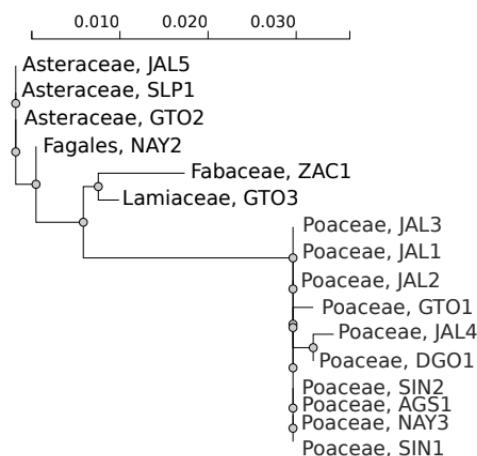


Figura 4. Afiliación taxonómica de las plantas ruderales colectadas *in situ*. Árbol de máxima verosimilitud utilizando las secuencias del gen ribosomal 16S y su clasificación taxonómica de la planta de cada sitio de acuerdo con el acierto con mayor porcentaje de identidad contra la base de datos NR del NCBI.

Diversidad del microbioma de suelos y raíces de *S. lycopersicum* y plantas ruderales: El modelo de dos pasos de adquisición del microbioma no se cumple en plantas ruderales

Se secuenciaron un total de 106 bibliotecas de amplicones de las muestras de suelos, rizósferas y endósferas de raíces de plantas ruderales y de *S. lycopersicum*, las cuales contenían un total de 8,035,907 secuencias pareadas crudas, que posteriormente se redujeron a 5,211,969 secuencias ensambladas (Anexos en línea: Tabla S1). Las secuencias ensambladas fueron agrupadas en un total de 2,570,541 OTUs al 97% de identidad, los cuales fueron reducidos a 271,940 OTUs después de haber descartado las secuencias clasificadas como mitocondrias ($n = 1926$) y cloroplastos ($n = 4561$), así como secuencias sin aciertos contra la base de datos ($n = 462$) y las secuencias que representaban *singletons* (secuencias únicas, $n = 2,291,652$).

Para describir la diversidad α de los microbiomas de suelos, rizósferas y endósferas, se calculó el número de OTUs observados en cada una de las muestras (Figura 5A; Anexos en línea: Tabla S2), encontrándose que la riqueza promedio de OTUs en las muestras de suelos (SI = $2,143 \pm 522$; SF = $3,149 \pm 1,161$; US = $2,882 \pm 860$) es menor a la riqueza encontrada tanto en las rizósferas como las endósferas de las plantas ruderales (RZ = $18,158 \pm 4,533$, EC = $19,885 \pm 6,742$), pero mayor al comparar contra las rizósferas y endósferas de *S. lycopersicum* (RT = $2,474 \pm 495$; ECT = $2,088 \pm 572$).

Adicionalmente, se utilizó el índice de Shannon (H') para estimar la diversidad de las comunidades bacterianas de suelos, rizósferas y endósferas. Los valores de la mediana muestran que las rizósferas y endósferas de las plantas ruderales (RZ = 8.58 ± 0.42 ; EC = 8.18 ± 0.70) son significativamente ($p < 0.005$) más diversas que los suelos de origen (SI = 7.12 ± 0.34) y las rizósferas y endósferas de las plantas de *S. lycopersicum* (RT = 6.84 ± 0.54 ; ECT = 6.51 ± 0.39) (Figuras 5B y 6B). A su vez, éstas últimas tienen valores por debajo de los de los suelos de origen así como de los suelos finales y control (SF = 7.35 ± 0.33 ; US = 7.24 ± 0.32) sin encontrarse diferencias significativas entre los grupos. Cabe mencionar que en los suelos colectados al final del experimento de jardín común hubo un incremento de riqueza y diversidad comparados con los suelos iniciales, pero no se encontraron diferencias significativas entre estos grupos (Figura 5; Anexos en línea: Tabla S2, Barajas *et al.*, 2020).

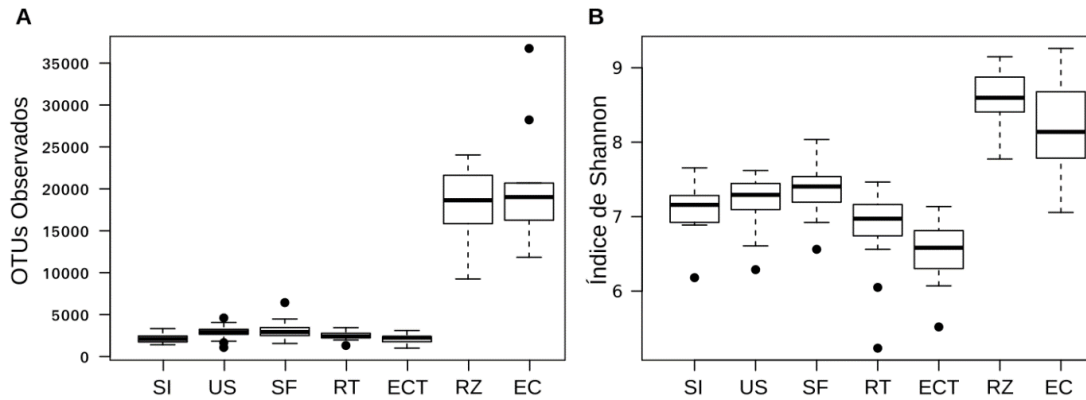


Figura 5. Riqueza y diversidad de los microbiomas de los suelos al inicio y final del experimento de jardín común (60 días), rizósferas y endósferas de *S. lycopersicum* y plantas ruderales. Gráficos de caja y bigotes de: **A)** Número de OTUs observados. **B)** Índice de diversidad de Shannon. SI: suelos al inicio del experimento, US: suelo control sin plantas al final del experimento, SF: suelo con plantas al final del experimento, RT: rizósfera de *S. lycopersicum*, ECT: endósfera de *S. lycopersicum*, RZ: rizósfera de plantas ruderales, EC: endósfera de plantas ruderales.

Para identificar los grupos taxonómicos compartidos entre los microbiomas de suelos y raíces de plantas ruderales y *S. lycopersicum*, comparamos los perfiles taxonómicos a nivel de género a través de diagramas de Venn (Figura 6A; Anexos en línea: Tabla S3, Barajas *et al.*, 2020). Encontramos que los microbiomas de los tres grupos (suelos y raíces de plantas ruderales y *S. lycopersicum*) tienen 586 géneros en común, mientras que los suelos comparten 68 géneros con las plantas ruderales y sólo 4 géneros con *S. lycopersicum*. Los microbiomas de ambos grupos de plantas comparten 143 géneros que no se encuentran incluidos entre los géneros comunes con el suelo y en particular, las raíces de *S. lycopersicum* cuentan con 14 géneros únicos, mientras que las plantas ruderales cuentan con 360 géneros únicos.

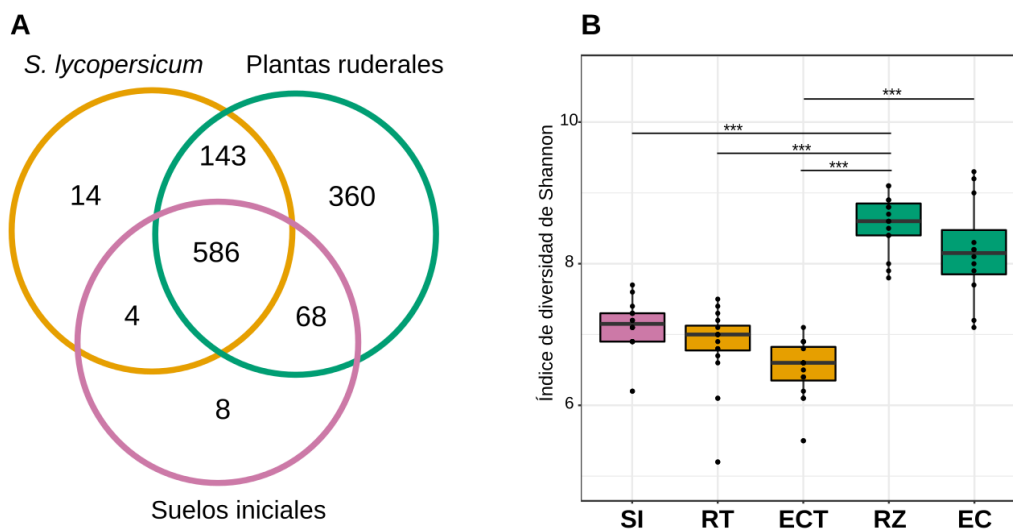


Figura 6. Diversidad α de suelos iniciales, rizósferas y endósferas de plantas ruderales y *S. lycopersicum* **A)** Diagrama de Venn representando el número de géneros bacterianos en las raíces (rizósfera + endósfera) de plantas ruderales, *S. lycopersicum* y suelos iniciales. **B)** Gráfico de caja y bigotes mostrando los valores del índice de diversidad de Shannon en los suelos iniciales (SI), rizósferas de *S. lycopersicum* (RT), endósferas de *S. lycopersicum* (ECT), rizósferas de plantas ruderales (RZ) y endósferas de plantas ruderales (EC).

Para identificar a los miembros del microbioma núcleo de raíz de *S. lycopersicum* y plantas ruderales, se comparó la composición de géneros de las muestras de rizósfera y endósfera en cada grupo de plantas a través de diagramas Upset (Figura 7). El microbioma núcleo de *S. lycopersicum* y plantas ruderales se conforman por 51 y 87 géneros bacterianos respectivamente.

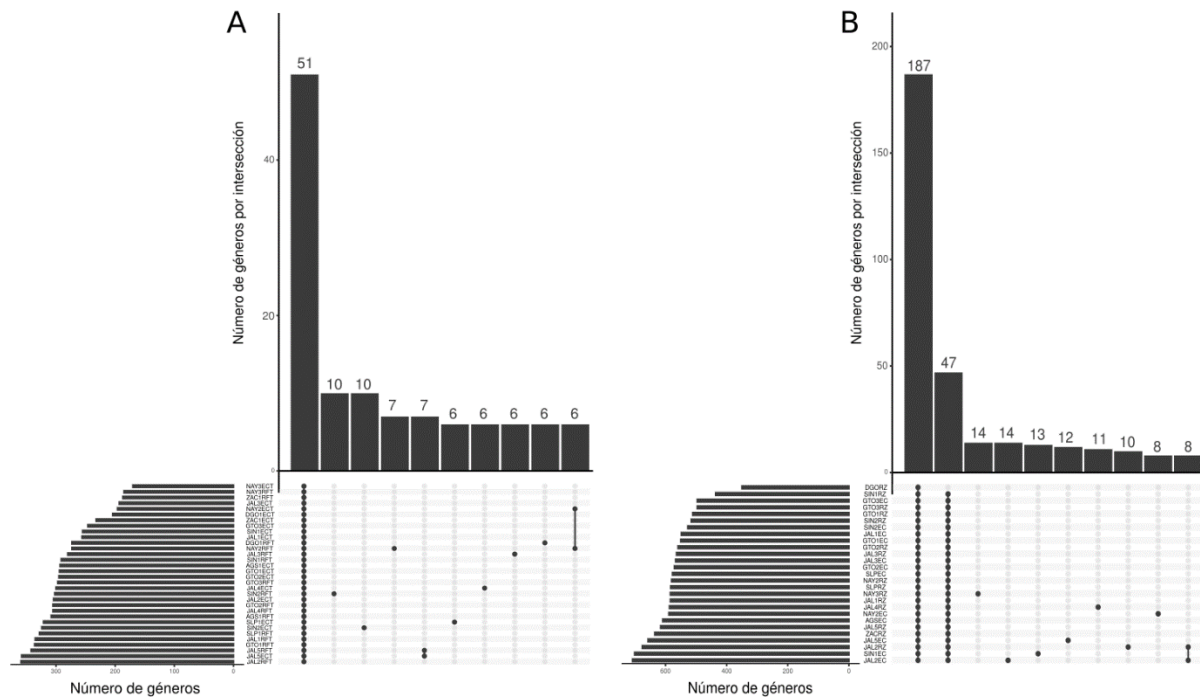


Figura 7. Microbioma núcleo de plantas ruderales y *S. lycopersicum*. A) Diagrama mostrando el número de géneros compartidos entre los microbiomas de *S. lycopersicum*. B) Diagrama mostrando el número de géneros compartidos entre los microbiomas de plantas ruderales.

Influencia del genotipo del hospedero en la estructura de comunidades bacterianas de raíz: La convergencia del microbioma de raíz de *S. lycopersicum*

Con el fin de evaluar el efecto del genotipo de los huéspedes, así como el papel de las propiedades del suelo en la composición del microbioma de raíz de las plantas, se llevaron a cabo distintos análisis de diversidad β entre las muestras de suelos, rizósferas y endósferas. Se utilizó la matriz de distancias Unifrac para realizar un agrupamiento jerárquico de las muestras, el cual nos permite describir las diferencias entre cada micronicho de acuerdo con las relaciones filogenéticas y la abundancia de los OTUs presentes en cada comunidad bacteriana (Figura 8, izquierda, Barajas *et al.*, 2020). En el dendrograma de la Figura 8 se pueden observar tres clados principales: El clado I, compuesto sólo de microbiomas de

rizósferas (10/16 muestras) y endósferas (13/16 muestras) de *S. lycopersicum*; el clado II, en el cual se encuentran dispuestos la mayoría de los microbiomas de suelos (SI, SF y US), junto con las rizósferas y endósferas de las plantas ruderales; y el clado III, un grupo en donde hallamos mezcladas a muestras de rizósferas y endósferas de ambos grupos de plantas, así como suelos. Es interesante notar que los suelos de este grupo corresponden a los suelos de origen (SI) y suelos finales del experimento de jardín común (SF) que son contraparte de las muestras asociadas a *S. lycopersicum* y que son de carácter ácido. El agrupamiento de los clados fue evaluado con un análisis de similitudes (ANOSIM) el cual muestra que la agrupación de cada clado cuenta con un soporte estadístico válido ($R = 0.7257$; $p < 0.001$; 999 permutaciones).

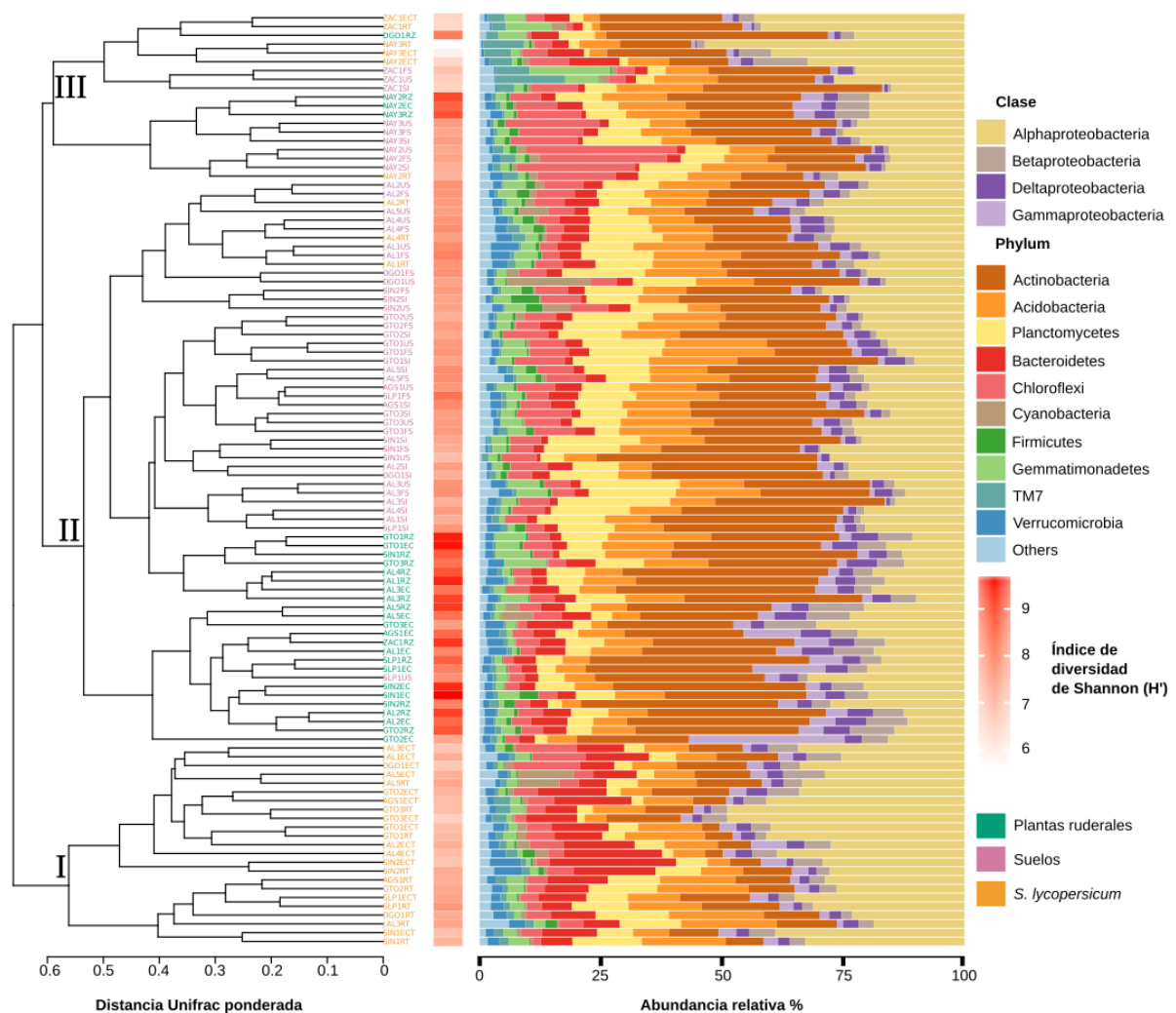


Figura 8. Influencia del genotipo del huésped y el suelo en la estructura y diversidad de las comunidades bacterianas. Dendrograma de agrupación jerárquica basado en las distancias UniFrac ponderadas mostrando la diversidad β entre los microbiomas de suelos, rizósferas y endósferas de plantas ruderales y *S. lycopersicum* (izquierda). El identificador de la locación se despliega en las puntas del dendrograma con un código alfanumérico de cuatro caracteres y un sufijo indicando el tipo de muestra. SI: suelos al inicio del experimento, US: suelo control sin plantas al final del experimento, SF: suelo con plantas al final del experimento, RT: rizósfera de *S. lycopersicum*, ECT: endósfera de *S. lycopersicum*, RZ: rizósfera de plantas ruderales, EC: endósfera de plantas ruderales. Gráfico de calor mostrando la diversidad de Shannon (H') calculada para cada muestra (centro). Gráfico de barras que presenta la abundancia relativa a nivel de phylum en cada muestra (derecha).

Para conocer la cohesividad dentro de cada clado, evaluamos los valores promedio de las distancias encontradas entre los microbiomas dentro de cada clado (Anexos en línea: Tabla S4, Barajas *et al.*, 2020). El valor para el clado I es igual a 0.5041, mientras que la del clado II es 0.5058 y en el clado III es de 0.4787. Al observar las distancias que se encuentran entre cualquier par de hojas en el clado I hacia cualquier hoja en el clado II o III es igual 0.6608, un valor mayor al promedio encontrado dentro de cualquiera de los tres clados.

Es interesante notar que los patrones generales de abundancia (Figura 8, derecha; Anexos en línea: Tabla S5, Barajas *et al.*, 2020) muestran que las proporciones de Actinobacteria en los microbiomas presentes en los clados II y III son mayores a las que se observan en el clado I, mientras que en éste se puede apreciar una mayor abundancia de Bacteroidetes y Proteobacteria, específicamente α -Proteobacteria.

Adicionalmente, se evaluó la diversidad beta a través de un análisis de ordenamiento no métrico multidimensional (NMDS) (Figura 9-A) en donde se observa la misma tendencia de agrupamiento de los grupos previamente descritos. A través de este análisis también pudimos evaluar el efecto de propiedades abióticas del suelo en la composición del microbioma sin poder encontrar diferencias entre los grupos dadas estas variables (Figura 9-B).

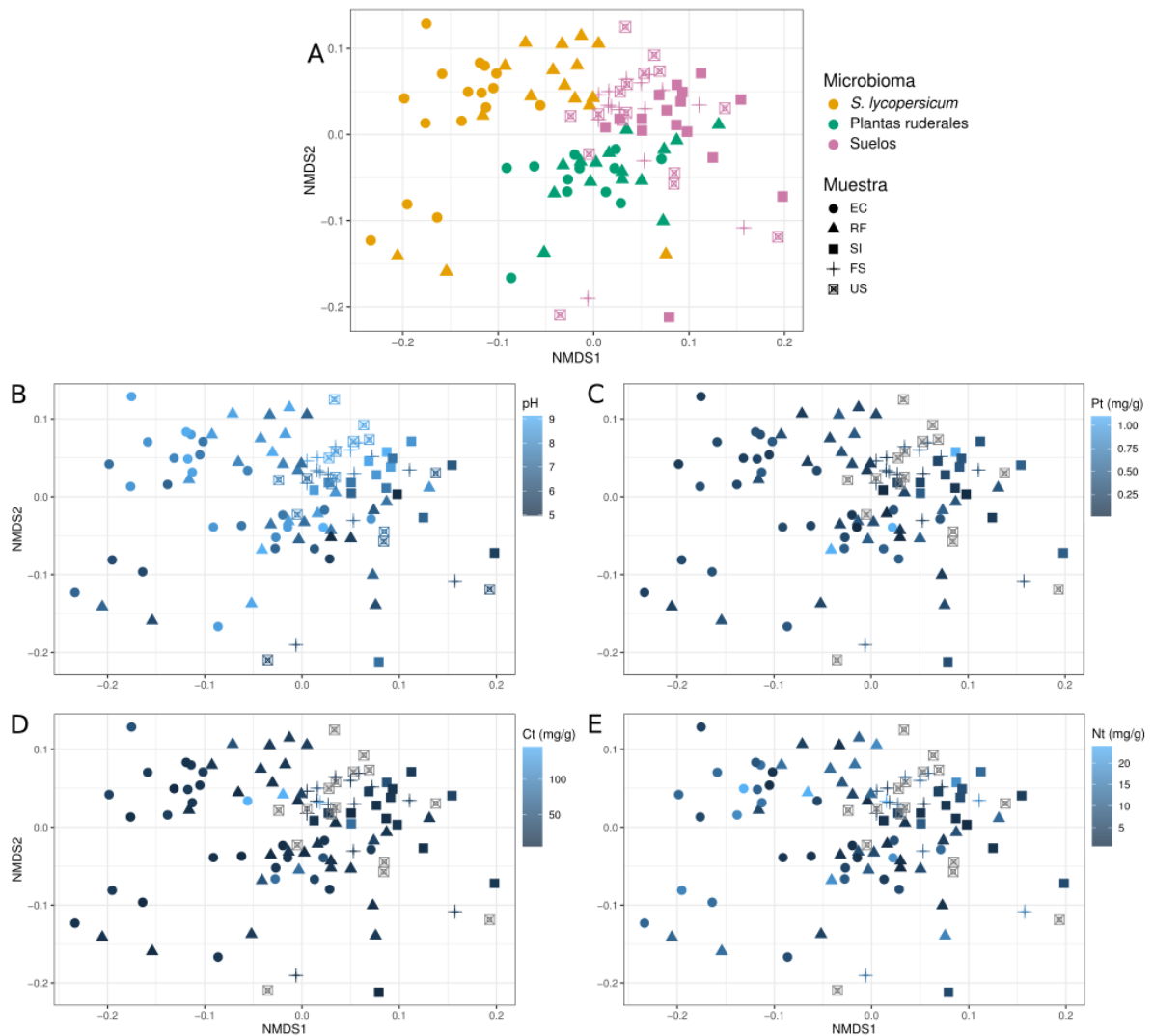


Figura 9. Análisis de diversidad beta de los microbiomas de suelos, rizósferas y endósferas de plantas ruderales y *S. lycopersicum*. Ordenamiento multidimensional no métrico (NMDS) basado en las abundancias de OTUs y distancias UniFrac ponderadas. Stress = 0.1888. Los colores corresponden en **A)** el tipo de muestra, **B)** los valores de pH del suelo, **C)** La cantidad de fósforo total en los suelos, **D)** La cantidad de carbono total en los suelos y **E)** La cantidad de nitrógeno total en los suelos.

Enriquecimiento de grupos taxonómicos en raíces de plantas domesticadas y ruderales: La selección activa del microbioma por parte de los hospederos

Debido a que encontramos diferencias evidentes en la composición taxonómica de los microbiomas dependiendo del nicho habitado, realizamos análisis de varianza entre los grupos de suelos, raíces de plantas ruderales y raíces de *S. lycopersicum* (Figura 10, Barajas *et al.*, 2020).

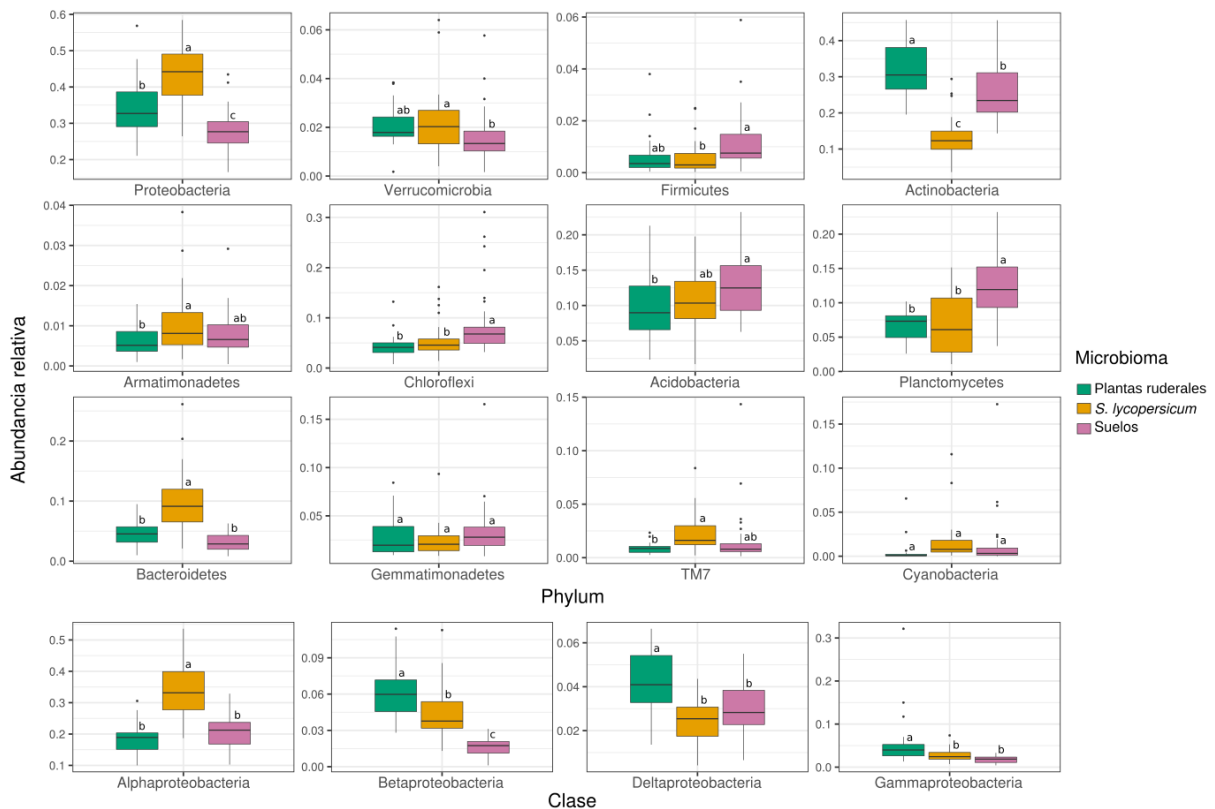


Figura 10. Abundancia relativa de phyla bacterianos en suelos y rizósferas de plantas ruderales y *S. lycopersicum*. En cada panel se muestra un diagrama de caja y bigotes representando la abundancia relativa de cada phylum. El phylum Proteobacteria se desglosa a nivel de clase. Las diferencias significativas fueron determinadas a través de una ANOVA seguida de una prueba de Tukey ($p < 0.001$).

Los microbiomas de suelos, así como el de rizósferas y endósferas de plantas ruderales se encuentra dominado por Actinobacteria, esta abundancia es significativamente mayor (ANOVA $p < 2E^{-17}$) a la encontrada en los microbiomas de *S. lycopersicum* (Figura 10, Barajas *et al.*, 2020). Por lo contrario, en el microbioma de raíz de *S. lycopersicum* encontramos que Proteobacteria es el phylum dominante y su abundancia es significativamente mayor (ANOVA $p < 1.8E^{-15}$) en comparación con los microbiomas de suelos y raíces de plantas ruderales. Debido a que Proteobacteria es un phylum diverso y polifilético, éste fue descrito a nivel de clase (α , β , γ , and δ -Proteobacteria). La clase α -Proteobacteria resultó ser la más abundante en las raíces de *S. lycopersicum*, presentando diferencias significativas (ANOVA $p < 2E^{-16}$) con los grupos de suelos y raíces de plantas ruderales, mientras que las demás clases de Proteobacteria (β , γ , and δ) se encuentran en mayor abundancia en los microbiomas de plantas ruderales y suelos ($p < 0.05$). El último phylum que presentó una mayor abundancia con diferencias significativas en las raíces de *S. lycopersicum* fue Bacteroidetes (ANOVA $p < 1.3E^{-15}$) (Figura 10). Ninguno de los demás grupos taxonómicos presenta diferencias significativas en abundancia entre los microbiomas de suelos y raíces de plantas.

A partir de los conjuntos de OTUs compartidos por los distintos tipos de microbiomas se realizó un análisis de abundancia diferencial basada en las veces de cambio de la abundancia de OTUs para identificar OTUs que se encontraran enriquecidos preferencialmente ya sea en las raíces de *S. lycopersicum* (RT, ECT) o en las raíces de plantas ruderales (RZ, EC) al comparar contra los suelos (SI, SF ó US) (Figura 11; Anexos en línea: Tabla S6, Barajas *et al.*, 2020). En la comparación de los microbiomas de *S. lycopersicum* contra los suelos iniciales, encontramos a 6 OTUs diferencialmente abundantes en rizósfera, los cuales corresponden a los géneros *Sphingobium*, *Caulobacter*, *Asticcacaulis*, *Arthrospira* y *Kaistobacter*, mientras que en la comparación contra la endósfera se identificaron a 14 OTUs enriquecidos, pertenecientes a los mismos géneros antes descritos, además de *Agrobacterium* y *Lacibacter* (Figuras 11A y 11B). En las comparaciones de los microbiomas de plantas ruderales contra suelos de origen encontramos un mayor número de OTUs enriquecidos; en el caso de las rizósferas, encontramos 45 OTUs diferencialmente abundantes, y en las endósferas hayamos 31 OTUs (Figuras 11C y 11D). La mayor parte de los OTUs enriquecidos en los microbiomas de las plantas ruderales fueron asignados dentro del phylum Actinobacteria, especialmente en las rizósferas, donde observamos a géneros como *Lentzea*, *Amycolatopsis*, *Couchioplanes*, *Actinoplanes* y *Geodermatophilus*. Por otro lado, en las endósferas encontramos un mayor número de miembros pertenecientes al phylum Proteobacteria, en concreto a los géneros *Methylibium*, *Roseateles*, *Rhizobium*, *Xenophilus*, *Arthrospira*, *Kaistobacter* y *Agrobacterium*, el cual también se encontraba en las rizósferas. Las comparaciones de los microbiomas de plantas contra SF y US no arrojaron diferencias significativas entre los grupos.

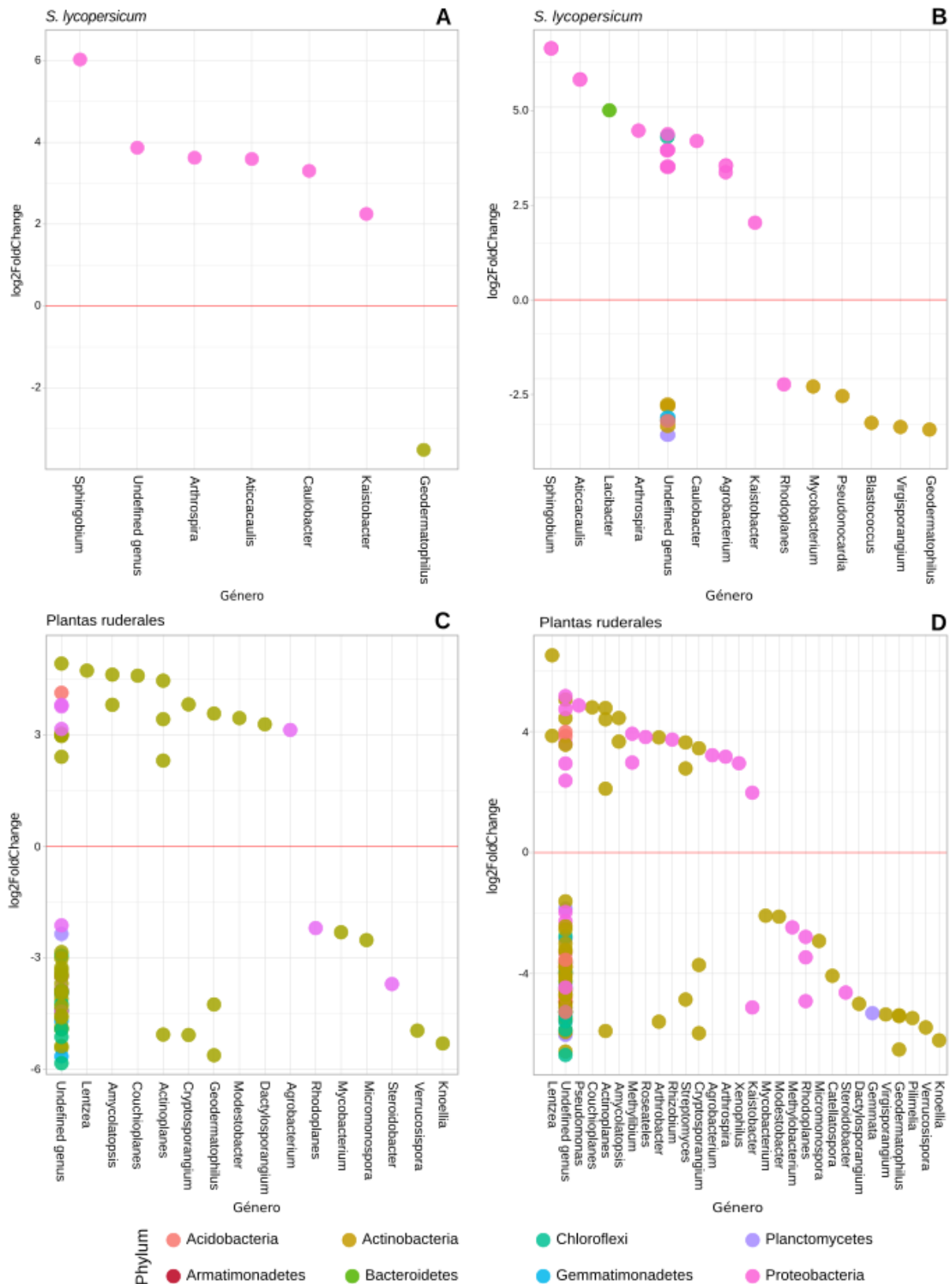


Figura 11. OTUs enriquecidos en la rizósfera y endósfera de plantas ruderales y de *S. lycopersicum*. Gráficos del Log₂ de las veces de cambio de las abundancias de OTUs en la comparación de los suelos iniciales contra **A)** Rizósferas de *S. lycopersicum* **B)** Endósferas de *S. lycopersicum* **C)** Rizósferas de plantas ruderales. **D)** Endósferas de plantas ruderales. Los OTUs se presentan a nivel de género y se colorean con base en el phylum al que corresponden. Los valores por encima de cero representan a OTUs enriquecidos en microbiomas de plantas.

Descripción del potencial metabólico de comunidades las microbianas de suelos y rizósferas

Se secuenciaron en total 17 metagenomas de muestras de suelos de origen ($n = 6$), rizósferas de plantas ruderales ($n = 6$) y rizósferas de plantas de *S. lycopersicum* ($n = 5$), obteniendo 50.1 Gb secuenciadas. Después del control de calidad, se obtuvieron 464,372,598 lecturas pareadas de secuenciación con un promedio de $27,316,035 \pm 2,943,233$ secuencias por metagenoma. Las lecturas de calidad fueron ensambladas en un total de 12,677,118 de contigs con un promedio de $745,713 \pm 366,001$ contigs por muestra, los cuales tienen un valor de N50 promedio de 176 ± 51 bp y la mayor longitud de contig igual a 45,645 bp (Anexos en línea: Tabla S7, Barajas *et al.*, 2020). A partir de los contigs metagenómicos, se logró hacer la predicción computacional de 12,272,971 proteínas con un promedio de $708,835 \pm 332,770$ proteínas predichas por muestra. Posteriormente, se realizó la anotación de proteínas contra la base de datos M5NR, contra la cual se tuvieron 3,147,929 aciertos. A partir de los aciertos obtenidos contra esta base de datos se obtuvieron dos juegos de sistemas de anotación distintos; el primero en base a Refseq y el segundo con la base de datos de Subsistemas de SEED. El número de proteínas anotadas por Refseq fue de 411,432, que corresponden sólo al 13.07% de los aciertos en M5NR, mientras que la anotación contra los subsistemas de SEED obtuvimos 138,627 proteínas, representando sólo el 4.4% de los aciertos contra M5NR.

Para describir el potencial metabólico de las comunidades microbianas de suelos y rizósferas utilizamos la base de datos de subsistemas de SEED, la cual se basa en una organización ontológica de la información, dividiéndose en cuatro diferentes niveles. Para realizar la descripción individual de los metagenomas, se utilizó el primer nivel de la clasificación jerárquica, compuesto de 28 distintas categorías (Figura 12, Barajas *et al.*, 2020). Este análisis nos mostró que, de acuerdo con las proteínas anotadas mediante este sistema, no podemos observar diferencias aparentes entre los perfiles metabólicos predichos en los metagenomas y únicamente fuimos capaces de encontrar pequeñas diferencias no significativas (TukeyHSD; $p > 0.05$) entre los metagenomas de suelos y rizósferas de *S. lycopersicum* en las categorías relacionadas a la adquisición de hierro ($p = 0.07$) y la producción de pared celular y cápsula ($p = 0.06$). La única diferencia significativa que se obtuvo fue entre las rizósferas de *S. lycopersicum* y plantas ruderales, encontrada en la categoría de genes relacionados al metabolismo de azufre ($p = 0.017$).

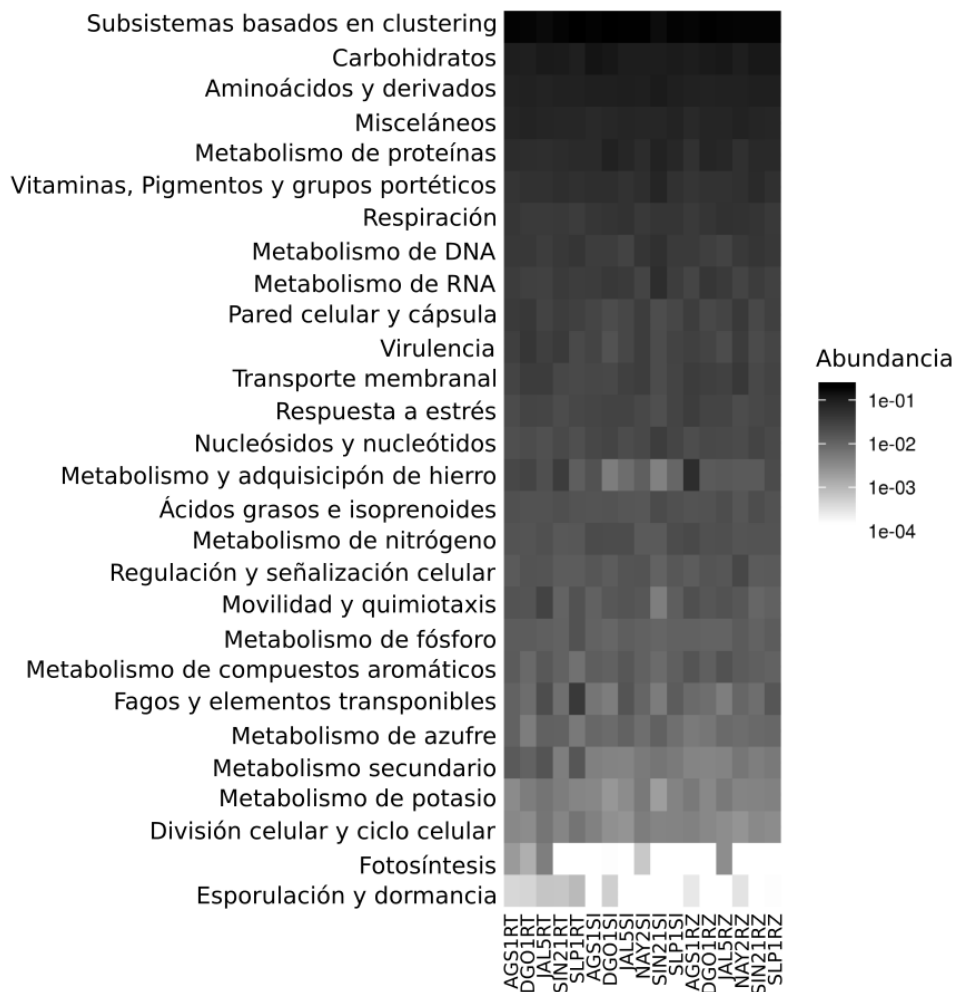


Figura 12. Perfiles metabólicos predichos de los metagenomas de suelos y rizósferas. Gráfico de calor mostrando la abundancia normalizada de proteínas anotadas en el nivel 1 de Subsistemas de SEED. Las muestras se despliegan en columnas con un código alfanumérico de cuatro caracteres y un sufijo indicando el tipo de muestra. SI: Suelos al inicio del experimento, RT: Rizósfera de *S. lycopersicum*, RZ: Rizósfera de plantas ruderales.

Para identificar a los grupos taxonómicos predominantes en los metagenomas, se clasificaron taxonómicamente las lecturas crudas de secuenciación a través de *binning* de las lecturas metagenómicas contra bases de datos de bacterias, arqueas y eucariontes, lo que nos permite llegar, en algunos casos, hasta el nivel de especie y cepa (para bacteria). Se lograron clasificar el 38% de las lecturas metagenómicas totales, de las cuales, el 88.14% corresponden al dominio Bacteria y el 7.09% tienen aciertos contra el dominio Eukarya, de estos, casi la mitad (3.38%) son aciertos contra Fungi (Anexos en línea: Tabla S8, Barajas *et al.*, 2020).

Basados en el *binning* metagenómico, la especie bacteriana con mayor abundancia en los suelos corresponde a *Solirubrobacter soli*, miembro del phylum Actinobacteria, y es importante notar que también se encuentra altamente abundante en los metagenomas de las rizósferas de ambos grupos de plantas. En los metagenomas de plantas ruderales, los

microorganismos más abundantes son *Sphingomonas* sp. URHD0057, miembro de la clase α -Proteobacteria, *S. soli* y *Rhodoplanes* sp. Z2-YC6860, miembro del orden *Rhizobiales*. Dentro de los metagenomas de rizósfera de *S. lycopersicum* pudimos detectar a múltiples especies del phylum Bacteroidetes: *Ohtaekwangia koreensis*, *Flavobacterium terrae*, *Niastella vici*, *Chryseolinea serpens* y al microorganismo identificado *Chitinophagaceae* bacterium IBVUCB2, todos miembros del phylum Bacteroidetes. En cuanto a los integrantes del reino Fungi, se pudieron detectar como las especies más abundantes en las rizósferas de tomates y plantas ruderales a *Rhizophagus irregularis* (Mucoromycota), *Melampsora larici-populina* (Basidiomycota) y *Saitoiella complicata* (Ascomycota).

Diversidad metagenómica de suelos y rizósferas de *S. lycopersicum* y plantas ruderales: El modelo de dos pasos se cumple a nivel metabólico

El análisis de la diversidad α del potencial metabólico de las comunidades microbianas de suelos y rizósferas se llevó a cabo utilizando el número observado total de proteínas predichas y basados en el índice de diversidad de Shannon. Con base en el número observado de proteínas totales predichas encontramos que, en promedio, los suelos cuentan con un mayor número de proteínas ($249,880 \pm 91,770$) que las rizósferas de plantas ruderales ($218,973 \pm 100,751$) y *S. lycopersicum* ($199,365 \pm 74,505$), aunque las medianas no difieran (SI = 221,364; RF = 219,757; RZ = 216,665) (Figura 13; Anexos en línea: Tabla S9, Barajas *et al.*, 2020). La diversidad de Shannon calculada para los suelos tiene un valor mayor (mediana=11.8), en comparación con la H' encontrada en *S. lycopersicum* (mediana = 11.3) y plantas ruderales (mediana = 11.3). Sin embargo, las diferencias entre los grupos no son significativas, pero la dispersión de los datos es mayor en las rizósferas de plantas ruderales que en los suelos (SI σ = 0.37; RZ σ = 0.66) (Figuras 13 y 14B; Anexos en línea: Tabla S9, Barajas *et al.*, 2020).

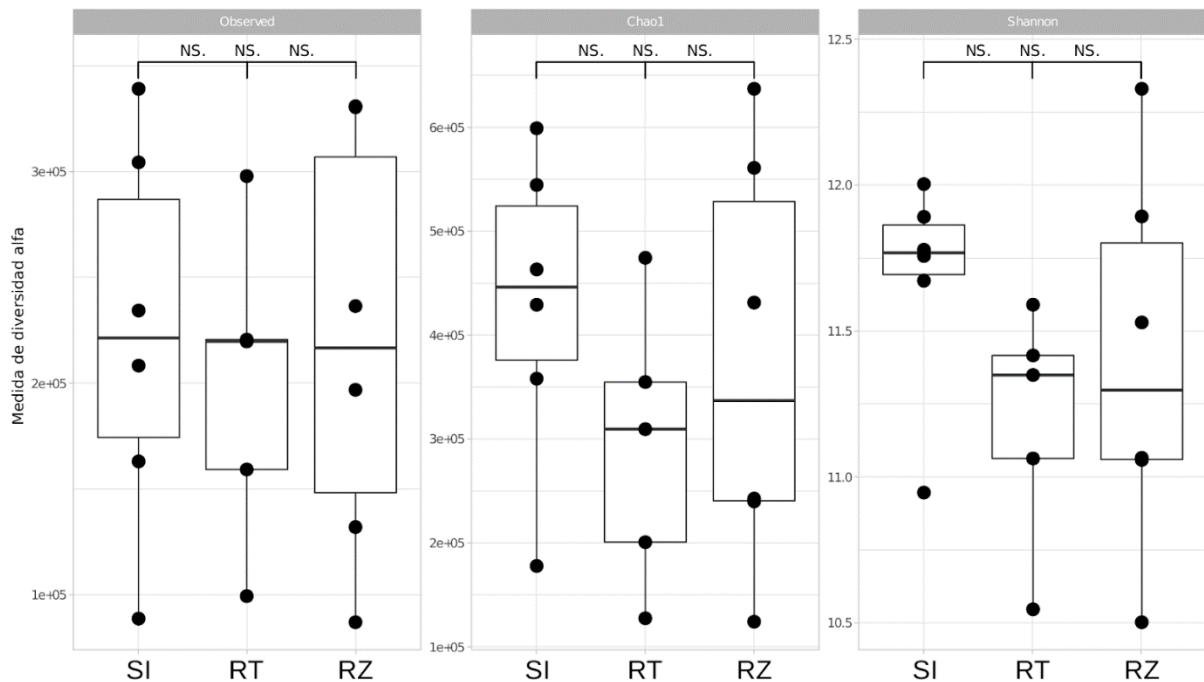


Figura 13. Riqueza y diversidad α de proteínas predichas en suelos y rizósferas de plantas ruderales y *S. lycopersicum*. Gráficos de caja y bigotes mostrando el número de proteínas observadas, el estimador de diversidad Chao1 y el índice de diversidad de Shannon para cada tipo de muestra. No se encontraron diferencias significativas entre los valores. SI: Suelos al inicio del experimento, RT: Rizósfera de *S. lycopersicum*, RZ: Rizósfera de plantas ruderales.

Para conocer la distribución de las proteínas predichas y anotadas contra Refseq entre los distintos metagenomas, se realizaron comparaciones a través de diagramas de Venn entre el grupo de suelos, rizósferas de plantas ruderales y rizósferas de plantas de *S. lycopersicum* (Figura 14A, Barajas *et al.*, 2020). En este análisis se consideró a una proteína presente dentro de un grupo con estar incluida en el mínimo de una muestra. Los metagenomas de suelos y rizósferas comparten un total de 43,305 proteínas predichas (40,553 con anotación). El suelo comparte 33,603 proteínas (19,782 con anotación) con las rizósferas de *S. lycopersicum* y 63,067 proteínas (40,091 con anotación) con las plantas ruderales; un 8.49% y 8.72% del total de proteínas, respectivamente. Las comunidades microbianas de rizósfera de *S. lycopersicum* comparten un mayor número de genes codificantes con las plantas ruderales (8.85%) que los que ambas comparten con el suelo (8.49%) (Figura 14A). Los conjuntos de proteínas únicas en los tres grupos resultaron ser mayoritarios en los suelos con el 88.8% de proteínas codificantes, seguido de las rizósferas de plantas ruderales y *S. lycopersicum* con 87.4% y 86.3%, respectivamente. Es interesante notar que los metagenomas de rizósfera de *S. lycopersicum* cuentan con la mayor proporción de proteínas anotadas (12.1%) en comparación con las plantas ruderales (9.97%) y los suelos (6.75%). La lista completa de proteínas anotadas se encuentra en Anexos en línea: Tabla S10 (Barajas *et al.*, 2020).

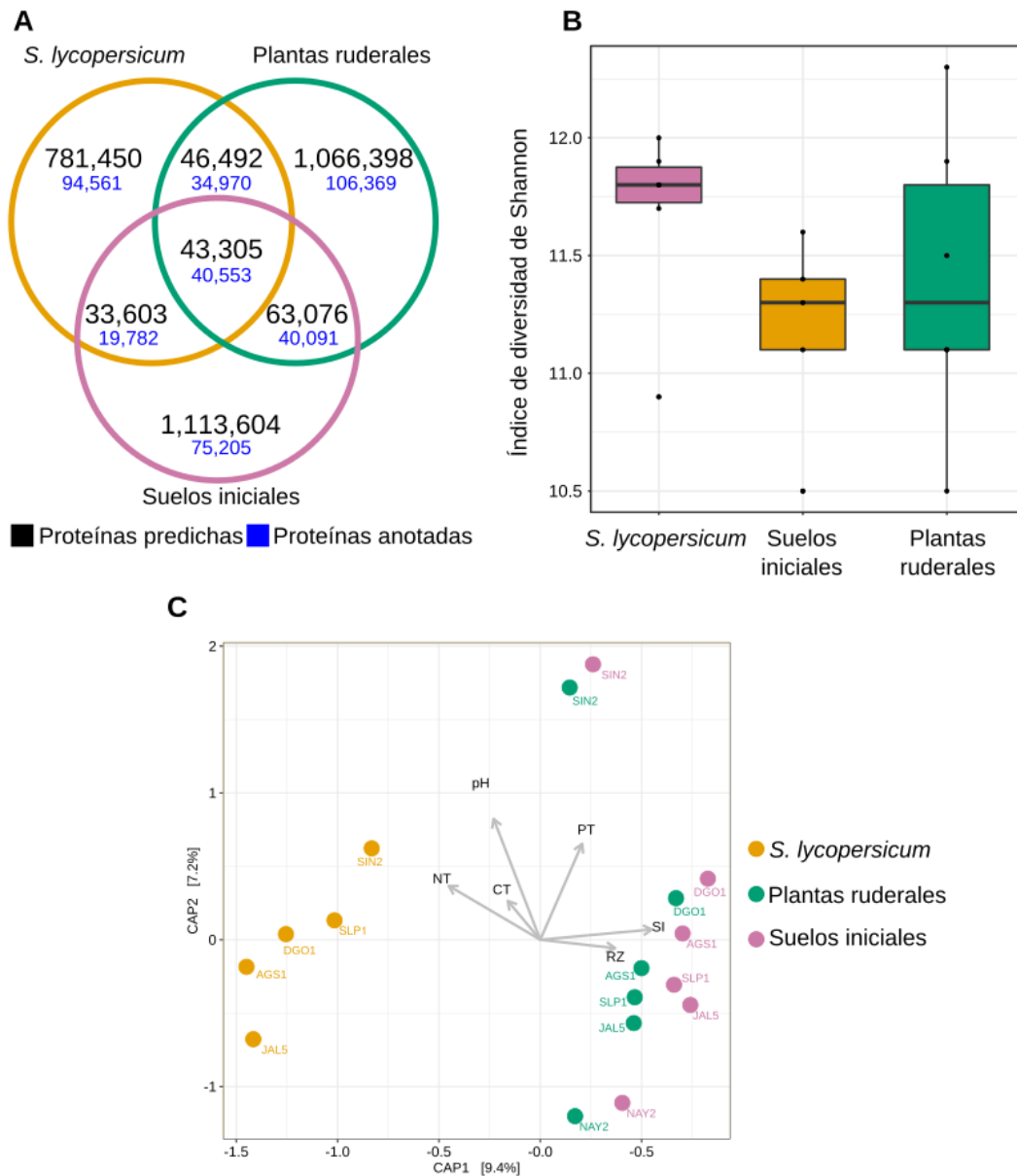


Figura 14. Diversidad metagenómica en suelos y rizósferas de plantas ruderales y *S. lycopersicum*. **A)** Diagrama de Venn mostrando las familias de proteínas predichas únicas y compartidas entre los metagenomas de suelos iniciales y rizósferas. **B)** Gráficos de caja y bigotes mostrando la diversidad de Shannon con base en el número total de proteínas predichas en suelos y rizósferas. **C)** Análisis Coordenadas Principales (CAP) calculado a partir del número total de proteínas predichas en cada muestra de suelos y rizósferas usando una matriz de disimilitudes de Bray-Curtis y constreñido por las variables ambientales desplegadas en los vectores. CT: Concentración de carbono total, NT: Concentración de nitrógeno total, PT: Concentración de fósforo total, RZ: Rizósfera de plantas ruderales, SI: Suelos iniciales.

Para comprender las diferencias en el potencial metabólico que son impuestas tanto por el microbioma generado por el suelo y por las distintas especies de plantas en la rizósfera, se comparó la composición de proteínas predichas en los metagenomas a través de un análisis de Coordenadas Principales (CAP) (Figura 14C). Este análisis fue constreñido de acuerdo con las variables fisicoquímicas medidas en los suelos (N, C, P y pH) y el tipo de muestra (SI, RF, RZ), con el objetivo de explorar la contribución de estas propiedades bióticas

y abióticas en las diferencias en la composición metabólica de las comunidades microbianas asociadas. El análisis de CAP logra explicar el 16.6% del total de la varianza en la composición metabólica de las comunidades microbianas de suelos y rizósferas. El CAP1 logra explicar el 9.4% de la varianza, formando dos grupos separados, uno compuesto por los metagenomas de rizósferas de *S. lycopersicum* y el segundo agrupamiento que contiene a los metagenomas de suelos y rizósferas de plantas ruderales. La separación de estos grupos se corroboró a través de un análisis de ANOSIM ($R = 0.4568$; $p < 0.001$; 999 permutaciones). Adicionalmente, la disimilitud entre las plantas ruderales y suelos es suficiente para no mezclar a ambos grupos, resultado soportado estadísticamente de acuerdo con un análisis de PERMANOVA (999 permutaciones sobre las distancias de Bray-Curtis, $p < 1E^{-4}$) que muestra las diferencias entre ambos micronichos. La disposición de los vectores que presentan a las variables bióticas y abióticas con las cuales se constriñe a los dos ejes del CAP en la Figura 14C muestran que existe una correlación positiva para las variable abiótica de concentración de fósforo y la variables bióticas tipo de muestra, específicamente suelos y rizósferas de plantas ruderales, en particular la correlación con el componente biótico remarca la separación de los microbiomas a nivel metabólico dependiendo del genotipo del huésped. Por otro lado, las variables abióticas de pH y concentraciones de carbono y nitrógeno presentan una correlación negativa en el gráfico.

Enriquecimiento de proteínas predichas en los metagenomas de rizósferas de *S. lycopersicum* y plantas ruderales.

Con el objetivo de investigar si existen genes codificantes que puedan estar enriquecidos dependiendo del micronicho habitado, se llevaron a cabo comparaciones pareadas entre los perfiles metabólicos predichos de los metagenomas de suelos y rizósferas con base en las proteínas anotadas por Refseq. Se utilizó el logaritmo base dos del valor de las veces de cambio para identificar a las proteínas diferencialmente abundantes ($p < 0.001$) entre cada uno de los grupos (Anexos en línea: Tabla S11, Barajas *et al.*, 2020). En la comparación entre rizósferas de *S. lycopersicum* y suelos encontramos 67 proteínas enriquecidas en *S. lycopersicum* las cuales están involucradas en procesos celulares como movilidad, quimiotaxis y formación de biofilm (*e.g.*, LuxR, CheY, diguanilato ciclase), genes relacionados a la producción de antibióticos (*e.g.*, β -lactamasa clase C), degradación de carbohidratos complejos (*e.g.*, xiloglucanasa, celulasa Cel5F), proteínas relacionadas al metabolismo de hierro (TonB), esporulación (SpolIIE) y proteínas relacionadas a sistemas de secreción (exosortasa) (Figura 15, Barajas *et al.*, 2020).

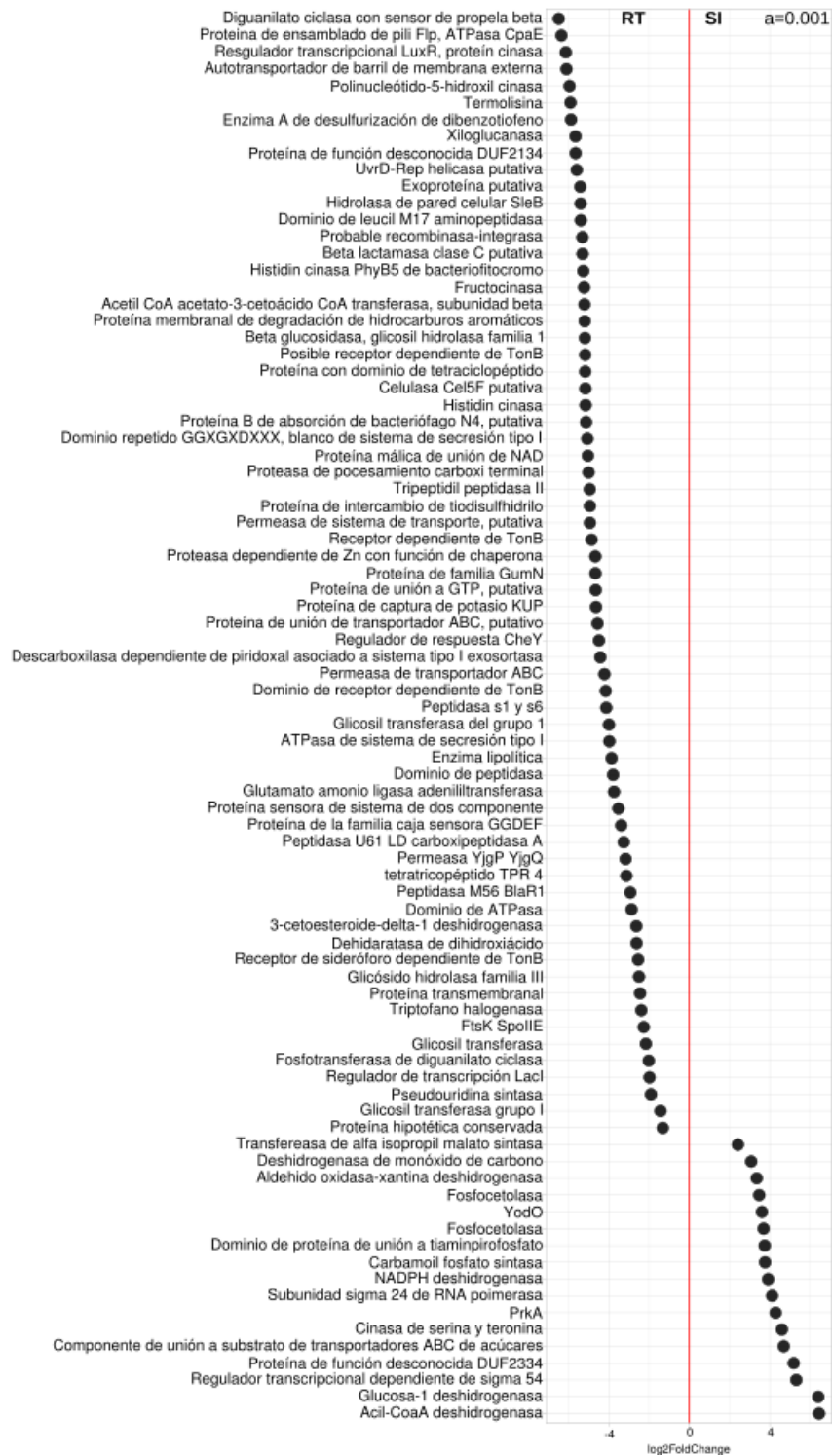


Figura 15. Proteínas diferencialmente abundantes en rizósferas de *S. lycopersicum* y suelos iniciales. Gráficos del Log2 de las veces de cambio de las abundancias de proteínas predichas anotadas en la comparación de los metagenomas de rizósferas de *S. lycopersicum* contra los metagenomas de suelos iniciales. Los valores negativos indican proteínas enriquecidas en las rizósferas de *S. lycopersicum* y los valores positivos indican proteínas enriquecidas en los suelos.

Al comparar las rizósferas de plantas ruderales y de *S. lycopersicum* pudimos hallar 16 proteínas en RT y 11 en RZ (Figura 16, Barajas *et al.*, 2020). Las proteínas enriquecidas en RZ tienen funciones relacionadas al transporte membranal (histidin cinasa de canal

osmosensible de K, transportadores y permeasas de la superfamilia MFS, transportadores ABC de azúcares), metabolismo de compuestos aromáticos (descarboxilasa de benzoilformato). Las proteínas diferencialmente abundante en RF tienen funciones involucradas en la degradación de aminoácidos (amino leucil peptidasa M17), proteínas de transferencia horizontal de material genético (e.g., integrasa-recombinasa, transposasa ISRSO17, proteína de absorción B del bacteriofago N4), reciclaje de peptidoglicano (LD descarboxilasa A) y además hemos encontrado proteínas similares a las encontradas en la comparación de RF contra SI (e.g., β -lactamasa clase C, glicósido hidrolasas) (Figura 15, Barajas *et al.*, 2020).

Finalmente, en la comparación de las rizósferas de plantas ruderales contra los suelos de origen hallamos que sólo dos proteínas predichas están enriquecidas en las plantas ruderales y participan en procesos de monitoreo ambiental (proteína sensora de un sistema de dos componentes) y metabolismo de hierro (dominio de un receptor de TonB) (Figura 17; Anexos en línea: Tabla S11, Barajas *et al.*, 2020).



Figura 16. Proteínas diferencialmente abundantes en rizósferas de *S. lycopersicum* y de plantas ruderales.

Gráficos del Log2 de las veces de cambio de las abundancias de proteínas predichas anotadas en la comparación de los metagenomas de rizósferas de *S. lycopersicum* contra los metagenomas de rizósfera de plantas ruderales. Los valores negativos indican proteínas enriquecidas en las rizósferas de *S. lycopersicum* y los valores positivos indican proteínas enriquecidas en las plantas ruderales.

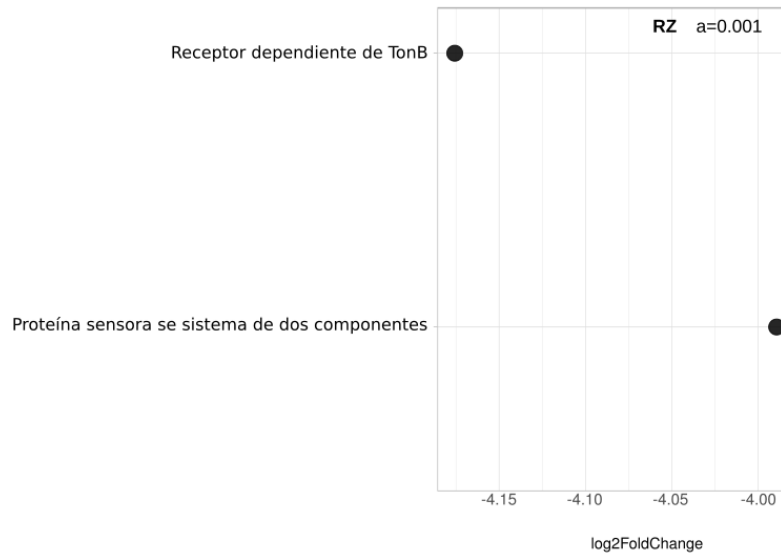


Figura 17. Proteínas diferencialmente abundantes en rizósferas de plantas ruderales y suelos iniciales. Gráficos del Log2 de las veces de cambio de las abundancias de proteínas predichas anotadas en la comparación de los metagenomas de rizósferas de plantas ruderales contra los metagenomas de suelos iniciales. Los valores negativos indican proteínas enriquecidas en las rizósferas de plantas ruderales y los valores positivos indican proteínas enriquecidas en los suelos.

El metagenoma núcleo de suelos y rizósferas de *S. lycopersicum* y plantas ruderales

Para investigar la composición del metagenoma núcleo de cada grupo (rizósferas de *S. lycopersicum*, rizósferas de plantas ruderales y suelos), en otras palabras, las proteínas predichas compartidas entre todos los metagenomas de cada grupo, se compararon los perfiles de proteínas predichas de estos microbiomas a través de diagramas de Venn. Encontramos que el metagenoma núcleo de rizósfera de *S. lycopersicum* se compone de 2,762 proteínas (Figura 18, Barajas *et al.*, 2020), mientras que el de las plantas ruderales contiene 694 y los suelos comparten 639 (Anexos en línea: Tabla S12, Barajas *et al.*, 2020).

Dentro del metagenoma núcleo de rizósfera de *S. lycopersicum* encontramos un gran número de genes codificantes para proteínas involucradas en el metabolismo central y procesos fundamentales de las células, como proteínas ribosomales, tRNA sintetisas, la acetil-CoA carboxilasa, oxoglutarato deshidrogenasa, alfa y beta glucosidasas, helicasas, ligasas, citocromo C oxidasas, entre otras. También identificamos múltiples genes codificantes interesantes involucrados en metabolismo de nitrógeno, como genes de desnitrificación (*nosZ*) y de nitrito reductasas (*nasA*, *nirB*, *nrfA*), junto con sintetisas y transferasas de glutamato y glutamina y transportadores ABC de NO_3/NO_2 . También

observamos la presencia de algunos otros genes que codifican para proteínas como la sintetasa de cianoficina (CphA), cianoficinas, amidohidrolasas de alantoato, alantoinasa y fosfolipasas de patatinas. Además de estos, identificamos distintas proteínas que participan en fenómenos de movilidad y quimiotaxis como las de biosíntesis del flagelo (FliP, FliA, FliG) y proteínas como CheY y CheB; proteínas de formación de pili (CpaE, CpaB, TadB, TadE), proteínas de ensamblaje de fimbrias y diguanilato ciclasas. Por otro lado, identificamos múltiples proteínas como transportadores ABC de azúcares, diferentes familias de glicósido hidrolasas y la enzima xiloglucanasa (Anexos en línea: Tabla S12, Barajas *et al.*, 2020).

Al analizar las proteínas compartidas entre todos los metagenomas de las plantas ruderales, observamos la misma tendencia que en el caso de *S. lycopersicum*, la mayor parte de las proteínas esenciales como chaperoninas, girasas, DNA polimerasas, proteínas ribosomales, por mencionar algunas.

Finalmente, comparamos las listas de anotación de los tres metagenomas núcleo de cada uno de los grupos mediante diagramas de Venn con el objetivo de encontrar las proteínas únicas y compartidas ubicuas en todas las muestras. Se encontró que las comunidades microbianas de rizósferas de *S. lycopersicum* cuentan con el mayor número de proteínas únicas (1,353), seguidas de los suelos (162) y las rizósferas de plantas ruderales (143). La intersección de los tres grupos consta de un total de 137 proteínas, mientras que suelos y rizósferas de *S. lycopersicum* comparten 133; suelos y rizósferas de ruderales comparten 62, y rizósferas de ruderales y *S. lycopersicum* 154. Entre los miembros que están compartidos por los tres grupos de metagenomas encontramos a proteínas consideradas esenciales para la vida como proteínas ribosomales, polimerasas de DNA y RNA, girasas y la chaperonina GroEL; siendo ubicuas en cada metagenoma (Anexos en línea: Tabla S13). Fuera de este tipo de proteínas esenciales, múltiples proteínas clasificadas como receptores de TonB y nitrilasas fueron encontradas tanto en las rizósferas de las plantas ruderales como en las de *S. lycopersicum*. Al analizar los miembros únicos de cada grupo, encontramos que, en general, volvemos a encontrar distintas proteínas esenciales en cada grupo de plantas. En las rizósferas de las plantas ruderales hay proteínas relacionadas a los sistemas de secreción tipo I, II, III y IV las cuales resultan interesantes, mientras que en las rizósferas de *S. lycopersicum* identificamos a múltiples transportadores ABC, las proteínas ExbD/ToIR sintetisas de péptidos no ribosomales (syringopeptina), las previamente mencionadas cianoficinas (CphA) y la cianoficina sintetasa, diguanilato ciclasas, proteínas de síntesis de flagelo y pili y enzimas de síntesis de polisacáridos con señal PEP-CTERM (Anexos en línea: Tabla S13, Barajas *et al.*, 2020).

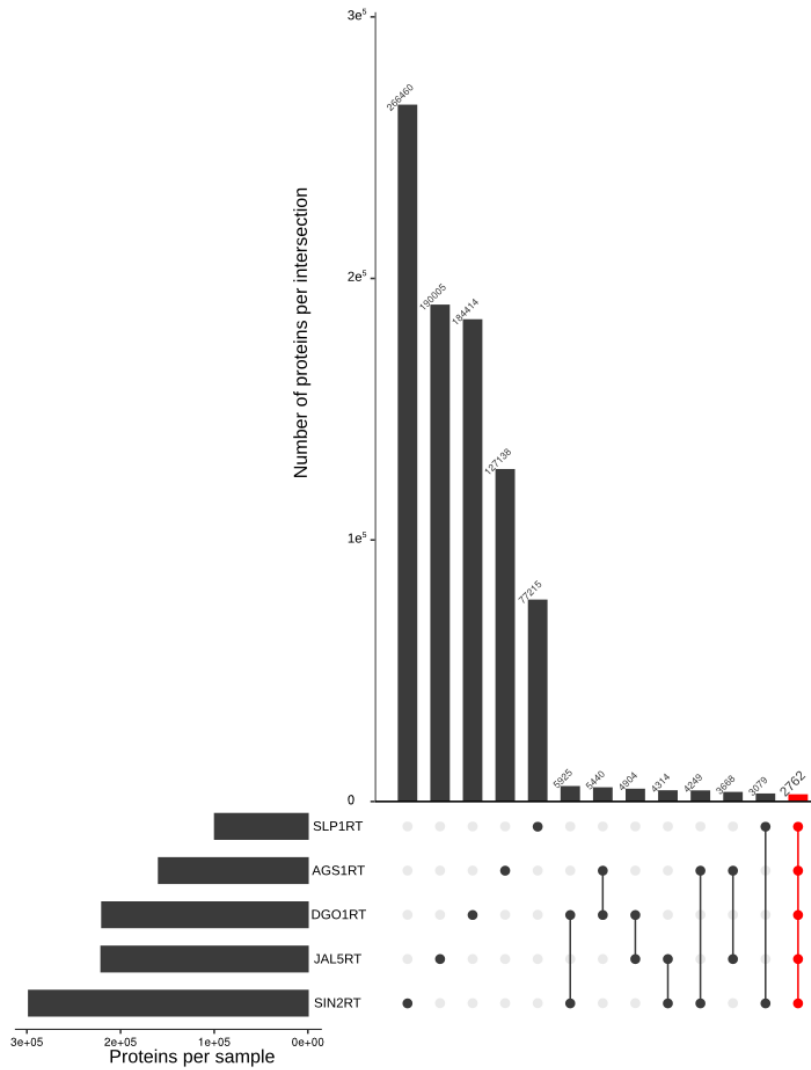


Figura 18. Metagenoma núcleo de *S. lycopersicum*. Diagrama mostrando el número de proteínas predichas únicas y compartidas entre todas las muestras de rizósferas de *S. lycopersicum*. En rojo se muestra el número de proteínas compartidas por todos los metagenomas

Discusión

En este trabajo se describió la composición y diversidad taxonómica de comunidades bacterianas de 16 muestras de suelos correspondientes a ocho distintos tipos de suelos, rizósferas y endósferas de plantas de *S. lycopersicum*, producto de un experimento de jardín común usando cada uno de los suelos como sustrato, así como suelo circundante a las plantas de jitomate y suelos control en donde no se crecieron plantas pero que recibieron el mismo tratamiento en el jardín común. Estos microbiomas fueron comparados contra los de rizósferas y endósferas de plantas ruderales, que se encontraban creciendo sobre cada uno de los suelos colectados sin ninguna condición de manejo.

Al realizar el experimento de jardín común en un invernadero se buscó someter a los múltiples tipos de suelo utilizados bajo un mismo régimen de riego y temperatura, dado que, al venir de distintos contextos geográficos, éstos se encontraban sujetos a distintas situaciones climáticas, particularmente, la cantidad de precipitación de lluvia en cada sitio. Al igual que los suelos, las plantas ruderales se encontraban sujetas a tales variables climáticas. Por otro lado, la idea de utilizar un mismo genotipo de planta en el jardín común nos permitió evaluar la influencia del huésped en la estructuración del microbioma de raíz a nivel taxonómico y metabólico.

Encontramos diferencias en las concentraciones de nutrientes totales al comparar dichas propiedades en los suelos al inicio contra los suelos al final del experimento de jardín común, en los cuales se cultivaron las plantas de *S. lycopersicum*. Particularmente, se observaron incrementos de N que varían desde una hasta 21 unidades (mg/g suelo) en más de la mitad de las muestras (Tabla 1). Esto refleja el aumento de la actividad microbiológica que puede ser explicada por el tratamiento de riego, ya que su actividad se ve favorecida con un mayor contenido de agua (Csonka., 1989). Estos incrementos pueden estar relacionados a los cambios en diversidad y abundancia de algunos grupos bacterianos, como el aumento de Proteobacterias y la disminución de Actinobacteria en los suelos finales. La presencia e incremento de estos grupos bacterianos puede contribuir en la fijación biológica de nitrógeno, ya que miembros de ambos phyla cuentan con genes codificantes de nitrogenasas en sus genomas (Boyd & Peters, 2013). Sin embargo, faltaría determinar la expresión y actividad de este tipo de enzimas en trabajos futuros. El suelo SIN2 mostró la mayor reducción en la concentración total de N (13.6 unidades), lo que puede explicarse debido a la lixiviación del nutriente, ya que este suelo era arena de una duna costera, un sustrato altamente poroso y con una pobre retención de agua. Otra explicación plausible puede ser la incorporación del

nutriente en la biomasa de las plantas, al igual que en el suelo SLP1; un suelo donde también se observó la disminución en la concentración de N. En estos dos suelos se obtuvieron las mayores producciones de biomasa en las plantas de *S. lycopersicum*. Adicionalmente, cabe mencionar que las diferencias en el desempeño de las plantas pueden haber sido producto de efectos mecánicos del suelo, ya que muchos de ellos al ser Vertisoles y Feozems presentan un alto contenido de arcilla y baja porosidad, inhibiendo el correcto desarrollo de las raíces.

Por otro lado, se observaron incrementos en la concentración total de C, indicando el incremento de la actividad biológica de las plantas al depositar compuestos de carbono a través de sus exudados de raíz, que pueden corresponder hasta el 21% del carbono fijado por las plantas (Marschner *et al.*, 2004) y que finalmente sirven como fuentes de carbono para los microorganismos del suelo (Canarini *et al.*, 2019). El mayor incremento de este nutriente en los suelos fue en SLP1, subiendo 67 unidades, en comparación con JAL1 donde el incremento es de 14 unidades. El incremento de carbono en SLP1 hace sentido con el hecho de haber sido el suelo donde las plantas tuvieron la mayor producción de biomasa y el sistema de raíces más prominente como se aprecia en el ejemplo de la Figura 3B.

Cuando buscamos las posibles relaciones entre los parámetros fisicoquímicos de los suelos con la estructura de sus microbiomas y el de las raíces, encontramos que sólo el pH fue capaz de explicar algunas diferencias en la composición lo cual se observa en la separación del clado III en la Figura 8, donde las muestras son de suelos de carácter ácido (Tabla 1). Resulta interesante notar que el pH ha sido descrito anteriormente en distintos trabajos como el mejor predictor de la estructura de comunidades bacterianas en suelos (Lauber *et al.*, 2009; Zhalnina *et al.*, 2015).

Al analizar los microbiomas de raíz (rizósfera y endósfera) de *S. lycopersicum* se encontró que éstos cuentan con valores de riqueza y diversidad (H') por debajo de los valores encontrados en los microbiomas de suelo. Este resultado va de acuerdo con el modelo de adquisición del microbioma de raíz en dos pasos (Bulgarelli *et al.*, 2015), el cual considera en primera instancia a las propiedades edáficas del suelo como un primer filtro y factor de selección de comunidades microbianas, y en segundo lugar toma en cuenta las propiedades intrínsecas del huésped. Esto se ha comprobado en organismos modelo como *Arabidopsis* (Lundberg *et al.*, 2012; Bulgarelli *et al.*, 2012) o en plantas domesticadas como cebada y arroz (Bulgarelli *et al.*, 2015; Edwards *et al.*, 2015).

La caída de diversidad y riqueza del microbioma de raíz de *S. lycopersicum* va de la

mano con el enriquecimiento de ciertos géneros bacterianos como *Caulobacter*, *Rhizobium*, *Asticcacaulis*, *Sphingobium*, *Sphingomonas*, and *Novosphingobium* los cuales han sido aislados de fuentes ambientales como cuerpos de agua (Chen *et al.*, 2013), suelos (Costa *et al.*, 2006), y rizósferas de distintas plantas (Young *et al.*, 2008; Schreiter *et al.*, 2014; Yang *et al.*, 2016). Algunos de estos géneros poseen características genómicas que pueden ser útiles para la colonización de las raíces de plantas. Por ejemplo, los géneros *Caulobacter* y *Asticcacaulis* que presentan apéndices que les facilitan la adherencia a sustratos sólidos (Poindexter, 1981) y los miembros de la familia *Sphingomonadaceae* cuentan con genes codificantes para enzimas degradadoras de compuestos complejos de carbono como lo son los derivados de plantas (Aylward *et al.*, 2013). Resulta interesante mencionar que miembros de las familias *Sphingomonadaceae* y *Bradyrhizobiaceae* han sido reportados previamente en microbiomas de *S. lycopersicum* (Tian *et al.*, 2015; Larousse *et al.* 2017), mientras que OTUs asignados al género *Sphingobium* conforman hasta el 50% de las lecturas en bibliotecas de amplicones del gen ribosomal 16S de microbiomas de jitomate (Lee *et al.*, 2016) y también se han encontrado como un género dominante en experimentos donde comparan rizósferas de jitomate bajo un déficit de hierro (Pii *et al.* 2016). En general, *Sphingobium* parece ser un género bastante cosmopolita en el contexto de plantas ya que se han aislado de rizósferas de jitomate, lechuga, maíz y *Populus* (Young *et al.* 2008; Yang *et al.* 2016; Schreiter *et al.* 2014). Además de esto, se ha visto que plantas de jitomate infectadas con el patógeno *Phytophthora parasitica* tienen una menor abundancia de las familias anteriormente señaladas en comparación con plantas sanas (Larousse *et al.*, 2017). También se ha encontrado que bacterias de este género en rizósfera confieren protección contra la podredumbre corchosa de la raíz de lechuga causado por *Sphingobium mellinum*, mostrando que al ser pertenecientes de un mismo género pueden competir por el nicho, confiriendo protección en última instancia a la planta (van Bruggen *et al.*, 2014). El hecho de que hayamos encontrado géneros sobrerrepresentados de tales familias en concomitancia con la ausencia de síntomas de pudrición de raíz en las plantas (Satour & Butler 1967) sugiere el uso de este tipo de microorganismos como indicadores de estados de salud de la planta.

Por otro lado, la riqueza y diversidad (H') observadas en los microbiomas de raíz de las plantas ruderales son significativamente superiores a las encontradas tanto en suelos como en las plantas de *S. lycopersicum*, rompiendo con los supuestos establecidos por el modelo de dos pasos; sin embargo, este tipo de resultados han sido encontrados en un meta-análisis que compara varios biomas en el cual se reportan valores más altos de diversidad en rizósferas que en suelos (Thompson *et al.*, 2017). Una de las explicaciones más plausibles es que las plantas ruderales *in situ* funcionan como un oasis de nutrientes y humedad para los microorganismos del suelo, ya que en un escenario natural como este tanto las plantas

como los microorganismos dependen fuertemente de la estacionalidad de las lluvias para poder llevar a cabo su actividad biológica, siendo el agua uno de los factores más limitantes para estos procesos (Manzoni *et al.*, 2012), lo cual resulta en el incremento en la diversidad bacteriana que observamos en las raíces de estas plantas. A pesar de esto, los microbiomas de raíz de plantas ruderales tienen una composición más parecida a los microbiomas de suelos, como es evidenciado a través del análisis de diversidad β donde se incorporan la mayoría de los microbiomas de plantas ruderales y suelos en el clado II en la Figura 8. Se conoce que el número de especies de plantas es un factor que modifica la diversidad, tasas de respiración y biomasa microbianas al incrementar la diversidad de plantas (Wu *et al.*, 2019). Esta es una consideración importante para tomar en el caso de las plantas ruderales, ya que éstas no se encontraban aisladas, sino que formaban parte de una comunidad completa de plantas formando una cobertura vegetal en donde existen también otras especies lo cual repercute en el suelo circundante, ampliando el efecto rizósfera.

Los microbiomas de raíz de *S. lycopersicum*, poseen una abundancia significativamente mayor de Proteobacteria y Bacteroidetes en comparación con los suelos y los microbiomas de raíz de las plantas ruderales que se encuentran significativamente enriquecidos en Actinobacteria. Estas diferencias en abundancia en suelos y raíces de plantas ruderales puede ser producto de factores climáticos como la ausencia de agua, la cual ha sido asociada al incremento de actinobacterias en sitios áridos, mientras que en sitios húmedos se hayan mayores proporciones de proteobacterias (Neilson *et al.*, 2017), como se observa en las raíces de *S. lycopersicum*, las cuales estuvieron recibiendo un tratamiento de riego constante. Cabe mencionar que la mayoría de las plantas ruderales que fueron colectadas, pastos, se clasificaron dentro de la familia *Poaceae*, en cuyos miembros se han reportado enriquecimientos de Actinobacteria bajo condiciones de sequía (Naylor *et al.*, 2017, Fitzpatrick *et al.*, 2018). Es importante recalcar que una diferencia radical en las comparaciones entre la estructura de microbioma de las plantas en el experimento de jardín común y las plantas ruderales es la hidratación constante de los suelos en el jardín común y que los miembros de Proteobacteria tienen un tiempo de duplicación menor que Actinobacteria, siendo favorecidos en condiciones de disponibilidad de agua (Ramin & Allison, 2019).

Además de esto, se ha visto que existen diferencias en la composición del microbioma de las plantas domesticadas en comparación con sus contrapartes salvajes; particularmente, en variedades salvajes de cebada (*Hordeum vulgare* spp. *spontaneum*) se encontró que hay una mayor abundancia de Bacteroidetes en la rizósfera, mientras que la variedad domesticada de la misma planta (*Hordeum vulgare* spp. *vulgare*) tiene un enriquecimiento de

los phyla Proteobacteria y Actinobacteria (Perez-Jaramillo *et al.*, 2018). La presión selectiva impuesta por la domesticación de las plantas ha creado diferencias en propiedades no solo de los productos que se desean obtener, sino también en características como la arquitectura de la raíz, lo cual ha sido correlacionado con la estructura de su microbioma (Perez-Jaramillo *et al.*, 2017). Algunos estudios han demostrado que existe una correlación entre la distancia evolutiva de los huéspedes y la estructura de sus microbiomas al comparar microbiomas de teocintle y otros pastos (Bouffaud *et al.* 2014; Peiffer *et al.* 2013; Redford *et al.* 2010); por lo tanto, la pérdida de Actinobacteria en los microbiomas de *S. lycopersicum*, sugiere un efecto compensatorio, producto de la domesticación.

Tomando en cuenta estos resultados es posible remarcar la importancia del papel activo del hospedero en la selección específica de sus colonizadores bacterianos. Dependiendo de factores genéticos de las plantas, los cuales varían entre especies e incluso genotipos (Baudoin *et al.*, 2002; Micallef *et al.*, 2009) diferentes microorganismos pueden ser atraídos a causa de la mezcla de exudados de raíz y propiedades particulares de sus paredes celulares (Lakshmanan *et al.*, 2012; Badri *et al.*, 2013; Carvalhais *et al.*, 2013; Lebeis *et al.*, 2015; Bulgarelli *et al.*, 2013).

A través de la secuenciación masiva de metagenomas totales fuimos capaces de analizar la composición metabólica predicha de los microbiomas de suelos y rizósferas de plantas ruderales. La riqueza de proteínas predichas observadas es equivalente en los suelos y rizósferas de ambos grupos de plantas cuando se toma en cuenta el total de genes codificantes predichos. Sin embargo, al calcular la diversidad de Shannon (H') de genes codificantes predichos encontramos que el valor de la mediana para los metagenomas de suelos se encuentra por encima de los valores de las rizósferas de plantas ruderales y de *S. lycopersicum*, no como en el caso de la información taxonómica. Esto indica en primer lugar que, a nivel metabólico, el modelo de dos pasos se cumple tanto en las plantas domesticadas como en las plantas ruderales. Esta suposición se hace con base en el valor de la mediana de los grupos, sin embargo, el grupo con la mayor dispersión de los valores de H' es el de las rizósferas de plantas ruderales, entre las cuales encontramos algunas muestras en particular que logran exceder la diversidad metabólica predicha de los suelos. Esta variabilidad puede corresponder a que las plantas ruderales colectadas pertenecen a distintos grupos filogenéticos y como se ha mencionado previamente, los microorganismos reclutados en las raíces de las plantas varían dependiendo del genotipo del huésped. Existen reportes en los cuales se ha analizado la diversidad funcional (metabólica) de suelos de pastizales y de campos de cultivo de trigo donde encontraron que los suelos de los pastizales tienen una mayor diversidad funcional en términos de genes involucrados en la degradación de materia

orgánica (Manoharan *et al.*, 2017) pero no han considerado la totalidad de genes codificantes en los metagenomas, sino que estas observaciones fueron hechas con base en proteínas anotadas contra la base de datos de Subsistemas de SEED a partir de bibliotecas generadas con primers enfocados a la identificación de dichos genes. Estos resultados también han sido observados en trabajos donde se han realizado descripciones metabólicas funcionales a partir de microarreglos con sondas de genes de metabolismo de carbono, nitrógeno, azufre y fósforo, entre otros (Watrud *et al.*, 2013). A pesar de esto, podemos sugerir que las diferencias en diversidad en las plantas domesticadas y ruderales sean también un producto de la domesticación, finalmente repercutiendo en la especialización del microbioma.

En segundo lugar, nos indica que la diversidad taxonómica no es forzosamente una aproximación válida para estimar la diversidad metabólica de las comunidades microbianas, sin embargo, aún falta investigar más a fondo cuál es la relación entre la abundancia de OTUs, sus pangenomas y los metagenomas para realizar una mejor descripción de la diversidad de genes codificantes como se ha hecho en el caso de *Prochlorococcus* (Delmont & Eren, 2018), en donde, a partir de su pangenoma se ha logrado ligar la abundancia de familias génicas específicas de distintos grupos ecológicos del género dentro de metagenomas marinos con diferentes nichos de intensidad de luz, lo cual ha ayudado a comprender mejor la ecología de este grupo taxonómico. Además de esto, es importante tomar en cuenta que la diversidad metagenómica puede estar sesgada debido a la ausencia de muchos genes en las bases de datos debido a que, en la actualidad, la mayoría de estos genes provienen de organismos cultivables que se han aislado principalmente de humanos u organismos modelo, lo cual también puede explicar el bajo porcentaje de aciertos que se pudieron encontrar en los metagenomas de plantas ruderales.

Por otro lado, también es importante mencionar que los resultados obtenidos en la anotación pueden depender tanto del ensamblado de los metagenomas como de los parámetros y bases de datos utilizadas para la asignación de taxa y funciones (Escobar-Zepeda *et al.*, 2018), por lo cual sería importante realizar un escrutinio en los parámetros para poder detectar estas diferencias en estudios posteriores.

La baja variabilidad y alta diversidad metabólica de las comunidades microbianas de los suelos sugiere que los microorganismos cuentan con repertorio extenso de genes los cuales les pueden ayudar a contender diferentes clases de estrés, principalmente abiótico, debido a la baja disponibilidad de nutrientes y agua en los suelos. Mientras que las bacterias asociadas a plantas nos sugieren una especialización metabólica que está relacionada a la colonización de la rizósfera y finalmente de las raíces de las plantas, en donde los

microorganismos requieren de características genómicas particulares para la interacción planta-bacteria (Levy *et al.*, 2017). El análisis de diversidad metabólica β de las comunidades microbianas muestra que las plantas ruderales y los suelos tienen una composición de proteínas predichas muy similar, diferenciándose de los metagenomas de rizósfera de *S. lycopersicum* lo cual evidencia que también existe una selección del microbioma a nivel metabólico cuando se trata de plantas con un genotipo fijo.

Al comparar los metagenomas para identificar los genes sobrerrepresentados entre los distintos micronichos nos encontramos con genes codificantes involucrados en procesos de movilidad y quimiotaxis en las rizósferas de *S. lycopersicum*, los cuales son importantes en los procesos de colonización como se ha visto en *Pseudomonas fluorescens* WCS36 cuando al mutar el gen *cheA*, el cual controla la quimiotaxis dirigida por flagelo, la eficiencia de colonización de raíces de jitomate es reducida (de Weert *et al.*, 2002). La alta abundancia de este tipo de genes en las rizósferas de *S. lycopersicum* hace sentido con los datos taxonómicos obtenidos que muestran una alta abundancia de Proteobacteria. Este tipo de genes se encuentran ampliamente distribuidos entre las clases α , β , y γ -Proteobacteria (Liu & Ochman, 2007). Entre otras proteínas enriquecidas en la rizósfera de *S. lycopersicum* encontramos una diguanilato ciclasa y a la proteína CpaE, las cuales están involucradas en la formación de biofilm y la producción de pili en la *Caulobacter crescentus* (Skerker & Shapiro, 2000). Otras propiedades que pueden ser relevantes para el proceso de colonización de raíces fue el enriquecimiento de una proteína identificada como xiloglucanasa, una enzima degradadora de xiloglucano, el cual es un heteropolisacárido que puede representar hasta una cuarta parte del contenido total de carbohidratos de plantas terrestres (Scheller & Ulvskov, 2010). En plantas de un mismo genotipo se ha observado el enriquecimiento de ciertos grupos bacterianos que están adaptados a estos ambientes con genes de quimiotaxis o de degradación de compuestos de pared celular de plantas (Bais *et al.* 2006; Compant *et al.*, 2010). Entre las funciones que resultaron diferencialmente abundantes en la rizósfera encontramos también al receptor dependiente de TonB, una proteína de membrana que reconoce sideróforos complejados con hierro (Noinaj *et al.* 2010) y junto con esta proteína, encontramos también en el metagenoma núcleo a ExbD, que es una proteína necesaria para la transducción de energía en el complejo de TonB para la internalización de substratos (Higgs *et al.*, 1998). Este tipo de propiedades funcionales en la rizósfera resultan relevantes debido a que se ha visto que existen fenómenos de competencia por hierro en este micronicho, donde las concentraciones de este elemento son generalmente bajas (Raaijmakers *et al.*, 2009).

Las comparaciones entre los cinco metagenomas de rizósferas de *S. lycopersicum* nos permitió encontrar la composición de su metagenoma núcleo, el cual contiene un número mucho mayor de proteínas que los metagenomas núcleo de los suelos y rizósferas de las plantas ruderales. Sugerimos que la domesticación del tomate disminuyó su diversidad genética comparado contra las plantas ruderales creciendo en el mismo suelo. Al usar un genotipo único de plantas para seleccionar su microbioma, podemos encontrar un mayor número de genes en común, mientras que comparando los microbiomas de plantas diversas encontramos un número reducido de genes comunes en sus microbiomas, debido a la diversidad genética del hospedero y a la distancia filogenética entre ellos. Además de esto, debemos considerar que debido al síndrome de domesticación se ha creado un cuello de botella en la diversidad genética de las plantas, generando líneas genéticamente más homogéneas, lo que repercute en sus características fenotípicas también, resultando en productos comestibles, libres de toxinas (Martínez *et al.*, 2020; Doebley *et al.*, 2006). De esta forma también existen descompensaciones en las habilidades para formar relaciones con algunos microorganismos, como ha sucedido con el maíz que ha perdido asociaciones micorrízicas (Toth *et al.*, 1990).

Dentro del metagenoma núcleo de *S. lycopersicum* encontramos distintas proteínas involucradas en la obtención de nitrógeno como las relacionadas a la biosíntesis de cianoficina (CphA), la cual está participa en escenarios limitantes por este nutriente. La cianoficina es un polímero de reserva compuesto por residuos de arginina y aspartato, y que aumenta la adecuación de las bacterias (Watzer & Forchhammer, 2018). Aunado a esto también se identificó a la cianoficinasa, una enzima que degrada el compuesto anteriormente mencionado (Richter *et al.*, 1999). También, encontramos genes de degradación de alantoína, una molécula que puede ser utilizada como la única fuente de nitrógeno para la producción de amonio (Cruz-Ramos *et al.*, 1997; Ma *et al.*, 2016) y genes anotados como nitrilasas, las cuales son producidas por plantas, hongos y bacterias tanto patógenas como promotoras de crecimiento vegetal (O'Reilly & Turner., 2003), catalizan la hidrólisis de nitritos, liberando amonio y pueden tener un impacto importante en la interacción planta-bacteria (Howden & Preston., 2009). Los genes *nasA*, *nirB* y *nrfA* participan en procesos que propician el almacenamiento de nitrógeno en el suelo, específicamente, el gen *NasA* es una enzima nitrato reductasa (Lin *et al.*, 1993), *NrfA* es una enzima que forma parte de la vía disimilatoria reductiva de nitrato a amonio (Putz *et al.*, 2018) y *NirB* es una nitrito reductasa que también da como producto amonio. Al generar amonio y nitrato, estas enzimas ayudan la retención de nitrógeno biodisponible en el suelo, lo cual resulta benéfico para el desarrollo de las plantas a través de la asimilación de estos recursos (Pandey *et al.*, 2020).

Proteínas anotadas como patatinas también fueron identificadas en el metagenoma núcleo de *S. lycopersicum*, las cuales son fosfolipasas que originalmente se encontraron en el genoma de papa (*Solanum tuberosum*), pero que interesantemente cuentan con homólogos en distintas bacterias (Banerji & Flieger, 2004). Este tipo de enzimas son utilizadas por las bacterias para invadir la membrana celular del hospedero mediante la secreción de efectores por el sistema de secreción tipo III (Finck-Barbançon *et al.*, 1997; Sato *et al.*, 2003; Phillips *et al.*, 2003). Es interesante notar que sus homólogos eucariontes tienen actividades antimicrobianas, específicamente inhiben el desarrollo de *Phytophthora infestans* (Bártová *et al.*, 2019) lo que da pie a una identificación más extensa de este tipo de genes provenientes de las bacterias como una estrategia para contrarrestar patógenos. Cabe mencionar que dentro de este conjunto de proteínas ubicuas en los metagenomas de rizósfera de *S. lycopersicum* encontramos también genes de biosíntesis de péptidos no ribosomales, de los cuales uno es específico para syringopeptina, una fitotoxina que es un factor de virulencia en *Pseudomonas syringae*, un patógeno común en tomate (Scholz-Schroeder *et al.*, 2001). Lo anterior muestra la prevalencia de genes que podrían relacionarse con interacciones negativas plantas-bacteria dentro del metagenoma núcleo de esta especie de planta. Múltiples genes codificantes relacionados con procesos de movilidad y quimiotaxis también se encuentran compartidos entre todos los metagenomas de rizósfera de *S. lycopersicum* y son altamente relevantes para la colonización de la raíz (Broek *et al.*, 1998; Dietel *et al.*, 2013). Por último, encontramos proteínas exosortasas y de biosíntesis de polisacáridos con señales PEP-CTERM. Estas proteínas tienen una función en la producción de cápsula y biofilm. La exosortasa reconoce a los polisacáridos con dicha señal para poder externalizarlos en la célula (Haft *et al.*, 2006, Haft *et al.*, 2012).

En resumen, las proteínas que encontramos como diferencialmente abundantes y presentes en el metagenoma núcleo de rizósfera de *S. lycopersicum* representan un conjunto de genes codificantes importantes y relevantes en la interacción planta-bacteria que cumplen funciones como el metabolismo y retención de nitrógeno en el suelo, la degradación de compuestos de carbono derivados del huésped y procesos de colonización y competencia por el nicho y se resumen en la Figura 19 (Barajas *et al.*, 2020).

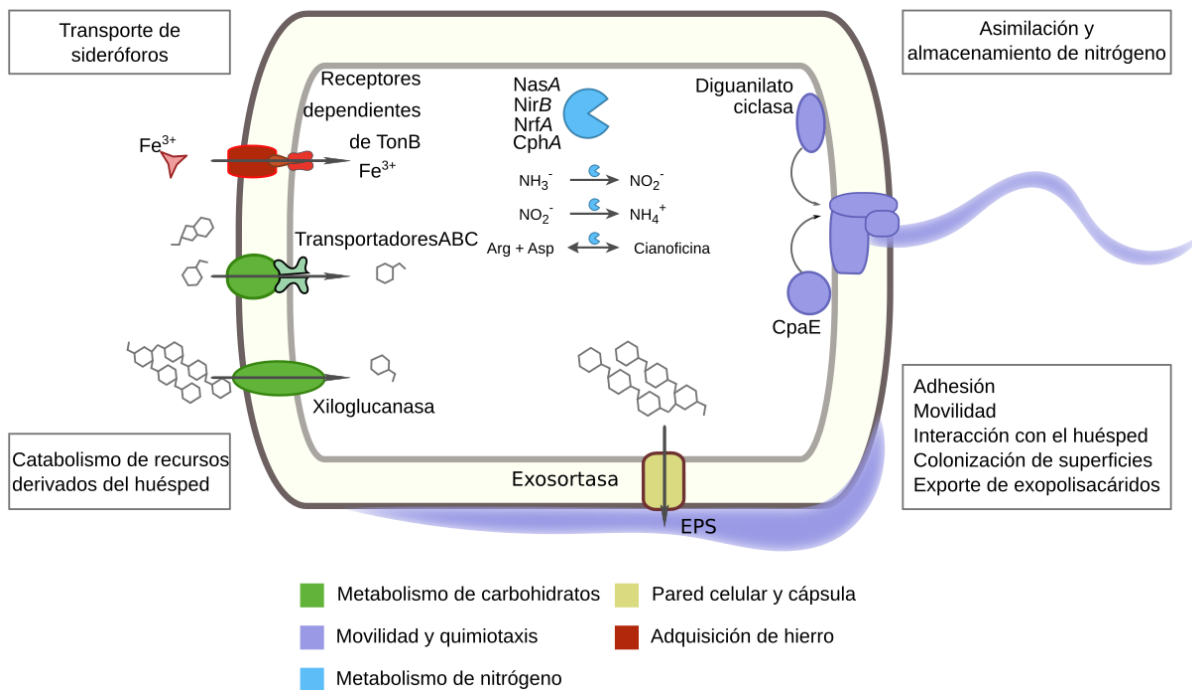


Figura 19. Esquema representativo de proteínas relevantes en el metagenoma de rizósfera de *S. lycopersicum*. El esquema presenta proteínas detectadas como diferencialmente abundantes en estos metagenomas o presentes en el metagenoma núcleo. Estas proteínas están involucradas en la colonización y competencia por el nicho de la rizósfera como la movilidad, quimiotaxis, competencia por hierro y degradación de carbohidratos complejos.

Conclusiones

Las propiedades abióticas del suelo consideradas en este estudio no explican las diferencias en la estructura de las comunidades bacterianas de raíces en plantas de distintas especies.

Bajo un escenario en donde las plantas y los suelos no reciben un manejo agronómico y están sujetos a las condiciones climáticas locales, las comunidades microbianas en raíces de plantas ruderales incrementan su diversidad taxonómica respecto al suelo de origen debido a que las plantas funcionan probablemente como un oasis de nutrientes y humedad. Esto se contrapone al modelo de adquisición del microbioma en dos pasos. En el caso de *S. lycopersicum* el modelo se cumple, donde la diversidad taxonómica de dichas comunidades decrece respecto a la del suelo, evidenciando un efecto de selección dependiente del huésped, posible producto del síndrome de domesticación.

La diversidad metabólica muestra que las comunidades microbianas de rizósferas de plantas ruderales son muy similares a las presentes en los suelos probablemente debido a un efecto rizósfera expandido ocasionado por una cobertura vegetal extensa y variada comparada contra una sola especie de planta bajo un tratamiento homogéneo.

Con 51 géneros bacterianos y 2,762 proteínas ubicuas, el metagenoma de rizósfera de *S. lycopersicum* muestra un comportamiento convergente tanto a nivel taxonómico como metabólico independientemente de la mayor diversidad de inóculos microbianos provenientes de suelos con características contrastantes. Observamos menor diversidad taxonómica, mayor convergencia de proteínas predichas compartidas del microbioma en *S. lycopersicum*, cuando se compara con los microbiomas y metagenomas de las plantas ruderales creciendo sobre el mismo suelo, sugiriendo una especialización del microbioma asociado a *S. lycopersicum*.

Las comunidades microbianas de rizósferas de *S. lycopersicum* muestran un mayor número de proteínas diferencialmente abundantes, con respecto al suelo y a las plantas ruderales, lo cual demuestra que también existe una especialización del microbioma a nivel metabólico dependiente del genotipo del hospedero. Dichos genes están involucrados en procesos como movilidad y quimiotaxis, obtención de hierro y degradación de carbohidratos complejos, lo cual sugiere que los microorganismos que colonizan la rizósfera requieren de estrategias particulares que les auxilien en la colonización de la raíz.

Perspectivas

Una parte por explorar sería estudiar la asociación de genes metagenómicos provenientes de géneros o especies bacterianas específicas. Para profundizar en este campo, se propone utilizar los pangenomas de los géneros bacterianos sobrerrepresentados para realizar el reclutamiento de genes metagenómicos y poder observar diferencias poblacionales a nivel de genes entre los nichos de rizósferas de plantas ruderales y tomates. Al ligar la información metagenómica y pangenómica sería posible identificar genes accesorios que estén asociados a las plantas ruderales y a los tomates y los cuales puedan explicar adaptaciones a cada nicho o alternativamente, que constituyan recursos útiles para su explotación en agroecosistemas. Adicionalmente, el mayor porcentaje de genes predichos en los metagenomas no tuvieron homología contra las bases de datos empleadas, los cuales, finalmente contribuyen a la amplia diversidad metagenómica observada. Por esta razón resulta conveniente realizar este puente entre la información metagenómica y pangenómica, probablemente utilizando genomas de organismos no cultivables. Además de este enfoque, una segunda opción a considerar es la generación una base de datos de proteínas predichas provenientes de metagenomas que no tengan homólogos en las bases de datos actuales, lo cual representa un recurso valioso para la identificación de familias de proteínas que puedan provenir de organismos no cultivados presentes en las comunidades microbianas de distintos ambientes; en particular, entornos tan complejos como el suelo, en el cual estamos empezando a comprender y a describir su diversidad metagenómica.

Uno de los resultados más relevantes en este trabajo es el incremento en la diversidad taxonómica de las comunidades bacterianas de rizósferas de las plantas ruderales, respecto al suelo de origen y el tomate. Un experimento interesante derivado de esto sería la obtención del microbioma rizosférico de plantas ruderales el cual será utilizado como inóculo para las plantas de tomate. De esta forma se podría maximizar la diversidad y el número de interacciones planta-bacteria con un huésped domesticado como el tomate con el objetivo de intentar recuperar parte de la diversidad microbiana que se ha perdido. Por otro lado, también es de interés realizar experimentos de jardín común con plantas ruderales (Poaceae) para evaluar la composición y diversidad del microbioma en distintos suelos bajo condiciones homogéneas, descartando diferencias producto de la diferencia de edad y condiciones climáticas en la estructuración de las comunidades microbianas de raíz.

El concepto del microbioma núcleo se ha planteado recientemente como un fundamento importante para el diseño de comunidades microbianas sintéticas que puedan ser aplicadas para el mejoramiento de la producción agrícola. En este trabajo se ha descrito el microbioma núcleo de la rizósfera de tomate bajo distintos suelos con diferentes características abióticas, albergando comunidades microbianas distintas. Esto nos ha permitido describir el conjunto de microorganismos y genes que están compartidos dentro del microbioma de raíz de una planta de interés agronómico en el contexto de una variedad de sustratos. Utilizando esta información se abre un camino que permite comenzar a diseñar comunidades bacterianas sintéticas, las cuales sean específicas para la especie del huésped y que al mismo tiempo cuentan con una serie de características metabólicas útiles para la interacción planta-bacteria y potencial benéfico para la nutrición de las plantas. Para este objetivo, es necesario realizar análisis de redes de co-ocurrencia de genes y de taxa para poder identificar aquellos que sean clave para ejercer alguna función de interés, por ejemplo, el bloqueo de patógenos, el mejoramiento de las condiciones nutrimentales para los huéspedes o el posible reclutamiento y refuerzo de otros microorganismos benéficos en la comunidad. Un primer paso para generar comunidades sintéticas es llevar a cabo el aislamiento dirigido de los géneros bacterianos como *Sphingobium*, *Caulobacter* y *Arthrobacter*, los cuales se encuentran sobrerrepresentados y compartidos entre todos los individuos de tomate bajo los 17 suelos diferentes. Este tipo de géneros bacterianos son relevantes en el contexto de plantas debido a que especies como *Arthrobacter koorensis* 5J12A y *Sphingobium* sp. AEW4 han mostrado propiedades como promotoras de crecimiento vegetal.

Para generar un cepario representativo del microbioma núcleo de la rizósfera de plantas de tomate o el de plantas ruderales, es posible tomar diferentes estrategias. Por un lado, se puede abordar un enfoque tradicional, utilizando medios selectivos para grupos bacterianos específicos que faciliten su aislamiento. En segunda instancia, se pueden generar sistemas de escrutinio de microorganismos utilizando a las plantas como factor de selección con el objetivo de favorecer el desarrollo de bacterias con un potencial funcional dentro de la comunidad. Estos experimentos pueden ser realizados en sistemas hidropónicos con tomate, los cuales facilitan el manejo de los individuos, así como llevar a cabo un monitoreo constante del desarrollo de las raíces y evaluar probables enfermedades. Al estar en un medio menos complejo que el suelo, es posible tener un mejor control sobre las condiciones abióticas que afectan a las comunidades microbianas y finalmente, se pueden obtener resultados de experimentos más controlados. Al utilizar este tipo de sistemas también sería posible comenzar a explorar el efecto en el desarrollo de las plantas durante la inoculación de diferentes microorganismos, ya sea de manera individual o en consorcios, lo

cual es relevante debido a que el tomate es una planta cultivada en sistemas hidropónicos y es importante conocer las limitantes y ventajas de la aplicación de microorganismos en este tipo de sistemas.

Referencias

- Aflitos, Saulo, Elio Schijlen, Hans De Jong, Dick De Ridder, Sandra Smit, Richard Finkers, Jun Wang, et al. 2014. "Exploring Genetic Variation in the Tomato (*Solanum Section Lycopersicon*) Clade by Whole-Genome Sequencing." *Plant Journal* 80 (1): 136–48. doi:10.1111/tpj.12616.
- Al-Busaidi, A., and P. Cookson. 2003. "Salinity – pH Relationships in Calcareous Soils of Oman. *Journal for Scientific Research.*" *Agricultural and Marine Sciences* 8 (1): 41–46.
- Alcaraz, Luis D., Mariana Peimbert, Hugo R. Barajas, Ana E. Dorantes-Acosta, John L. Bowman, and Mario A. Arteaga-Vázquez. 2018. "Marchantia Liverworts as a Proxy to Plants' Basal Microbiomes." *Scientific Reports* 8 (1): 12712. doi:10.1038/s41598-018-31168-0.
- Alcaraz, Luis David, Shamayim Martínez-Sánchez, Ignacio Torres, Enrique Ibarra-Laclette, and Luis Herrera-Estrella. 2016. "The Metagenome of *Utricularia Gibba*'s Traps: Into the Microbial Input to a Carnivorous Plant." *PLoS ONE* 11 (2): e0148979. doi:10.1371/journal.pone.0148979.
- Anderson, David M., Hiromi Sato, Aaron T. Dirck, Jimmy B. Feix, and Dara W. Frank. 2015. "Ubiquitin Activates Patatin-Like Phospholipases from Multiple Bacterial Species." Edited by G. A. O'Toole. *Journal of Bacteriology* 197 (3): 529–41. doi:10.1128/JB.02402-14.
- Badri, Dayakar V., Jacqueline M. Chaparro, Ruifu Zhang, Qirong Shen, and Jorge M. Vivanco. 2013. "Application of Natural Blends of Phytochemicals Derived from the Root Exudates of Arabidopsis to the Soil Reveal That Phenolic-Related Compounds Predominantly Modulate the Soil Microbiome." *Journal of Biological Chemistry* 288 (7): 4502–12. doi:10.1074/jbc.M112.433300.
- Badri, Dayakar V., and Jorge M. Vivanco. 2009. "Regulation and Function of Root Exudates." *Plant, Cell and Environment* 32 (6): 666–81. doi:10.1111/j.1365-3040.2009.01926.x.
- Bais, Harsh P., Tiffany L. Weir, Laura G. Perry, Simon Gilroy, and Jorge M. Vivanco. 2006. "The Role of Root Exudates in Rhizosphere Interactions With Plants and Other Organisms." *Annual Review of Plant Biology* 57 (1): 233–66. doi:10.1146/annurev.arplant.57.032905.105159.
- Banerjee, Samiran, Bobbi Helgason, Lianfeng Wang, Tristrom Winsley, Belinda C. Ferrari, and Steven D. Siciliano. 2016. "Legacy Effects of Soil Moisture on Microbial Community Structure and N₂O Emissions." *Soil Biology and Biochemistry* 95 (December). Elsevier Ltd: 40–50. doi:10.1016/j.soilbio.2015.12.004.
- Banerji, Sangeeta, and Antje Flieger. 2004. "Patatin-like Proteins: A New Family of Lipolytic Enzymes Present in Bacteria?" *Microbiology* 150 (3): 522–25. doi:10.1099/mic.0.26957-0.
- Barajas Hugo R. Fac de Químca 2016. Metagenómica de rizósferas en distintos tipos de suelos de México.
- Barajas, Hugo R., Shamayim Martínez-Sánchez, Miguel F. Romero, Cristóbal Hernández Álvarez, Luis Servín-González, Mariana Peimbert, Rocío Cruz-Ortega, Felipe García-Oliva, and Luis D. Alcaraz. 2020. "Testing the Two-Step Model of Plant Root Microbiome Acquisition Under Multiple Plant Species and Soil Sources." *Frontiers in Microbiology* 11 (October): 1–17. doi:10.3389/fmicb.2020.542742.
- Barea, J. M. and Azcon, R. 1975. "Possible synergistic interactions between endogone and phosphate-solubilizing bacteria in low-phosphate soils." p. 409-417 p. In: Mosse, B. and Tinker, P. B., eds. Endomycorrhizas. London: Academic Press.
- Bártová, Veronika, Jan Bárta, and Markéta Jarošová. 2019. "Antifungal and Antimicrobial Proteins and Peptides of Potato (*Solanum Tuberosum* L.) Tubers and Their Applications." *Applied Microbiology and Biotechnology* 103 (14): 5533–47.

- doi:10.1007/s00253-019-09887-9.
- Baudoin, E., E. Benizri, and A. Guckert. 2002. "Impact of Growth Stage on the Bacterial Community Structure along Maize Roots, as Determined by Metabolic and Genetic Fingerprinting." *Applied Soil Ecology* 19 (2): 135–45. doi:10.1016/S0929-1393(01)00185-8.
- Beleggia, Romina, Domenico Rau, Giovanni Laido, Cristiano Platani, Franca Nigro, Mariagiovanna Fragasso, Pasquale De Vita, et al. 2016. "Evolutionary Metabolomics Reveals Domestication-Associated Changes in Tetraploid Wheat Kernels." *Molecular Biology and Evolution* 33 (7): 1740–53. doi:10.1093/molbev/msw050.
- Bellucci, Elisa, Elena Bitocchi, Alberto Ferrarini, Andrea Benazzo, Eleonora Biagetti, Sebastian Klie, Andrea Minio, et al. 2014. "Decreased Nucleotide and Expression Diversity and Modified Coexpression Patterns Characterize Domestication in the Common Bean." *Plant Cell* 26 (5): 1901–12. doi:10.1105/tpc.114.124040.
- Berendsen, Roeland L., Corné M.J. Pieterse, and Peter A.H.M. Bakker. 2012. "The Rhizosphere Microbiome and Plant Health." *Trends in Plant Science* 17 (8): 478–86. doi:10.1016/j.tplants.2012.04.001.
- Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30 (15): 2114–20. doi:10.1093/bioinformatics/btu170.
- Bouffaud, Marie Lara, Marie Andrée Poirier, Daniel Muller, and Yvan Moëgne-Loccoz. 2014. "Root Microbiome Relates to Plant Host Evolution in Maize and Other *Poaceae*." *Environmental Microbiology* 16 (9): 2804–14. doi:10.1111/1462-2920.12442.
- Boyd, Eric S., and John W. Peters. 2013. "New Insights into the Evolutionary History of Biological Nitrogen Fixation." *Frontiers in Microbiology* 4 (AUG): 1–12. doi:10.3389/fmicb.2013.00201.
- Brader, Günter, Stéphane Compant, Birgit Mitter, Friederike Trognitz, and Angela Sessitsch. 2014. "Metabolic Potential of Endophytic Bacteria." *Current Opinion in Biotechnology* 27 (June): 30–37. doi:10.1016/j.copbio.2013.09.012.
- Bray, J. Roger, and J. T. Curtis. 1957. "An Ordination of the Upland Forest Communities of Southern Wisconsin." *Ecological Monographs* 27 (4): 325–49. doi:10.2307/1942268.
- Brisson, Vanessa L., Jennifer E. Schmidt, Trent R. Northen, John P. Vogel, and Amélie C.M. Gaudin. 2019. "Impacts of Maize Domestication and Breeding on Rhizosphere Microbial Community Recruitment from a Nutrient Depleted Agricultural Soil." *Scientific Reports* 9 (1): 1–14. doi:10.1038/s41598-019-52148-y.
- Brundrett, M. C. (2002). "Coevolution of roots and mycorrhizas of land plants." *New phytologist*, 154(2), 275-304. doi.org/10.1046/j.1469-8137.2002.00397.x
- Buchfink, Benjamin, Chao Xie, and Daniel H. Huson. 2015. "Fast and Sensitive Protein Alignment Using DIAMOND." *Nature Methods* 12 (1): 59–60. doi:10.1038/nmeth.3176.
- Bulgarelli, Davide, Ruben Garrido-Oter, Philipp C. Münch, Aaron Weiman, Johannes Dröge, Yao Pan, Alice C. McHardy, and Paul Schulze-Lefert. 2015. "Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley." *Cell Host and Microbe* 17 (3): 392–403. doi:10.1016/j.chom.2015.01.011.
- Bulgarelli, Davide, Matthias Rott, Klaus Schlaeppi, Emiel Ver Loren van Themaat, Nahal Ahmadinejad, Federica Assenza, Philipp Rauf, et al. 2012. "Revealing Structure and Assembly Cues for Arabidopsis Root-Inhabiting Bacterial Microbiota." *Nature* 488 (7409). Nature Publishing Group: 91–95. doi:10.1038/nature11336.
- Bulgarelli, Davide, Klaus Schlaeppi, Stijn Spaepen, Emiel Ver Loren van Themaat, and Paul Schulze-Lefert. 2013. "Structure and Functions of the Bacterial Microbiota of Plants." *Annual Review of Plant Biology* 64 (1): 807–38. doi:10.1146/annurev-arplant-050312-120106.
- Canarini, Alberto, Christina Kaiser, Andrew Merchant, Andreas Richter, and Wolfgang Wanek. 2019. "Root Exudation of Primary Metabolites: Mechanisms and Their Roles in Plant Responses to Environmental Stimuli." *Frontiers in Plant Science* 10 (February). doi:10.3389/fpls.2019.00157.
- Caporaso, J. Gregory, Justin Kuczynski, Jesse Stombaugh, Kyle Bittinger, Frederic D.

- Bushman, Elizabeth K. Costello, Noah Fierer, et al. 2010. "QIIME Allows Analysis of High-Throughput Community Sequencing Data." *Nature Methods* 7 (5). Nature Publishing Group: 335–36. doi:10.1038/nmeth.f.303.
- Carvalhais, Lilia C., Paul G. Dennis, Dayakar V. Badri, Gene W. Tyson, Jorge M. Vivanco, and Peer M. Schenk. 2013. "Activation of the Jasmonic Acid Plant Defence Pathway Alters the Composition of Rhizosphere Bacterial Communities." *PLoS ONE* 8 (2): 1–5. doi:10.1371/journal.pone.0056457.
- Chen, Hong, Mareike Jogler, Manfred Rohde, Hans Peter Klenk, Hans Jürgen Busse, Brian J. Tindall, Cathrin Spröer, and Jörg Overmann. 2013. "*Sphingobium limneticum* Sp. Nov. and *Sphingobium boeckii* Sp. Nov., Two Freshwater Planktonic Members of the Family *Sphingomonadaceae*, and Reclassification of *Sphingomonas suberifaciens* as *Sphingobium suberifaciens* Comb. Nov." *International Journal of Systematic and Evolutionary Microbiology* 63 (PART2): 735–43. doi:10.1099/ijs.0.040105-0.
- Chen, Lin, Xiuli Xin, Jiabao Zhang, Marc Redmile-Gordon, Guangsen Nie, and Qingyun Wang. 2019. "Soil Characteristics Overwhelm Cultivar Effects on the Structure and Assembly of Root-Associated Microbiomes of Modern Maize." *Pedosphere* 29 (3). Soil Science Society of China: 360–73. doi:10.1016/S1002-0160(17)60370-9.
- Cheng, Zhiqiang, Shaonan Lei, Ye Li, Wei Huang, Rongqin Ma, Juan Xiong, Ting Zhang, et al. 2020. "Revealing the Variation and Stability of Bacterial Communities in Tomato Rhizosphere Microbiota." *Microorganisms* 8 (2): 1–15. doi:10.3390/microorganisms8020170.
- Coleman-Derr, Devin, Damaris Desgarennes, Citlali Fonseca-Garcia, Stephen Gross, Scott Clingenpeel, Tanja Woyke, Gretchen North, Axel Visel, Laila P. Partida-Martinez, and Susannah G. Tringe. "Plant compartment and biogeography affect microbiome composition in cultivated and native Agave species." *New Phytologist* 209, no. 2 (2016): 798-811.
- Compant, Stéphane, Christophe Clément, and Angela Sessitsch. 2010. "Plant Growth-Promoting Bacteria in the Rhizo- and Endosphere of Plants: Their Role, Colonization, Mechanisms Involved and Prospects for Utilization." *Soil Biology and Biochemistry* 42 (5): 669–78. doi:10.1016/j.soilbio.2009.11.024.
- Conway, Jake R., Alexander Lex, and Nils Gehlenborg. 2017. "UpSetR: An R Package for the Visualization of Intersecting Sets and Their Properties." Edited by John Hancock. *Bioinformatics* 33 (18): 2938–40. doi:10.1093/bioinformatics/btx364.
- Costa, Rodrigo, Monika Götz, Nicole Mrotzek, Jana Lottmann, Gabriele Berg, and Kornelia Smalla. 2006. "Effects of Site and Plant Species on Rhizosphere Community Structure as Revealed by Molecular Analysis of Microbial Guilds." *FEMS Microbiology Ecology* 56 (2): 236–49. doi:10.1111/j.1574-6941.2005.00026.x.
- Crowther, T. W., J. van den Hoogen, J. Wan, M. A. Mayes, A. D. Keiser, L. Mo, C. Averill, and D. S. Maynard. 2019. "The Global Soil Community and Its Influence on Biogeochemistry." *Science* 365 (6455). doi:10.1126/science.aav0550.
- Cruz-Ramos, Hugo, Phillipe Glaser, Lewis V. Wray, and Susan H. Fisher. 1997. "The *Bacillus subtilis* ureABC Operon." *Journal of Bacteriology* 179 (10): 3371–73. doi:10.1128/JB.179.10.3371-3373.1997.
- Csonka, L. N. 1989. "Physiological and Genetic Responses of Bacteria to Osmotic Stress." *Microbiological Reviews* 53 (1): 121–47. doi:10.1128/mmr.53.1.121-147.1989.
- Daniel, Rolf. 2005. "The Metagenomics of Soil." *Nature Reviews. Microbiology* 3 (6): 470–78. doi:10.1038/nrmicro1160.
- Dakora, Felix D, and Donald A Phillips. 2002. "Root Exudates as Mediators of Mineral Acquisition in Low-Nutrient Environments." *Plant and Soil* 245 (1): 35–47. doi:10.1023/A:1020809400075.
- de Weert, Sandra, Hans Vermeiren, Ine H. M. Mulders, Irene Kuiper, Nico Hendrickx, Guido V. Bloemberg, Jos Vanderleyden, René De Mot, and Ben J. J. Lugtenberg. 2002. "Flagella-Driven Chemotaxis Towards Exudate Components Is an Important Trait for Tomato Root Colonization by *Pseudomonas fluorescens*." *Molecular Plant-Microbe Interactions* 15 (11): 1173–80. doi:10.1094/MPMI.2002.15.11.1173.

- Delmont, Tom O., and E. Murat Eren. 2018. "Linking Pangenomes and Metagenomes: The *Prochlorococcus* Metapangenome." *PeerJ* 2018 (1): e4320. doi:10.7717/peerj.4320.
- Dennis, Paul G., Anthony J. Miller, and Penny R. Hirsch. 2010. "Are Root Exudates More Important than Other Sources of Rhizodeposits in Structuring Rhizosphere Bacterial Communities?" *FEMS Microbiology Ecology* 72 (3): 313–27. doi:10.1111/j.1574-6941.2010.00860.x.
- DeSantis, T. Z., P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen. 2006. "Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB." *Applied and Environmental Microbiology* 72 (7): 5069–72. doi:10.1128/AEM.03006-05.
- Dietel, Kristin, Barbara Beator, Anto Budiharjo, Ben Fan, and Rainer Borriss. 2013. "Bacterial Traits Involved in Colonization of *Arabidopsis thaliana* Roots by *Bacillus amyloliquefaciens* FZB42." *Plant Pathology Journal* 29 (1): 59–66. doi:10.5423/PPJ.OA.10.2012.0155.
- Doebley, John F., Brandon S. Gaut, and Bruce D. Smith. 2006. "The Molecular Genetics of Crop Domestication." *Cell*. doi:10.1016/j.cell.2006.12.006.
- Dong, Chun Juan, Ling Ling Wang, Qian Li, and Qing Mao Shang. 2019. "Bacterial Communities in the Rhizosphere, Phyllosphere and Endosphere of Tomato Plants." *PLoS ONE* 14 (11): 1–17. doi:10.1371/journal.pone.0223847.
- Edwards, Joseph, Cameron Johnson, Christian Santos-Medellín, Eugene Lurie, Natraj Kumar Podishetty, Srijak Bhatnagar, Jonathan A. Eisen, Venkatesan Sundaresan, and L. Dangl Jeffery. 2015. "Structure, Variation, and Assembly of the Root-Associated Microbiomes of Rice." *Proceedings of the National Academy of Sciences of the United States of America* 112 (8): E911–20. doi:10.1073/pnas.1414592112.
- Elphinstone, J. G. 2005. "The current bacterial wilt situation: A global overview." Pages 9-28 in: *Bacterial Wilt Disease and the *Ralstonia solanacearum* Species Complex*. C. Allen, P. Prior, and A. C. Hayward, eds. American Phytopathological Society, St. Paul, MN.
- El-Tarabily, K. A., A. H. Nassar, G. E. St. J. Hardy, and K. Sivasithamparam. 2009. "Plant Growth Promotion and Biological Control of *Pythium aphanidermatum*, a Pathogen of Cucumber, by Endophytic Actinomycetes." *Journal of Applied Microbiology* 106 (1): 13–26. doi:10.1111/j.1365-2672.2008.03926.x.
- Engelbrektson, Anna, Victor Kunin, Anna Engelbrektson, Victor Kunin, Tijana Glavina del Rio, Philip Hugenholtz, and Susannah Green Tringe. 2012. "Defining the Core *Arabidopsis thaliana* Root Microbiome." *Nature* 488 (7409). Nature Publishing Group: 86–90. doi:10.1038/nature11237.
- Escobar-Zepeda, Alejandra, Elizabeth Ernestina Godoy-Lozano, Luciana Raggi, Lorenzo Segovia, Enrique Merino, Rosa María Gutiérrez-Rios, Katy Juarez, Alexei F. Licea-Navarro, Liliana Pardo-Lopez, and Alejandro Sanchez-Flores. "Analysis of sequencing strategies and tools for taxonomic annotation: defining standards for progressive metagenomics." *Scientific reports* 8, no. 1 (2018): 1-13.
- Fakruddin, M. 2015. "Methods for Analyzing Diversity of Microbial Communities in Natural Environments." *Microbes in Soil and Their Agricultural Prospects* 42 (1): 123–51. doi:10.4038/cjsbs.v42i1.5896.
- Fernandez-Illescas, Coral Pilar, Amilcare Porporato, Francesco Laio, and Ignacio Rodriguez-Iturbe. 2001. "The Ecohydrological Role of Soil Texture in a Water-Limited Ecosystem." *Water Resources Research* 37 (12): 2863–72. doi:10.1029/2000WR000121.
- Fierer, Noah. 2017. "Embracing the Unknown: Disentangling the Complexities of the Soil Microbiome." *Nature Reviews Microbiology* 15 (10). Nature Publishing Group: 579–90. doi:10.1038/nrmicro.2017.87.
- Fierer, Noah, J. W. Leff, B. J. Adams, U. N. Nielsen, S. T. Bates, C. L. Lauber, S. Owens, J. A. Gilbert, D. H. Wall, and J. G. Caporaso. 2012. "Cross-Biome Metagenomic Analyses of Soil Microbial Communities and Their Functional Attributes." *Proceedings of the National Academy of Sciences* 109 (52): 21390–95. doi:10.1073/pnas.1215210110.
- Fierer, Noah, Christian L Lauber, Kelly S Ramirez, Jesse Zaneveld, Mark a Bradford, and Rob Knight. 2012. "Comparative Metagenomic, Phylogenetic and Physiological Analyses

- of Soil Microbial Communities across Nitrogen Gradients." *The ISME Journal* 6 (5). Nature Publishing Group: 1007–17. doi:10.1038/ismej.2011.159.
- Fierer, Noah, and Robert B. Jackson. 2006. "The Diversity and Biogeography of Soil Bacterial Communities." *Proceedings of the National Academy of Sciences of the United States of America* 103 (3): 626–31. doi:10.1073/pnas.0507535103.
- Finck-Barbançon, Viviane, Joanne Goranson, Lei Zhu, Teiji Sawa, Jeanine P. Wiener-Kronish, Suzanne M. J. Fleiszig, Christine Wu, Liane Mende-Mueller, and Dara W. Frank. 1997. "ExoU Expression by *Pseudomonas aeruginosa* Correlates with Acute Cytotoxicity and Epithelial Injury." *Molecular Microbiology* 25 (3): 547–57. doi:10.1046/j.1365-2958.1997.4891851.x.
- Fitzpatrick, Connor R., Isai Salas-González, Jonathan M. Conway, Omri M. Finkel, Sarah Gilbert, Dor Russ, Paulo José Pereira Lima Teixeira, and Jeffery L. Dangl. 2020. "The Plant Microbiome: From Ecology to Reductionism and beyond." *Annual Review of Microbiology* 74: 81–100. doi:10.1146/annurev-micro-022620-014327.
- Fitzpatrick, Connor R., Julia Copeland, Pauline W. Wang, David S. Guttman, Peter M. Kotanen, and Marc T.J. Johnson. 2018a. "Assembly and Ecological Function of the Root Microbiome across Angiosperm Plant Species." *Proceedings of the National Academy of Sciences of the United States of America* 115 (6): E1157–65. doi:10.1073/pnas.1717617115.
- Flores-Núñez, Víctor M., Citlali Fonseca-García, Damaris Desgarenes, Emiley Eloefadros, Tanja Woyke, and Laila P. Partida-Martínez. "Functional signatures of the epiphytic prokaryotic microbiome of agaves and cacti." *Frontiers in microbiology* 10 (2020): 3044.
- Fowler, Jonathan H., Javier Narváez-Vásquez, Dale N. Aromdee, Véronique Pautot, Frances M. Holzer, and Linda L. Walling. 2009. "Leucine Aminopeptidase Regulates Defense and Wound Signaling in Tomato Downstream of Jasmonic Acid." *The Plant Cell* 21 (4): 1239–51. doi:10.1105/tpc.108.065029.
- Garbeva, P, J a van Veen, and J D van Elsas. 2004. "Microbial Diversity in Soil: Selection Microbial Populations by Plant and Soil Type and Implications for Disease Suppressiveness." *Annual Review of Phytopathology* 42 (29): 243–70. doi:10.1146/annurev.phyto.42.012604.135455.
- Fonseca-García, Citlali, Damaris Desgarenes, Víctor M. Flores-Núñez, and Laila P. Partida-Martínez. "The microbiome of desert CAM plants: lessons from amplicon sequencing and metagenomics." In *Metagenomics*, pp. 231-254. Academic Press, 2018.
- Fulthorpe, R. R., A. N. Rhodes, and J. M. Tiedje. 1998. "High Levels of Endemicity of 3-Chlorobenzoate-Degrading Soil Bacteria." *Applied and Environmental Microbiology* 64 (5): 1620–27. doi:10.1128/aem.64.5.1620-1627.1998.
- Girvan, Martina S., Juliet Bullimore, Jules N. Pretty, A. Mark Osborn, and Andrew S. Ball. 2003. "Soil Type Is the Primary Determinant of the Composition of the Total and Active Bacterial Communities in Arable Soils." *Applied and Environmental Microbiology* 69 (3): 1800–1809. doi:10.1128/AEM.69.3.1800-1809.2003.
- Graham, Katherine E., Aaron J. Prussin, Linsey C. Marr, Lauren M. Sassoubre, and Alexandria B. Boehm. 2018. "Microbial Community Structure of Sea Spray Aerosols at Three California Beaches." *FEMS Microbiology Ecology* 94 (3): 1–10. doi:10.1093/femsec/fiy005.
- Haft, Daniel H., Ian T. Paulsen, Naomi Ward, and Jeremy D. Selengut. 2006. "Exopolysaccharide-Associated Protein Sorting in Environmental Organisms: The PEP-CTERM/SpsH System. Application of a Novel Phylogenetic Profiling Heuristic." *BMC Biology* 4: 1–16. doi:10.1186/1741-7007-4-29.
- Haft, Dniel H., Samuel H. Payne, and Jeremy D. Selengut. 2012. "Archaeosortases and Exosortases Are Widely Distributed Systems Linking Membrane Transit with Posttranslational Modification." *Journal of Bacteriology* 194 (1): 36–48. doi:10.1128/JB.06026-11.
- Hartmann, Anton, Michael Rothballer, and Michael Schmid. 2008. "Lorenz Hiltner, a Pioneer in Rhizosphere Microbial Ecology and Soil Bacteriology Research." *Plant and Soil* 312

- (1–2): 7–14. doi:10.1007/s11104-007-9514-z.
- Hayat, Rifat, Safdar Ali, Ummay Amara, Rabia Khalid, and Iftikhar Ahmed. 2010. "Soil Beneficial Bacteria and Their Role in Plant Growth Promotion: A Review." *Annals of Microbiology* 60 (4): 579–98. doi:10.1007/s13213-010-0117-1.
- Higgs, Penelope I., Paul S. Myers, and Kathleen Postle. 1998. "Interactions in the TonB-Dependent Energy Transduction Complex: ExbB and ExbD Form Homomultimers." *Journal of Bacteriology* 180 (22): 6031–38. doi:10.1128/jb.180.22.6031-6038.1998.
- Hill, Tom C.J., Kerry A. Walsh, James A. Harris, and Bruce F. Moffett. 2003. "Using Ecological Diversity Measures with Bacterial Communities." *FEMS Microbiology Ecology* 43 (1): 1–11. doi:10.1111/j.1574-6941.2003.tb01040.x.
- Howden, Andrew J.M., and Gail M. Preston. 2009. "Nitrilase Enzymes and Their Role in Plant-Microbe Interactions." *Microbial Biotechnology* 2 (4): 441–51. doi:10.1111/j.1751-7915.2009.00111.x.
- Hueck, Christoph J. 1998. "Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants." *Microbiology and Molecular Biology Reviews* 62 (2): 379–433. doi:10.1128/mmbr.62.2.379-433.1998.
- Huffman, Edward W.D. 1977. "Performance of a New Automatic Carbon Dioxide Coulometer." *Microchemical Journal* 22 (4): 567–73. doi:10.1016/0026-265X(77)90128-X.
- Hyatt, Doug, Gwo Liang Chen, Philip F. LoCascio, Miriam L. Land, Frank W. Larimer, and Loren J. Hauser. 2010. "Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification." *BMC Bioinformatics* 11 (1): 119. doi:10.1186/1471-2105-11-119.
- Kennedy, Katherine, Michael W. Hall, Michael D.J. Lynch, Gabriel Moreno-Hagelsieb, and Josh D. Neufeld. 2014. "Evaluating Bias of Illumina-Based Bacterial 16S rRNA Gene Profiles." *Applied and Environmental Microbiology* 80 (18): 5717–22. doi:10.1128/AEM.01451-14.
- Kieft, Thomas L., Edith soroker, and Mary K. firestone. 1987. "Microbial Biomass Response to a Rapid Increase in Water Potential When Dry Soil Is Wetted." *Soil Biology and Biochemistry* 19 (2): 119–26. doi:10.1016/0038-0717(87)90070-8.
- Kuramae, Eiko E., Etienne Yergeau, Lina C. Wong, Agata S. Pijl, Johannes A. Van Veen, and George A. Kowalchuk. 2012. "Soil Characteristics More Strongly Influence Soil Bacterial Communities than Land-Use Type." *FEMS Microbiology Ecology* 79 (1): 12–24. doi:10.1111/j.1574-6941.2011.01192.x.
- Legendre, Pierre, Daniel Borcard, and Pedro R Peres-Neto. 2005. "Concepts & Synthesis Emphasizing New Ideas To Stimulate Research in Ecology Analyzing Beta Diversity: Partitioning the Spatial Variation of Community Composition Data." *Ecological Monographs* 75 (4): 435–50.
- Miller, Elizabeth Theresa, Richard Svanbäck, and Brendan J.M. Bohannan. 2018. "Microbiomes as Metacommunities: Understanding Host-Associated Microbes through Metacommunity Ecology." *Trends in Ecology and Evolution* 33 (12). Elsevier Ltd: 926–35. doi:10.1016/j.tree.2018.09.002
- Iannucci, Anna, Mariagiovanna Fragasso, Romina Beleggia, Franca Nigro, and Roberto Papa. 2017. "Evolution of the Crop Rhizosphere: Impact of Domestication on Root Exudates in Tetraploid Wheat (*Triticum turgidum* L.)." *Frontiers in Plant Science* 8 (December). doi:10.3389/fpls.2017.02124.
- Inceoğlu, Özgül, Waleed Abu Al-Soud, Joana Falcão Salles, Alexander V. Semenov, and Jan Dirk van Elsas. 2011. "Comparative Analysis of Bacterial Communities in a Potato Field as Determined by Pyrosequencing." Edited by Jack Anthony Gilbert. *PLoS ONE* 6 (8): e23321. doi:10.1371/journal.pone.0023321.
- IUSS Working Group WRB. 2015. World Reference Base for Soil Resources 2014, update 2015. 750 'International soil classification system for naming soils and creating legends for soil maps'. 751 World Soil Resources Reports No. 106. FAO, Rome.
- Kamutando, Casper N., Surendra Vikram, Gilbert Kamgan-Nkuekam, Thulani P. Makhalanyaane, Michelle Greve, Johannes J. Le Roux, David M. Richardson, Don A. Cowan, and Angel Valverde. 2019. "The Functional Potential of the Rhizospheric

- Microbiome of an Invasive Tree Species, *Acacia dealbata*." *Microbial Ecology* 77 (1): 191–200. doi:10.1007/s00248-018-1214-0.
- Katz, Micah, Bradley M. Hover, and Sean F. Brady. 2016. "Culture-Independent Discovery of Natural Products from Soil Metagenomes." *Journal of Industrial Microbiology and Biotechnology* 43 (2–3). Springer Berlin Heidelberg: 129–41. doi:10.1007/s10295-015-1706-6.
- Kennedy K., M. W. Hall, M. D. J. Lynch, G. Moreno-Hagelsieb, and J. D. Neufeld, "Evaluating bias of Illumina-based bacterial 16S rRNA gene profiles.," *Applied and environmental microbiology*, vol. 80, no. 18, pp. 5717– 5722, 2014.
- Lakshmanan, Venkatachalam, Sherry L. Kitto, Jeffrey L. Caplan, Yi Huang Hsueh, Daniel B. Kearns, Yu Sung Wu, and Harsh P. Bais. 2012. "Microbe-Associated Molecular Patterns-Triggered Root Responses Mediate Beneficial Rhizobacterial Recruitment in Arabidopsis." *Plant Physiology* 160 (3): 1642–61. doi:10.1104/pp.112.200386.
- Lamelas, Araceli, Damaris Desgarnes, Daniel López-Lima, Luc Villain, Alexandro Alonso-Sánchez, Alejandro Artacho, Amparo Latorre, Andrés Moya, and Gloria Carrión. 2020. "The Bacterial Microbiome of *Meloidogyne*-Based Disease Complex in Coffee and Tomato." *Frontiers in Plant Science* 11 (February): 1–13. doi:10.3389/fpls.2020.00136.
- Lang, R., 1920. *Verwitterung und bodenbildung als einfuehrung in die bodenkunde*. 759 Schweizerbart Science Publishers, Stuttgart.
- Langmead, Ben, and Steven L. Salzberg. 2012. "Fast Gapped-Read Alignment with Bowtie 2." *Nature Methods* 9 (4): 357–59. doi:10.1038/nmeth.1923.
- Larousse, Marie, Corinne Rancurel, Camille Syska, Ferran Palero, Catherine Etienne, Benoît Industri, Xavier Nesme, Marc Bardin, and Eric Galiana. 2017. "Tomato Root Microbiota and *Phytophthora parasitica*-Associated Disease." *Microbiome* 5 (1). Microbiome: 56. doi:10.1186/S40168-017-0273-7.
- Lauber, Christian L., Micah Hamady, Rob Knight, and Noah Fierer. 2009. "Pyrosequencing-Based Assessment of Soil pH as a Predictor of Soil Bacterial Community Structure at the Continental Scale." *Applied and Environmental Microbiology* 75 (15): 5111–20. doi:10.1128/AEM.00335-09.
- Lebeis, Sarah L., Sur Herrera Paredes, Derek S. Lundberg, Natalie Breakfield, Jase Gehring, Meredith McDonald, Stephanie Malfatti, et al. 2015. "Salicylic Acid Modulates Colonization of the Root Microbiome by Specific Bacterial Taxa." *Science* 349 (6250): 860–64. doi:10.1126/science.aaa8764.
- Lee, Shin Ae, Yiseul Kim, Jeong Myeong Kim, Bora Chu, Jae-Ho Joa, Mee Kyung Sang, Jaekyeong Song, and Hang-Yeon Weon. 2019. "A Preliminary Examination of Bacterial, Archaeal, and Fungal Communities Inhabiting Different Rhizocompartments of Tomato Plants under Real-World Environments." *Scientific Reports* 9 (1): 9300. doi:10.1038/s41598-019-45660-8.
- Lee, Shin Ae, Jiyoung Park, Bora Chu, Jeong Myeong Kim, Jae Ho Joa, Mee Kyung Sang, Jaekyeong Song, and Hang Yeon Weon. 2016. "Comparative Analysis of Bacterial Diversity in the Rhizosphere of Tomato by Culture-Dependent and -Independent Approaches." *Journal of Microbiology* 54 (12): 823–31. doi:10.1007/s12275-016-6410-3.
- Li, Weizhong, and Adam Godzik. 2006. "Cd-Hit: A Fast Program for Clustering and Comparing Large Sets of Protein or Nucleotide Sequences." *Bioinformatics* 22 (13): 1658–59. doi:10.1093/bioinformatics/btl158.
- Lin, J. T., B. S. Goldman, and V. Stewart. 1993. "Structures of Genes *nasA* and *nasB*, Encoding Assimilatory Nitrate and Nitrite Reductases in *Klebsiella pneumoniae* M5a1." *Journal of Bacteriology* 175 (8): 2370–78. doi:10.1128/jb.175.8.2370-2378.1993.
- Liu, Renyi, and Howard Ochman. 2007. "Origins of Flagellar Gene: Operons and Secondary Flagellar Systems." *Journal of Bacteriology* 189 (19): 7098–7104. doi:10.1128/JB.00643-07.
- Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2." *Genome Biology* 15 (12): 550. doi:10.1186/s13059-014-0550-8.

- Lozupone, Catherine A., Micah Hamady, Scott T. Kelley, and Rob Knight. 2007. "Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities." *Applied and Environmental Microbiology* 73 (5): 1576–85. doi:10.1128/AEM.01996-06.
- Lozupone, Catherine, Micah Hamady, and Rob Knight. 2006. "UniFrac - An Online Tool for Comparing Microbial Community Diversity in a Phylogenetic Context." *BMC Bioinformatics* 7 (1): 371. doi:10.1186/1471-2105-7-371.
- Ma, Li Jun, David M. Geiser, Robert H. Proctor, Alejandro P. Rooney, Kerry O'Donnell, Frances Trail, Donald M. Gardiner, John M. Manners, and Kemal Kazan. 2013. "Fusarium Pathogenomics." *Annual Review of Microbiology* 67: 399–416. doi:10.1146/annurev-micro-092412-155650.
- Ma, Pikyee, Simon G. Patching, Ekaterina Ivanova, Jocelyn M. Baldwin, David Sharples, Stephen A. Baldwin, and Peter J. F. Henderson. 2016. "Allantoin Transport Protein, Pucl, from *Bacillus subtilis*: Evolutionary Relationships, Amplified Expression, Activity and Specificity." *Microbiology* 162 (5): 823–36. doi:10.1099/mic.0.000266.
- Magurran, A.E. 2004. *Measuring Biological Diversity*. Blackwell.
- Manoharan, Lokeshwaran, Sandeep K. Kushwaha, Dag Ahrén, and Katarina Hedlund. 2017. "Agricultural Land Use Determines Functional Genetic Diversity of Soil Microbial Communities." *Soil Biology and Biochemistry* 115: 423–32. doi:10.1016/j.soilbio.2017.09.011.
- Manzoni, Stefano, Joshua P. Schimel, and Amilcare Porporato. 2012. "Responses of Soil Microbial Communities to Water Stress: Results from a Meta-Analysis." *Ecology* 93 (4): 930–38. doi:10.1890/11-0026.1.
- Marasco, Ramona, María J. Mosqueira, Marco Fusi, Jean Baptiste Ramond, Giuseppe Merlino, Jenny M. Booth, Gillian Maggs-Kölling, Don A. Cowan, and Daniele Daffonchio. 2018. "Rhizosphere Microbial Community Assembly of Sympatric Desert Speargrasses Is Independent of the Plant Host." *Biological Sciences* 0605 Microbiology 06 Biological Sciences 0602 Ecology 06 Biological Sciences 0607 Plant Biology." *Microbiome* 6 (1). Microbiome: 1–18. doi:10.1186/s40168-018-0597-y.
- Marques, Joana M., Thais F. da Silva, Renata E. Vollu, Arie F. Blank, Guo Chun Ding, Lucy Seldin, and Kornelia Smalla. 2014. "Plant Age and Genotype Affect the Bacterial Community Composition in the Tuber Rhizosphere of Field-Grown Sweet Potato Plants." *FEMS Microbiology Ecology* 88 (2): 424–35. doi:10.1111/1574-6941.12313.
- Marschner, Petra, David Crowley, and Ching Hong Yang. 2004. "Development of Specific Rhizosphere Bacterial Communities in Relation to Plant Species, Nutrition and Soil Type." *Plant and Soil* 261 (1–2): 199–208. doi:10.1023/B:PLSO.0000035569.80747.c5.
- Martínez-Romero, Esperanza, José Luis Aguirre-Noyola, Nataly Taco-Taype, Julio Martínez-Romero, and Doris Zuñiga-Dávila. 2020. "Plant Microbiota Modified by Plant Domestication." *Systematic and Applied Microbiology* 43 (5). Elsevier GmbH.: 126106. doi:10.1016/j.syapm.2020.126106.
- Martínez Sánchez DS. Fac. Ciencias 2017. Análisis del microbioma de rizósfera visto a través de diferentes suelos y hospederos.
- Masella, Andre P., Andrea K. Bartram, Jakub M. Truszkowski, Daniel G. Brown, and Josh D. Neufeld. 2012. "PANDAseq: Paired-End Assembler for Illumina Sequences." *BMC Bioinformatics* 13 (1). BioMed Central Ltd: 31. doi:10.1186/1471-2105-13-31.
- McMurdie, Paul J., and Susan Holmes. 2013. "Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data." Edited by Michael Watson. *PLoS ONE* 8 (4): e61217. doi:10.1371/journal.pone.0061217.
- Mendes, Lucas W., Eiko E. Kuramae, Acácio A. Navarrete, Johannes A. Van Veen, and Siu M. Tsai. 2014. "Taxonomical and Functional Microbial Community Selection in Soybean Rhizosphere." *ISME Journal* 8 (8): 1577–87. doi:10.1038/ismej.2014.17.
- Menzel, Peter, Kim Lee Ng, and Anders Krogh. 2016. "Fast and Sensitive Taxonomic Classification for Metagenomics with Kaiju." *Nature Communications* 7 (1). Nature Publishing Group: 11257. doi:10.1038/ncomms11257.

- Micallef, Shirley A., Michael P. Shiaris, and Adán Colón-Carmona. 2009. "Influence of Arabidopsis Thaliana Accessions on Rhizobacterial Communities and Natural Variation in Root Exudates." *Journal of Experimental Botany* 60 (6): 1729–42. doi:10.1093/jxb/erp053.
- Mira, Alex, Ravindra Pushker, and Francisco Rodríguez-Valera. 2006. "The Neolithic Revolution of Bacterial Genomes." *Trends in Microbiology* 14 (5): 200–206. doi:10.1016/j.tim.2006.03.001.
- Murphy, J., and J. P. Riley. 1962. "A Modified Single Solution Method for the Determination of Phosphate in Natural Waters." *Analytica Chimica Acta* 27 (C): 31–36. doi:10.1016/S0003-2670(00)88444-5.
- Nawrocki, Eric P., Diana L. Kolbe, and Sean R. Eddy. 2009. "Infernal 1.0: Inference of RNA Alignments." *Bioinformatics* 25 (10): 1335–37. doi:10.1093/bioinformatics/btp157.
- Naylor, Dan, Stephanie Degraaf, Elizabeth Purdom, and Devin Coleman-Derr. 2017. "Drought and Host Selection Influence Bacterial Community Dynamics in the Grass Root Microbiome." *ISME Journal* 11 (12). Nature Publishing Group: 2691–2704. doi:10.1038/ismej.2017.118.
- Neilson, Julia W., Katy Califf, Cesar Cardona, Audrey Copeland, Will van Treuren, Karen L. Josephson, Rob Knight, et al. 2017. "Significant Impacts of Increasing Aridity on the Arid Soil Microbiome." Edited by Haiyan Chu. *mSystems* 2 (3): 1–15. doi:10.1128/msystems.00195-16.
- Noinaj, Nicholas, Maude Guillier, Travis J. Barnard, and Susan K. Buchanan. 2010. "TonB-Dependent Transporters: Regulation, Structure, and Function." *Annual Review of Microbiology* 64 (75): 43–60. doi:10.1146/annurev.micro.112408.134247.
- Nurk, Sergey, Dmitry Meleshko, Anton Korobeynikov, and Pavel A. Pevzner. 2017. "MetaSPAdes: A New Versatile Metagenomic Assembler." *Genome Research* 27 (5): 824–34. doi:10.1101/gr.213959.116.
- O'Reilly, C., and P. D. Turner. 2003. "The Nitrilase Family of CN Hydrolysing Enzymes - A Comparative Study." *Journal of Applied Microbiology* 95 (6): 1161–74. doi:10.1046/j.1365-2672.2003.02123.x.
- Overbeek, Ross, Robert Olson, Gordon D. Pusch, Gary J. Olsen, James J. Davis, Terry Disz, Robert A. Edwards, et al. 2014. "The SEED and the Rapid Annotation of Microbial Genomes Using Subsystems Technology (RAST)." *Nucleic Acids Research* 42 (D1): 206–14. doi:10.1093/nar/gkt1226.
- Pandey, C. B., Upendra Kumar, Megha Kaviraj, K. J. Minick, A. K. Mishra, and J. S. Singh. 2020. "DNRA: A Short-Circuit in Biological N-Cycling to Conserve Nitrogen in Terrestrial Ecosystems." *Science of the Total Environment* 738. Elsevier B.V.: 139710. doi:10.1016/j.scitotenv.2020.139710.
- Panke-Buisse, Kevin, Angela C. Poole, Julia K. Goodrich, Ruth E. Ley, and Jenny Kao-Kniffin. 2015. "Selection on Soil Microbiomes Reveals Reproducible Impacts on Plant Function." *ISME Journal* 9 (4). Nature Publishing Group: 980–89. doi:10.1038/ismej.2014.196.
- Pautot, V., F. M. Holzer, B. Reisch, and L. L. Walling. 1993. "Leucine Aminopeptidase: An Inducible Component of the Defense Response in *Lycopersicon Esculentum* (Tomato)." *Proceedings of the National Academy of Sciences* 90 (21): 9906–10. doi:10.1073/pnas.90.21.9906.
- Pautot, Véronique, Frances M. Holzer, Josette Chaufaux, and Linda L. Walling. 2001. "The Induction of Tomato Leucine Aminopeptidase Genes (LapA) After *Pseudomonas syringae* Pv. Tomato Infection Is Primarily a Wound Response Triggered by Coronatine." *Molecular Plant-Microbe Interactions* 14 (2): 214–24. doi:10.1094/MPMI.2001.14.2.214.
- Peiffer, Jason A., Aymé Spor, Omry Koren, Zhao Jin, Susannah Green Tringe, Jeffery L. Dangl, Edward S. Buckler, and Ruth E. Ley. 2013. "Diversity and Heritability of the Maize Rhizosphere Microbiome under Field Conditions." *Proceedings of the National Academy of Sciences of the United States of America* 110 (16): 6548–53. doi:10.1073/pnas.1302837110.

- Pérez-Jaramillo, Juan E., Víctor J. Carrión, Mattias de Hollander, and Jos M. Raaijmakers. 2018. "The Wild Side of Plant Microbiomes." *Microbiome* 6 (1). Microbiome: 143. doi:10.1186/s40168-018-0519-z.
- Pérez-Jaramillo, Juan E., Mattias De Hollander, Camilo A. Ramírez, Rodrigo Mendes, Jos M. Raaijmakers, and Víctor J. Carrión. 2019. "Deciphering Rhizosphere Microbiome Assembly of Wild and Modern Common Bean (*Phaseolus vulgaris*) in Native and Agricultural Soils from Colombia." *Microbiome* 7 (1). Microbiome: 1–16. doi:10.1186/s40168-019-0727-1.
- Pérez-Jaramillo, Juan E., Rodrigo Mendes, and Jos M. Raaijmakers. 2016. "Impact of Plant Domestication on Rhizosphere Microbiome Assembly and Functions." *Plant Molecular Biology* 90 (6). Springer Netherlands: 635–44. doi:10.1007/s11103-015-0337-7.
- Peters, N. K., J. W. Frost, and S. R. Long. 1986. "A Plant Flavone, Luteolin, Induces Expression of *Rhizobium meliloti* Nodulation Genes." *Science* 233 (4767): 977–80. doi:10.1126/science.3738520.
- Phillips, Rebecca M., David A. Six, Edward A. Dennis, and Partho Ghosh. 2003. "In Vivo Phospholipase Activity of the *Pseudomonas aeruginosa* Cytotoxin ExoU and Protection of Mammalian Cells with Phospholipase A 2 Inhibitors." *Journal of Biological Chemistry* 278 (42): 41326–32. doi:10.1074/jbc.M302472200.
- Pii, Youry, Luigimaria Borruso, Lorenzo Brusetti, Carmine Crecchio, Stefano Cesco, and Tanja Mimmo. 2016. "The Interaction between Iron Nutrition, Plant Species and Soil Type Shapes the Rhizosphere Microbiome." *Plant Physiology and Biochemistry* 99 (February). Elsevier Masson SAS: 39–48. doi:10.1016/j.plaphy.2015.12.002.
- Poindexter, J. S. 1981. "The Caulobacters: Ubiquitous Unusual Bacteria." *Microbiological Reviews*. doi:10.1128/membr.45.1.123-179.1981.
- Price, Morgan N., Paramvir S. Dehal, and Adam P. Arkin. 2009. "Fasttree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix." *Molecular Biology and Evolution* 26 (7): 1641–50. doi:10.1093/molbev/msp077.
- Pruitt, Kim D., Tatiana Tatusova, and Donna R. Maglott. 2007. "NCBI Reference Sequences (RefSeq): A Curated Non-Redundant Sequence Database of Genomes, Transcripts and Proteins." *Nucleic Acids Research* 35 (SUPPL. 1): D61–65. doi:10.1093/nar/gkl842.
- Purugganan, Michael D., and Dorian Q. Fuller. 2009a. "The Nature of Selection during Plant Domestication." *Nature* 457 (7231): 843–48. doi:10.1038/nature07895.
- Putz, Martina, Philipp Schleusner, Tobias Rütting, and Sara Hallin. 2018. "Relative Abundance of Denitrifying and DNRA Bacteria and Their Activity Determine Nitrogen Retention or Loss in Agricultural Soil." *Soil Biology and Biochemistry* 123 (February). Elsevier: 97–104. doi:10.1016/j.soilbio.2018.05.006.
- Ramin, Kelly I., and Steven D. Allison. 2019. "Bacterial Tradeoffs in Growth Rate and Extracellular Enzymes." *Frontiers in Microbiology* 10 (December): 1–10. doi:10.3389/fmicb.2019.02956.
- Redford, Amanda J., Robert M. Bowers, Rob Knight, Yan Linhart, and Noah Fierer. 2010. "The Ecology of the Phyllosphere: Geographic and Phylogenetic Variability in the Distribution of Bacteria on Tree Leaves." *Environmental Microbiology* 12 (11): 2885–93. doi:10.1111/j.1462-2920.2010.02258.x.
- Richter, Ralf, Mahdi Hejazi, Regine Kraft, Karl Ziegler, and Wolfgang Lockau. 1999. "Cyanophycinase, a Peptidase Degrading the Cyanobacterial Reserve Material Multi-L-Arginyl-Poly-L-Aspartic Acid (Cyanophycin)." *European Journal of Biochemistry* 263: 163–69.
- Roesch, Luiz F.W., Roberta R. Fulthorpe, Alberto Riva, George Casella, Alison K.M. Hadwin, Angela D. Kent, Samira H. Daroub, Flavio A.O. Camargo, William G. Farmerie, and Eric W. Triplett. 2007. "Pyrosequencing Enumerates and Contrasts Soil Microbial Diversity." *ISME Journal* 1 (4): 283–90. doi:10.1038/ismej.2007.53.
- Rovira, A. D. 1956. "A study Of The Development Of The Root Surface Microflora During The Initial Stages Of Plant Growth." *Journal of Applied Bacteriology* 19 (1): 72–79. doi:10.1111/j.1365-2672.1956.tb00048.x.

- Rukshana, Fatima, Clayton R. Butterly, Jeff A. Baldock, and Caixian Tang. 2011. "Model Organic Compounds Differ in Their Effects on pH Changes of Two Soils Differing in Initial pH." *Biology and Fertility of Soils* 47 (1): 51–62. doi:10.1007/s00374-010-0498-0.
- Sato, Hiromi. 2003. "The Mechanism of Action of the *Pseudomonas aeruginosa*-Encoded Type III Cytotoxin, ExoU." *The EMBO Journal* 22 (12): 2959–69. doi:10.1093/emboj/cdg290.
- Saxton, K. E., and W. J. Rawls. 2006. "Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions." *Soil Science Society of America Journal* 70 (5): 1569–78. doi:10.2136/sssaj2005.0117.
- Scheller, Henrik Vibe, and Peter Ulvskov. 2010. "Hemicelluloses." *Annual Review of Plant Biology* 61 (1): 263–89. doi:10.1146/annurev-arplant-042809-112315.
- Schlaeppi, Klaus, Nina Dombrowski, Ruben Garrido Oter, Emiel Ver Loren Van Themaat, and Paul Schulze-Lefert. 2014. "Quantitative Divergence of the Bacterial Root Microbiota in *Arabidopsis thaliana* Relatives." *Proceedings of the National Academy of Sciences of the United States of America* 111 (2): 585–92. doi:10.1073/pnas.1321597111.
- Scholz-Schroeder, B. K., M. L. Hutchison, I. Grgurina, and D. C. Gross. 2001. "The Contribution of Syringopeptin and Syringomycin to Virulence of *Pseudomonas syringae* Pv. *syringae* Strain B301D on the Basis of *sypA* and *syrB1* Biosynthesis Mutant Analysis." *Molecular Plant-Microbe Interactions* 14 (3): 336–48. doi:10.1094/MPMI.2001.14.3.336.
- Schreiter, Susanne, Guo Chun Ding, Holger Heuer, Günter Neumann, Martin Sandmann, Rita Grosch, Siegfried Kropf, and Kornelia Smalla. 2014. "Effect of the Soil Type on the Microbiome in the Rhizosphere of Field-Grown Lettuce." *Frontiers in Microbiology* 5 (APR): 144. doi:10.3389/fmicb.2014.00144.
- Shange, Raymon S., Ramble O. Ankumah, Abasiofiok M. Ibekwe, Robert Zabawa, and Scot E. Dowd. 2012. "Distinct Soil Bacterial Communities Revealed under a Diversely Managed Agroecosystem." *PLoS ONE* 7 (7). doi:10.1371/journal.pone.0040338.
- Shu, Yue Zhong, Maredith Arcuri, Michael Kozlowski, Richard Wang, Kin Sing Lam, Li Ping Chang, Dolores Pirnik, and Patricia Rose. 1994. "Haloemodins, a New Class of Endothelin-1 Type B (ETB) Receptor Binding Inhibitors." *Journal of Antibiotics* 47 (11): 1328–32. doi:10.7164/antibiotics.47.1328.
- Simpson, E. H. 1949. "Measurement of Diversity." *Nature* 163 (4148): 688–688. doi:10.1038/163688a0.
- Singer, Esther, Jason Bonnette, Tanja Woyke, and Thomas E. Juenger. 2019. "Conservation of Endophyte Bacterial Community Structure Across Two *Panicum* Grass Species." *Frontiers in Microbiology* 10 (SEP). doi:10.3389/fmicb.2019.02181.
- Skerker, J. M., and Lucy Shapiro. 2000. "Identification and Cell Cycle Control of a Novel Pilus System in *Caulobacter crescentus*." *The EMBO Journal* 19 (13): 3223–34. doi:10.1093/emboj/19.13.3223.
- Somasundaram, Sutharsan, Satoshi Fukuzono, and Morio Iijima. "Dynamics of root border cells in rhizosphere soil of *Zea mays* L.: crushed cells during root penetration, survival in soil, and long term soil compaction effect." *Plant production science* 11, no. 4 (2008): 440-446.
- Sprent, Janet I. 2008. "60Ma of Legume Nodulation. What's New? What's Changing?" *Journal of Experimental Botany* 59 (5): 1081–84. doi:10.1093/jxb/erm286.
- Sreevidya, M., S. Gopalakrishnan, H. Kudapa, and R. K. Varshney. 2016. "Exploring Plant Growth-Promotion Actinomycetes from Vermicompost and Rhizosphere Soil for Yield Enhancement in Chickpea." *Brazilian Journal of Microbiology* 47 (1). Sociedade Brasileira de Microbiologia: 85–95. doi:10.1016/j.bjm.2015.11.030.
- SSSA. Glossary of Soil Science Terms. Disponible en: www.soils.org/sssagloss/index.php.
- Stackebrandt, E. & Goebel, B.M., 1994. 'Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology.' *International journal of systematic bacteriology*. 44(4), pp.846–849.
- Sumner, M.E., 2000. Handbook of Soil Science. CRC Press, Boca Raton.

- Tatusov, R L, M Y Galperin, D a Natale, and E V Koonin. 2000. "The COG Database: A Tool for Genome-Scale Analysis of Protein Functions and Evolution." *Nucleic Acids Research* 28 (1): 33–36. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=102395&tool=pmcentrez&rendertype=abstract>.
- Thompson, Luke R., Jon G. Sanders, Daniel McDonald, Amnon Amir, Joshua Ladau, Kenneth J. Locey, Robert J. Prill, et al. 2017. "A Communal Catalogue Reveals Earth's Multiscale Microbial Diversity." *Nature* 551 (7681): 457–63. doi:10.1038/nature24621.
- Tian, Bao Yu, Yi Cao, and Ke Qin Zhang. 2015. "Metagenomic Insights into Communities, Functions of Endophytes, and Their Associates with Infection by Root-Knot Nematode, *Meloidogyne incognita*, in Tomato Roots." *Scientific Reports* 5 (1). Nature Publishing Group: 17087. doi:10.1038/srep17087.
- Toju, Hirokazu, Kabir G. Peay, Masato Yamamichi, Kazuhiko Narisawa, Kei Hiruma, Ken Naito, Shinji Fukuda, et al. 2018. "Core Microbiomes for Sustainable Agroecosystems." *Nature Plants* 4 (5). Springer US: 247–57. doi:10.1038/s41477-018-0139-4.
- Truyens, Sascha, Nele Weyens, Ann Cuypers, and Jaco Vangronsveld. 2015. "Bacterial Seed Endophytes: Genera, Vertical Transmission and Interaction with Plants." *Environmental Microbiology Reports* 7 (1): 40–50. doi:10.1111/1758-2229.12181.
- Upchurch, Rima, Chih Yu Chiu, Karin Everett, Glen Dyszynski, David C. Coleman, and William B. Whitman. 2008. "Differences in the Composition and Diversity of Bacterial Communities from Agricultural and Forest Soils." *Soil Biology and Biochemistry* 40 (6): 1294–1305. doi:10.1016/j.soilbio.2007.06.027.
- van Bruggen, A. H.C., I. M. Francis, and K. N. Jochimsen. 2014. "Non-Pathogenic Rhizosphere Bacteria Belonging to the Genera *Rhizorhapis* and *Sphingobium* Provide Specific Control of Lettuce Corky Root Disease Caused by Species of the Same Bacterial Genera." *Plant Pathology* 63 (6): 1384–94. doi:10.1111/ppa.12212.
- Van Der Heijden, Marcel G.A., Richard D. Bardgett, and Nico M. Van Straalen. 2008. "The Unseen Majority: Soil Microbes as Drivers of Plant Diversity and Productivity in Terrestrial Ecosystems." *Ecology Letters* 11 (3): 296–310. doi:10.1111/j.1461-0248.2007.01139.x.
- Vande Broek, Ann, Mark Lambrecht, and Jos Vanderleyden. 1998. "Bacterial Chemotactic Motility Is Important for the Initiation of Wheat Root Colonization by *Azospirillum brasilense*." *Microbiology* 144 (9): 2599–2606. doi:10.1099/00221287-144-9-2599.
- Vermeer, J., and M. E. McCully. 1982. "The Rhizosphere in Zea: New Insight into Its Structure and Development." *Planta* 156 (1): 45–61. doi:10.1007/BF00393442.
- Vieira, Selma, Johannes Sikorski, Sophie Dietz, Katharina Herz, Marion Schruppf, Helge Bruelheide, Dierk Scheel, Michael W. Friedrich, and Jörg Overmann. 2020. "Drivers of the Composition of Active Rhizosphere Bacterial Communities in Temperate Grasslands." *ISME Journal* 14 (2). Springer US: 463–75. doi:10.1038/s41396-019-0543-4.
- Watrud, Lidia S., Jay R. Reichman, Michael A. Bollman, Bonnie M. Smith, E. Henry Lee, Julie D. Jastrow, Michael D. Casler, et al. 2013. "Chemistry and Microbial Functional Diversity Differences in Biofuel Crop and Grassland Soils in Multiple Geographies." *BioEnergy Research* 6 (2): 601–19. doi:10.1007/s12155-012-9279-y.
- Watzer, Björn, and Karl Forchhammer. 2018. "Cyanophycin Synthesis Optimizes Nitrogen Utilization in the Unicellular Cyanobacterium *Synechocystis* Sp. Strain PCC 6803." Edited by Claire Vieille. *Applied and Environmental Microbiology* 84 (20): 1–18. doi:10.1128/AEM.01298-18.
- Whittaker, R H. 1972. "Evolution And Measurement Of Species Diversity." *TAXON* 21 (2–3): 213–51. doi:10.2307/1218190
- Wilke, Andreas, Travis Harrison, Jared Wilkening, Dawn Field, Elizabeth M. Glass, Nikos Kyrpides, Konstantinos Mavrommatis, and Folker Meyer. 2012. "The M5nr: A Novel Non-Redundant Database Containing Protein Sequences and Annotations from Multiple Sources and Associated Tools." *BMC Bioinformatics* 13 (1): 141. doi:10.1186/1471-2105-13-141.

- Wilkinson, Leland. 2011. "ggplot2: Elegant Graphics for Data Analysis by WICKHAM, H." *Biometrics* 67 (2): 678–79. doi:10.1111/j.1541-0420.2011.01616.x.
- Wissuwa, Matthias, Mark Mazzola, and Christine Picard. 2009. "Novel Approaches in Plant Breeding for Rhizosphere-Related Traits." *Plant and Soil*. doi:10.1007/s11104-008-9693-2.
- Wu, Huili, Wenhua Xiang, Shuai Ouyang, David I. Forrester, Bo Zhou, Lingxiu Chen, Tida Ge, et al. 2019. "Linkage between Tree Species Richness and Soil Microbial Diversity Improves Phosphorus Bioavailability." Edited by Rachel Gallery. *Functional Ecology* 33 (8): 1549–60. doi:10.1111/1365-2435.13355.
- Wu, Linkun, Juanying Wang, Hongmiao Wu, Jun Chen, Zhigang Xiao, Xianjin Qin, Zhongyi Zhang, and Wenxiong Lin. 2018. "Comparative Metagenomic Analysis of Rhizosphere Microbial Community Composition and Functional Potentials under *Rehmannia glutinosa* Consecutive Monoculture." *International Journal of Molecular Sciences* 19 (8): 1–17. doi:10.3390/ijms19082394.
- Xu, Jin, Yunzeng Zhang, Pengfan Zhang, Pankaj Trivedi, Nadia Riera, Yayu Wang, Xin Liu, et al. 2018. "The Structure and Function of the Global Citrus Rhizosphere Microbiome." *Nature Communications* 9 (1): 4894. doi:10.1038/s41467-018-07343-2.
- Yang, Zhenping, Wenping Yang, Shengcai Li, Jiaomin Hao, Zhifeng Su, Min Sun, Zhiqiang Gao, and Chunlai Zhang. 2016. "Variation of Bacterial Community Diversity in Rhizosphere Soil of Sole-Cropped versus Intercropped Wheat Field after Harvest." *PLoS ONE* 11 (3): 1–18. doi:10.1371/journal.pone.0150618.
- Young, C. C., A. B. Arun, Peter Kämpfer, H. J. Busse, W. A. Lai, W. M. Chen, F. T. Shen, and P. D. Rekha. 2008. "*Sphingobium rhizovicinum* Sp. Nov., Isolated from Rhizosphere Soil of *Fortunella Hindsii* (Champ. Ex Benth.) Swingle." *International Journal of Systematic and Evolutionary Microbiology* 58 (8): 1801–6. doi:10.1099/ijs.0.65564-0.
- Yu, Ke, Corné M.J. Pieterse, Peter A.H.M. Bakker, and Roeland L. Berendsen. 2019. "Beneficial Microbes Going Underground of Root Immunity." *Plant Cell and Environment* 42 (10): 2860–70. doi:10.1111/pce.13632.
- Yurgel, Svetlana N., Jacob T. Nearing, Gavin M. Douglas, and Morgan G. I. Langille. 2019. "Metagenomic Functional Shifts to Plant Induced Environmental Changes." *Frontiers in Microbiology* 10 (July): 1–12. doi:10.3389/fmicb.2019.01682.
- Zerbino, Daniel R., and Ewan Birney. 2008. "Velvet: Algorithms for de Novo Short Read Assembly Using de Bruijn Graphs." *Genome Research* 18 (5): 821–29. doi:10.1101/gr.074492.107.
- Zhalnina, Kateryna, Raquel Dias, Patricia Dörr de Quadros, Austin Davis-Richardson, Flavio A. O. Camargo, Ian M. Clark, Steve P. McGrath, Penny R. Hirsch, and Eric W. Triplett. 2015a. "Soil pH Determines Microbial Diversity and Composition in the Park Grass Experiment." *Microbial Ecology* 69 (2): 395–406. doi:10.1007/s00248-014-0530-2.
- Zhalnina, Kateryna, Raquel Dias, Patricia Dörr de Quadros, Austin Davis-Richardson, Flavio A O Camargo, Ian M Clark, Steve P McGrath, Penny R Hirsch, and Eric W Triplett. 2015b. "Soil pH Determines Microbial Diversity and Composition in the Park Grass Experiment." *Microbial Ecology* 69 (2): 395–406. doi:10.1007/s00248-014-0530-2.
- Zhang, X F, L. Zhao, S J Xu, Y Z Liu, H Y Liu, and G D Cheng. 2013. "Soil Moisture Effect on Bacterial and Fungal Community in Beilu River (Tibetan Plateau) Permafrost Soils with Different Vegetation Types." *Journal of Applied Microbiology* 114 (4): 1054–65. doi:10.1111/jam.12106.
- Zobel, Martin, Eddy van der Maarel, and Cecilia Dupré. 1998. "Species Pool: The Concept, Its Determination and Significance for Community Restoration." *Applied Vegetation Science* 1 (1): 55–66. doi:10.2307/1479085.

Anexo I

Producción de artículos científicos durante el doctorado.

1. **Barajas, Hugo R.**, Shamayim Martinez-Sanchez, Miguel F. Romero, Cristóbal Hernandez-Alvarez, Luis Servin-Gonzalez, Mariana Peimbert & Luis D. Alcaraz. "Testing the two-step model of plant root microbiome acquisition under multiple plant species and soil sources." *Frontiers in Microbiology* (2020): 2445. <https://doi.org/10.3389/fmicb.2020.542742>
2. Romero Miguel, Diana Gallego, Jazmin Blaz, Arely Lechuga, José F. Martínez, **Hugo R. Barajas**, Corina Hayano-Kanashiro et al. "Rhizosphere metagenomics of mine tailings colonizing plants: assembling and selecting synthetic bacterial communities to enhance in situ bioremediation." *bioRxiv* (2019): 664805.
3. **Barajas, Hugo R.**, Miguel F. Romero, Shamayim Martínez-Sánchez, and Luis D. Alcaraz. "Global genomic similarity and core genome sequence diversity of the *Streptococcus* genus as a toolkit to identify closely related bacterial species in complex environments." *PeerJ* 6 (2019): e6233. <https://doi.org/10.7717/peerj.6233>
4. Alcaraz, Luis D., Mariana Peimbert, **Hugo R. Barajas**, Ana E. Dorantes-Acosta, John L. Bowman, and Mario A. Arteaga-Vázquez. "Marchantia liverworts as a proxy to plants' basal microbiomes." *Scientific reports* 8, no. 1 (2018): 1-12. <https://doi.org/10.1038/s41598-018-31168-0>
5. Martínez-Carranza, Enrique, **Hugo Barajas**, Luis-David Alcaraz, Luis Servín-González, Gabriel-Yaxal Ponce-Soto, and Gloria Soberón-Chávez. "Variability of bacterial essential genes among closely related bacteria: The case of *Escherichia coli*." *Frontiers in microbiology* 9 (2018): 1059. <https://doi.org/10.3389/fmicb.2018.01059>



Testing the Two-Step Model of Plant Root Microbiome Acquisition Under Multiple Plant Species and Soil Sources

Hugo R. Barajas¹, Shamayim Martínez-Sánchez¹, Miguel F. Romero¹, Cristóbal Hernández Álvarez¹, Luis Servín-González², Mariana Peimbert³, Rocío Cruz-Ortega⁴, Felipe García-Oliva⁵ and Luis D. Alcaraz^{1*}

¹ Departamento de Biología Celular, Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City, Mexico, ² Departamento de Biología Molecular y Biotecnología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Mexico City, Mexico, ³ Departamento de Ciencias Naturales, Unidad Cuajimalpa, Universidad Autónoma Metropolitana, Mexico City, Mexico, ⁴ Laboratorio de Alelopatía, Departamento de Ecología Funcional, Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City, Mexico, ⁵ Instituto de Investigaciones en Ecosistemas y Sustentabilidad, Universidad Nacional Autónoma de México, Morelia, Mexico

OPEN ACCESS

Edited by:

Alok Kumar Srivastava,
National Bureau of Agriculturally
Important Microorganisms (ICAR),
India

Reviewed by:

Murali Gopal,
Central Plantation Crops Research
Institute (ICAR), India
Siu Mui Tsai,
University of São Paulo, Brazil

*Correspondence:

Luis D. Alcaraz
lalcaraz@ciencias.unam.mx

Specialty section:

This article was submitted to
Microbial Symbioses,
a section of the journal
Frontiers in Microbiology

Received: 13 March 2020

Accepted: 11 September 2020

Published: 09 October 2020

Citation:

Barajas HR, Martínez-Sánchez S,
Romero MF, Álvarez CH,
Servín-González L, Peimbert M,
Cruz-Ortega R, García-Oliva F and
Alcaraz LD (2020) Testing
the Two-Step Model of Plant Root
Microbiome Acquisition Under
Multiple Plant Species and
Soil Sources.
Front. Microbiol. 11:542742.
doi: 10.3389/fmicb.2020.542742

The two-step model for plant root microbiomes considers soil as the primary microbial source. Active selection of the plant's bacterial inhabitants results in a biodiversity decrease toward roots. We collected sixteen samples of *in situ* ruderal plant roots and their soils and used these soils as the main microbial input for single genotype tomatoes grown in a greenhouse. Our main goal was to test the soil influence in the structuring of rhizosphere microbiomes, minimizing environmental variability, while testing multiple plant species. We massively sequenced the 16S rRNA and shotgun metagenomes of the soils, *in situ* plants, and tomato roots. We identified a total of 271,940 bacterial operational taxonomic units (OTUs) within the soils, rhizosphere and endospheric microbiomes. We annotated by homology a total of 411,432 (13.07%) of the metagenome predicted proteins. Tomato roots did follow the two-step model with lower α -diversity than soil, while ruderal plants did not. Surprisingly, ruderal plants are probably working as a microenvironmental oasis providing moisture and plant-derived nutrients, supporting larger α -diversity. Ruderal plants and their soils are closer according to their microbiome community composition than tomato and its soil, based on OTUs and protein comparisons. We expected that tomato β -diversity clustered together with their soil, if it is the main rhizosphere microbiome structuring factor. However, tomato microbiome β -diversity was associated with plant genotype in most samples (81.2%), also supported by a larger set of enriched proteins in tomato rhizosphere than soil or ruderals. The most abundant bacteria found in soils was the Actinobacteria *Solirubrobacter soli*, ruderals were dominated by the Proteobacteria *Sphingomonas* sp. URGHD0057, and tomato mainly by the Bacteroidetes *Ohtaekwangia koreensis*, *Flavobacterium terrae*, *Niastella vici*, and *Chryseolinea serpens*. We calculated a metagenomic tomato root core of 51 bacterial genera and 2,762 proteins, which could

Disponible en: <https://www.biorxiv.org/content/10.1101/664805v1>

bioRxiv preprint doi: <https://doi.org/10.1101/664805>; this version posted June 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Rhizosphere metagenomics of mine tailings colonizing plants: assembling and selecting synthetic bacterial communities to enhance *in situ* bioremediation

Miguel Romero^{1,2,3*}, Diana Gallego^{1,2*}, Jazmin Blaz^{1,2}, Arely Lechuga^{1,2}, José F. Martínez⁴, Hugo R. Barajas^{1,2}, Corina Hayano-Kanashiro⁵, Mariana Peimbert⁶, Rocío Cruz-Ortega⁷, Francisco E. Molina-Freaner³, and Luis D. Alcaraz^{1,2*}

¹ Departamento de Biología Celular, Facultad de Ciencias, Universidad Nacional Autónoma de México. 04510 Coyoacán, Ciudad de México, México.

² Laboratorio Nacional de Ciencias de la Sostenibilidad (LANCIS), Instituto de Ecología, Universidad Nacional Autónoma de México. 04510 Coyoacán, Ciudad de México, México.

³ Posgrado en Ciencias Biológicas, Unidad de Posgrado, Edificio 1er piso, Ciudad Universitaria. Universidad Nacional Autónoma de México. 04510. Ciudad de México.

⁴ Departamento de Ecología de la Biodiversidad, Instituto de Ecología, Universidad Nacional Autónoma de México. 04510 Coyoacán, Ciudad de México, México.

⁵ Departamento de Investigaciones Científicas y Tecnológicas, Universidad de Sonora, Hermosillo, México.

⁶ Departamento de Ciencias Naturales, Universidad Autónoma Metropolitana, Unidad Cuajimalpa. 05348, Cuajimalpa, Ciudad de México, México.

⁷ Departamento de Ecología Funcional, Instituto de Ecología, Universidad Nacional Autónoma de México. 04510 Coyoacán, Ciudad de México, México.

*Contributed equally to this work

*Correspondence:

Luis D. Alcaraz

lalcaraz@ciencias.unam.mx

Keywords: Metagenomics, rhizosphere metagenome, synthetic bacterial community, mine tailing, phytostabilization



Global genomic similarity and core genome sequence diversity of the *Streptococcus* genus as a toolkit to identify closely related bacterial species in complex environments

Hugo R. Barajas¹, Miguel F. Romero¹, Shamayim Martínez-Sánchez¹ and Luis D. Alcaraz^{1,2}

¹ Departamento de Biología Celular, Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City, Mexico

² Laboratorio Nacional de Ciencias de la Sostenibilidad, Instituto de Ecología. Universidad Nacional Autónoma de México, Mexico city, Mexico

ABSTRACT

Background: The *Streptococcus* genus is relevant to both public health and food safety because of its ability to cause pathogenic infections. It is well-represented (>100 genomes) in publicly available databases. Streptococci are ubiquitous, with multiple sources of isolation, from human pathogens to dairy products. The *Streptococcus* genus has traditionally been classified by morphology, serum types, the 16S ribosomal RNA (rRNA) gene, and multi-locus sequence types subject to in-depth comparative genomic analysis.

Methods: Core and pan-genomes described the genomic diversity of 108 strains belonging to 16 *Streptococcus* species. The core genome nucleotide diversity was calculated and compared to phylogenomic distances within the genus *Streptococcus*. The core genome was also used as a resource to recruit metagenomic fragment reads from streptococci dominated environments. A conventional 16S rRNA gene phylogeny reconstruction was used as a reference to compare the resulting dendrograms of average nucleotide identity (ANI) and genome similarity score (GSS) dendrograms.

Results: The core genome, in this work, consists of 404 proteins that are shared by all 108 *Streptococcus*. The average identity of the pairwise compared core proteins decreases proportionally to GSS lower scores, across species. The GSS dendrogram recovers most of the clades in the 16S rRNA gene phylogeny while distinguishing between 16S polytomies (unresolved nodes). The GSS is a distance metric that can reflect evolutionary history comparing orthologous proteins. Additionally, GSS resulted in the most useful metric for genus and species comparisons, where ANI metrics failed due to false positives when comparing different species.

Discussion: Understanding of genomic variability and species relatedness is the goal of tools like GSS, which makes use of the maximum pairwise shared orthologous sequences for its calculation. It allows for long evolutionary distances (above species) to be included because of the use of amino acid alignment scores, rather than nucleotides, and normalizing by positive matches. Newly sequenced species and strains could be easily placed into GSS dendrograms to infer overall genomic

Submitted 9 March 2018
Accepted 7 December 2018
Published 14 January 2019

Corresponding author
Luis D. Alcaraz,
lalcaraz@ciencias.unam.mx

Academic editor
Kay Nieselt

Additional Information and
Declarations can be found on
page 12

DOI 10.7717/peerj.6233

© Copyright
2019 Barajas et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

How to cite this article Barajas HR, Romero MF, Martínez-Sánchez S, Alcaraz LD. 2019. Global genomic similarity and core genome sequence diversity of the *Streptococcus* genus as a toolkit to identify closely related bacterial species in complex environments. *PeerJ* 6:e6233 DOI 10.7717/peerj.6233

OPEN *Marchantia* liverworts as a proxy to plants' basal microbiomes

Luis D. Alcaraz¹, Mariana Peimbert², Hugo R. Barajas¹, Ana E. Dorantes-Acosta³, John L. Bowman⁴ & Mario A. Arteaga-Vázquez³

Received: 21 April 2017
Accepted: 13 August 2018
Published online: 23 August 2018

Microbiomes influence plant establishment, development, nutrient acquisition, pathogen defense, and health. Plant microbiomes are shaped by interactions between the microbes and a selection process of host plants that distinguishes between pathogens, commensals, symbionts and transient bacteria. In this work, we explore the microbiomes through massive sequencing of the 16S rRNA genes of microbiomes two *Marchantia* species of liverworts. We compared microbiomes from *M. polymorpha* and *M. paleacea* plants collected in the wild relative to their soils substrates and from plants grown *in vitro* that were established from gemmae obtained from the same populations of wild plants. Our experimental setup allowed identification of microbes found in both native and *in vitro* *Marchantia* species. The main OTUs (97% identity) in *Marchantia* microbiomes were assigned to the following genera: *Methylobacterium*, *Rhizobium*, *Paenibacillus*, *Lysobacter*, *Pirellula*, *Steroidobacter*, and *Bryobacter*. The assigned genera correspond to bacteria capable of plant-growth promotion, complex exudate degradation, nitrogen fixation, methylo-trophs, and disease-suppressive bacteria, all hosted in the relatively simple anatomy of the plant. Based on their long evolutionary history *Marchantia* is a promising model to study not only long-term relationships between plants and their microbes but also the transgenerational contribution of microbiomes to plant development and their response to environmental changes.

All multicellular eukaryotes have a microbiome composed of prokaryotes, primarily bacteria, and eukaryotes, both uni- and multicellular. Species of the microbiome can be commensal, pathogenic, symbiotic, essential, or neutral. Land plants are no exception, with three distinct habitats colonized by microbial species — the rhizosphere underground and at the substrate surface, the phyllosphere on the aerial plant surfaces, and the endosphere within the plant body^{1–3}. Land plants evolved from a freshwater algal ancestor and upon the transition to terrestrial habitats was exposed to novel microbial communities. It has long been recognized that early, and perhaps the first, land plants harbored mycorrhizal fungi, perhaps in a symbiotic relationship from the onset⁴. Furthermore, it has been suggested that early land plants may have inherited a microbiome from their algal ancestor, perhaps consisting of nitrogen-fixing bacterial and methanotrophs⁵.

Extant land plants consist of bryophytes (liverworts, mosses, hornworts), a monophyletic or paraphyletic group of gametophyte dominant plants and the monophyletic vascular plants. Bryophytes are largely *poikilohydric* — their water potential is equilibrated rapidly to external water availability. Their gametophyte bodies are in close contact with the substrate surface and ventral rhizoids act in water absorption and distribution. In contrast, the diploid sporophyte dominant vascular plants are homoiohydric, with a root system through which water is obtained. Early land plants established symbiotic associations with mycorrhizae, and it is thought that Palaeozoic drops in CO₂ reduced phosphate (P) absorption in non-vascular plants which in addition to competition for light favoured the vascular plants Earth dominance⁶.

Extensive sequencing approaches have fostered recent studies on plant-microbe interactions, and plant microbiomes from model plants like *Arabidopsis thaliana* and agriculturally relevant species are publicly available^{7–10}, and these have contributed to the study of plant-microbe interactions. Proposed models for the establishment of the plant root microbiome include the two-step model that considers the root microbiome as a product of

¹Departamento de Biología Celular, Facultad de Ciencias, Universidad Nacional Autónoma de México, Ciudad Universitaria, UNAM, 04510, Coyoacán, Mexico City, Mexico. ²Departamento de Ciencias Naturales, Universidad Autónoma Metropolitana, Unidad Cuajimalpa, Av. Vasco de Quiroga 4871, Col. Santa Fe Cuajimalpa, 05348, Mexico City, Mexico. ³University of Veracruz, Institute for Biotechnology and Applied Ecology (INBIOTECA), Avenida de las Culturas Veracruzanas 101, Colonia Emiliano Zapata, 91090, Xalapa, Veracruz, Mexico. ⁴School of Biological Sciences, Monash University, Melbourne, Victoria, 3800, Australia. Correspondence and requests for materials should be addressed to L.D.A. (email: lalcaraz@ciencias.unam.mx) or M.A.A.-V. (email: maarteaga@uv.mx)



Variability of Bacterial Essential Genes Among Closely Related Bacteria: The Case of *Escherichia coli*

Enrique Martínez-Carranza¹, Hugo Barajas², Luis-David Alcaraz^{2,3}, Luis Servín-González¹, Gabriel-Yaxal Ponce-Soto⁴ and Gloria Soberón-Chávez^{1*}

¹ Departamento de Biología Molecular y Biotecnología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Mexico City, Mexico, ² Departamento de Biología Celular, Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City, Mexico, ³ Laboratorio Nacional de Ciencias de la Sostenibilidad, Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City, Mexico, ⁴ Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City, Mexico

OPEN ACCESS

Edited by:

Vasco Ariston De Carvalho Azevedo,
Universidade Federal de Minas
Gerais, Brazil

Reviewed by:

Bei-Wen Ying,
University of Tsukuba, Japan
Jun Li,
Zhejiang University of Technology,
China

*Correspondence:

Gloria Soberón-Chávez
gloria@biomedicas.unam.mx

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 22 February 2018

Accepted: 04 May 2018

Published: 29 May 2018

Citation:

Martínez-Carranza E, Barajas H,
Alcaraz L-D, Servín-González L,
Ponce-Soto G-Y and
Soberón-Chávez G (2018) Variability
of Bacterial Essential Genes Among
Closely Related Bacteria: The Case
of *Escherichia coli*.
Front. Microbiol. 9:1059.
doi: 10.3389/fmicb.2018.01059

The definition of bacterial essential genes has been widely pursued using different approaches. Their study has impacted several fields of research such as synthetic biology, the construction of bacteria with minimal chromosomes, the search for new antibiotic targets, or the design of strains with biotechnological applications. Bacterial genomes are mosaics that only share a small subset of gene-sequences (core genome) even among members of the same species. It has been reported that the presence of essential genes is highly variable between closely related bacteria and even among members of the same species, due to the phenomenon known as "non-orthologous gene displacement" that refers to the coding for an essential function by genes with no sequence homology due to horizontal gene transfer (HGT). The existence of dormant forms among bacteria and the high incidence of HGT have been proposed to be driving forces of bacterial evolution, and they might have a role in the low level of conservation of essential genes among related bacteria by non-orthologous gene displacement, but this correlation has not been recognized. The aim of this mini-review is to give a brief overview of the approaches that have been taken to define and study essential genes, and the implications of non-orthologous gene displacement in bacterial evolution, focusing mainly in the case of *Escherichia coli*. To this end, we reviewed the available literature, and we searched for the presence of the essential genes defined by mutagenesis in the genomes of the 63 best-sequenced *E. coli* genomes that are available in NCBI database. We could not document specific cases of non-orthologous gene displacement among the *E. coli* strains analyzed, but we found that the quality of the genome-sequences in the database is not enough to make accurate predictions about the conservation of essential-genes among members of this bacterial species.

Keywords: essential-genes, core-genome, non-orthologous gene displacement, *Escherichia coli*, whole genome sequence databases