



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**Maestría y Doctorado en Ciencias Bioquímicas**

**“Metagenómica de rizósferas en distintos tipos de suelo de México”**

**TESIS**

**QUE PARA OPTAR POR EL GRADO DE:**

Maestro en Ciencias

**PRESENTA:**

Hugo Rafael Barajas de la Torre

**TUTOR PRINCIPAL**

Dr. Luis David Alcaraz Peraza

Instituto de Ecología, UNAM

**MIEMBROS DEL COMITÉ TUTOR**

Dr. Luis Servín González

Instituto de Investigaciones Biomédicas, UNAM

Dr. Felipe García Oliva

Instituto de Investigaciones en Ecosistemas y Sustentabilidad, UNAM

**Ciudad de México, abril de 2016**



Universidad Nacional  
Autónoma de México



**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.

## **Metagenómica de rizósferas en distintos tipos de suelo de México**

### **Reconocimientos**

Esta tesis de maestría se realizó bajo la tutoría del Dr. Luis David Alcaraz Peraza en el Laboratorio Nacional de Ciencias de la Sostenibilidad del Instituto de Ecología de la Universidad Nacional Autónoma de México y el análisis fisicoquímico de los suelos y plantas se llevó a cabo bajo la asesoría del Dr. Felipe García Oliva en el Laboratorio de Biogeoquímica de Suelos del Instituto de Investigación en Ecosistemas y Sustentabilidad de la Universidad Nacional Autónoma de México.

Se reconoce el apoyo técnico de: Dr. Felipe García Oliva y M. en C. Rodrigo Velázquez Durán en el área de análisis fisicoquímicos de suelos; Dra. Rocío Cruz Ortega por facilitar las semillas de tomate, espacio de invernadero y equipo; Biol. Shamayim Martínez Sánchez y Francisco González en el muestreo; Biol. Jazmín Blaz y Q.F.B. Selene Molina por la ayuda en el procesamiento de muestras.

Este proyecto fue apoyado financieramente por los proyectos DGAPA-PAPIIT-UNAM IA200514 y CONACyT 237387. Durante mis estudios de maestría he sido beneficiado con una beca otorgada por el CONACyT.

Miembros del Comité Tutor:

Dr. Luis Servín González Instituto de Investigaciones Biomédicas, UNAM

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

## Índice de Contenido

<a href="#">Resumen</a>	<a href="#">3</a>
<a href="#">Introducción</a>	<a href="#">4</a>
<a href="#">Diversidad taxonómica y funcional en el microbiana en el suelo</a>	<a href="#">4</a>
<a href="#">Factores implicados en la modificación de la estructura de las comunidades microbianas</a>	<a href="#">5</a>
<a href="#">El microbioma de las plantas y las estrategias para su análisis.</a>	<a href="#">6</a>
<a href="#">El suelo, un ambiente heterogéneo</a>	<a href="#">9</a>
<a href="#">Los principales suelos de México</a>	<a href="#">10</a>
<a href="#">Antecedentes</a>	<a href="#">11</a>
<a href="#">Objetivo general</a>	<a href="#">16</a>
<a href="#">Objetivos particulares</a>	<a href="#">16</a>
<a href="#">Hipótesis</a>	<a href="#">16</a>
<a href="#">Metodología</a>	<a href="#">17</a>
<a href="#">Esquema general del trabajo</a>	<a href="#">17</a>
<a href="#">Colecta de distintos tipos de suelo</a>	<a href="#">18</a>
<a href="#">Análisis físico-químico de los suelos.</a>	<a href="#">19</a>
<a href="#">Germinación de plantas</a>	<a href="#">20</a>
<a href="#">Crecimiento de plantas en distintos tipos de suelo</a>	<a href="#">21</a>
<a href="#">Toma de muestras de suelo final, rizósfera y compartimento endofítico</a>	<a href="#">22</a>
<a href="#">Extracción de DNA de suelos.</a>	<a href="#">24</a>
<a href="#">Extracción de DNA de rizósfera y compartimento endofítico.</a>	<a href="#">24</a>
<a href="#">Amplificación del gen ribosomal 16S rRNA.</a>	<a href="#">25</a>
<a href="#">Secuenciación masiva.</a>	<a href="#">25</a>
<a href="#">Procesamiento y análisis de datos de secuenciación.</a>	<a href="#">27</a>
<a href="#">Resultados y discusión</a>	<a href="#">31</a>
<a href="#">Clasificación y características ambientales de los tipos de suelo colectados en este trabajo.</a>	<a href="#">31</a>
<a href="#">Crecimiento y biomasa de las plantas utilizadas en el experimento de jardín común con los múltiples suelos colectados.</a>	<a href="#">33</a>
<a href="#">Análisis de fisicoquímico de los suelos.</a>	<a href="#">35</a>
<a href="#">Construcción de bibliotecas del gen rRNA 16S</a>	<a href="#">42</a>
<a href="#">Control de calidad de las secuencias del gen rRNA 16S</a>	<a href="#">42</a>
<a href="#">Perfiles filogenéticos de los compartimentos de raíz en JAL5.</a>	<a href="#">46</a>
<a href="#">Predicción metabólica en los compartimentos de raíz de JAL5.</a>	<a href="#">51</a>
<a href="#">Conclusiones</a>	<a href="#">55</a>
<a href="#">Perspectivas</a>	<a href="#">56</a>
<a href="#">Bibliografía</a>	<a href="#">58</a>

## Resumen

El suelo es un ambiente heterogéneo en donde los microorganismos presentes en ellos se encuentran sujetos a distintos factores bióticos y abióticos. Estos factores estructuran a las comunidades bacterianas presentes en el suelo, las cuales son elementales en el mantenimiento de la funcionalidad del mismo al estar involucradas en el ciclo de nutrientes.

La rizósfera, se define como la capa de suelo estrechamente unida a las raíces de una planta y es el sitio donde se lleva a cabo el mayor reclutamiento e interacciones microbianas. Las plantas, por medio de los exudados radiculares generan estos ambientes y actúan como fuentes de nutrientes para los microorganismos.

En este trabajo se colectaron 16 muestras de suelo representativas de 8 tipos principales de suelos de México, los cuales fueron caracterizados fisicoquímicamente y utilizados para crecer plantas de tomate (*Solanum lycopersicum*). Se evaluó la producción de biomasa de las plantas en cada suelo y se estudió el microbioma de uno de los suelos antes (S0) y después (SF) del crecimiento de las plantas. Se analizó la diversidad taxonómica y funcional predicha de la rizósfera (RF) y los compartimentos endofíticos (EC) de raíces en un suelo. Además, estos microbiomas se compararon contra otros sistemas de plantas, específicamente, rizósferas de *Arabidopsis thaliana* y endófitos de raíz de otra variedad de *S. lycopersicum*.

Se encontró que los suelos y compartimentos se encuentran dominados por los *phyla* *Proteobacteria*, *Bacteroidetes*, *Firmicutes* y *Verrucomicrobia* y que la diversidad de los compartimentos es alta ( $H'$ (shannon); SF=7.75, RF=6.94, EC=6.63). Al comparar contra otros sistemas se vio que las especies e incluso variedad de planta reclutan distintos microbiomas. La mayor producción de biomasa

fue en el suelo SLP1, siendo más del doble que en los demás suelos, lo cual nos lleva a querer caracterizar su microbioma para explorar su potencial funcional.

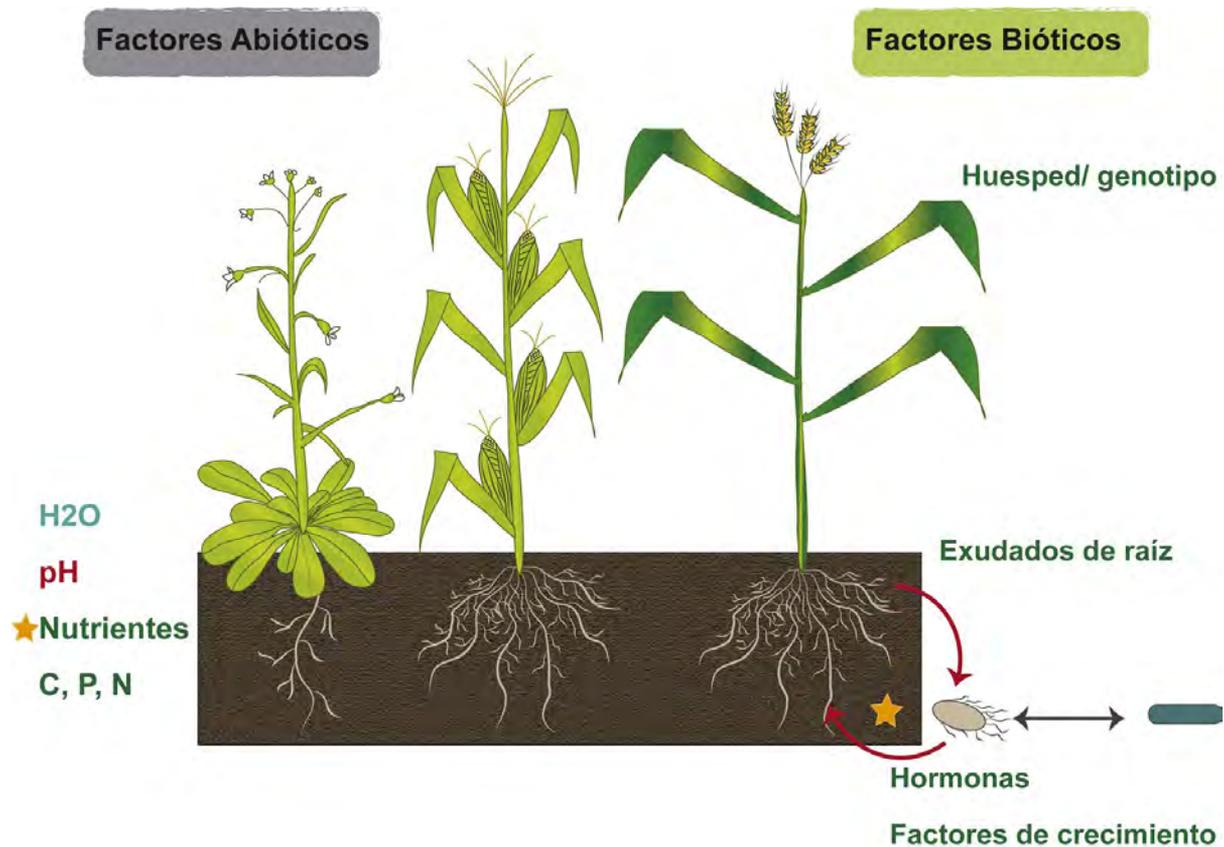
## Introducción

### Diversidad taxonómica y funcional en el microbioma del suelo

El suelo es un ambiente heterogéneo; los diferentes componentes de las fracciones sólidas que lo conforman representan distintos microhábitats para los organismos presentes en éstos, los cuales están sujetos a factores bióticos y abióticos que varían entre los horizontes del suelo (Uksa et al., 2015), actuando como sitios de selección para los microorganismos (Garbeva et al., 2004). En el suelo, los procariontes son los organismos más abundantes, representando el mayor porcentaje de biomasa en dicho ambiente (Daniel, 2005); se estima que es posible encontrar hasta  $10^{11}$  células y hasta 21,000 unidades taxonómicas operativas (OTUs) por gramo de suelo (Roesch et al., 2007).

Las bacterias son elementales para el mantenimiento de la funcionalidad del suelo debido a que están involucradas en la degradación de materia orgánica, remoción de toxinas y los ciclos de carbono, fósforo, nitrógeno y azufre (Garbeva et al., 2004; Madsen, 2011). La estabilidad, supervivencia y el metabolismo de los microorganismos en el suelo están determinados en gran medida por la presencia y disponibilidad de nutrientes y el contenido de agua en los suelos. Una fracción de los microorganismos presentes en el suelo muere después de periodos de sequía e hidratación (Kieft et al., 1987).

A nivel de *Phyla* se ha encontrado que los suelos se encuentran dominados principalmente por *Proteobacteria*, *Firmicutes*, *Actinobacteria*, *Bacteroidetes* y *Verucomicrobia* (Kuramae et al., 2012) Estas tendencias son similares entre distintos tipos de suelos, pero con variabilidad en la diversidad que se puede encontrar, sobre todo a nivel de especies y en escalas biogeográficas amplias (Fulthorpe et al., 2008; Fierer & Jackson, 2005).



**Figura 1. Factores estructurantes de las comunidades bacterianas de suelos y rizósferas.** Diversos factores bióticos y abióticos moldean a las comunidades bacterianas y las interacciones que tienen con las plantas y el medio. Modificada de Laksmanan et al., 2014.

## Factores implicados en la modificación de la estructura de las comunidades microbianas

La vida microbiana puede ser afectada por distintos factores que alteran tanto la estructura de las comunidades como su diversidad (Figura 1). Las hipótesis más aceptadas que explican estos fenómenos son: I) el tipo de plantas que habitan los suelos determinan la estructura de la comunidad. A través de los exudados radiculares las plantas pueden modificar las características físico químicas del suelo y a su vez actúan como las mayores fuentes de carbono y energía (Grayston *et al.*, 1996; Bekku *et al.*, 1997; Bais *et al.*, 2004). y II) el tipo de suelo ha sido propuesto

como el factor determinante en la estructura de las comunidades microbianas, ya que la combinación de la textura del suelo, el pH, la cantidad de materia orgánica y la presencia de nutrientes clave como nitrógeno, fósforo y hierro determina nichos estables (Garbeva *et al.*, 2004; Girvan *et al.*, 2002; Pii *et al.*, 2015). De todas estas variables, el pH ha sido postulado como una de las características más importantes del suelo para predecir los cambios que pueden haber en la composición taxonómica de las comunidades bacterianas presentes en éstos, siendo un parámetro influenciado por la variación en distintas características como la salinidad, cantidad de cationes o compuestos orgánicos (Lauber *et al.*, 2009).

Las concentraciones de nutrientes (C, N y P) y sus relaciones estequiométricas también han sido documentadas como predictores y variables estructurantes de las comunidades bacterianas de suelos. En particular se ha visto que la relación C:N del suelo puede estar diferenciando a las comunidades bacterianas en los suelos (Kuramae *et al.*, 2012; Peralta *et al.*, 2013). En estos estudios se han comparado suelos de distintos ecosistemas y suelos con distintos tratamientos de manejo para cultivo, las relaciones C:N difieren entre cada muestra y explican la mayor parte de la varianza observada entre las distintas comunidades.

La diversidad que se puede encontrar en función de los tipos de suelo varía de acuerdo a las condiciones climáticas en el que el suelo esté presente. Se ha visto que en los suelos de climas áridos existe una diversidad más baja que en suelos donde se practica la agricultura. Las principales diferencias encontradas fueron a nivel de *phylum*, donde los suelos de cultivo presentan una mayor proporción de *Firmicutes* (Köberl *et al.*, 2011). Bajo esta línea, otros autores han demostrado que la cantidad de humedad en los suelos también es determinante en la estructura de las comunidades (Zhang *et al.*, 2012).

## **El microbioma de las plantas y las estrategias para su análisis.**

Las comunidades bacterianas que habitan en el suelo e interactúan directamente con las plantas y en particular con sus raíces han sido evaluadas desde los años 1940s (Lochhead *et al.*, 1940). Diferentes estudios han sido realizados para evaluar el microbioma tanto de partes aéreas, denominadas filósfera (Lindow & Brandl, 2003), como el de las raíces de éstas.

A lo largo de los años diferentes estrategias han sido utilizadas para describir la microbiota habitante en el suelo y las plantas; desde técnicas clásicas dependientes de cultivo (Barea *et al.*, 1976), pasando por técnicas moleculares como perfiles de ácidos grasos (Griffiths *et al.*, 2008), análisis de restricción, determinación de ribotipos (Fulthorpe *et al.*, 1998) y más recientemente secuenciación masiva de bibliotecas de clonas del gen 16S rRNA (Chelius & Triplett, 2001) o bibliotecas de amplicones del mismo gen (Caporaso *et al.*, 2012) y metagenomas totales (Ottesen *et al.*, 2013).

Gracias al desarrollo y los avances en las tecnologías de secuenciación masiva se ha podido incrementar la cantidad de información disponible de organismos no cultivables en estos ambientes. Este tipo de enfoques, utilizando genes marcadores como el 16S rRNA, nos permiten realizar análisis taxonómicos más detallados y con un mayor nivel de profundidad (Kennedy *et al.*, 2014). Hoy en día contamos con tecnologías de secuenciación masiva tales como la pirosecuenciación, utilizada por los secuenciadores 454 Roche se basa en la detección de un pirofosfato liberado durante la incorporación de nucleótidos (dNTPs) específicos por la polimerasa a la cadena creciente de DNA complementaria al molde (Mardis, 2008). En cada ciclo se agrega un nucleótido distinto seguido de un lavado para eliminar a los nucleótidos no incorporados. Esta tecnología genera lecturas de hasta 330 bases. La tecnología de secuenciación por síntesis, utilizada por los secuenciadores de Illumina se basa en la amplificación clonal de una sola secuencia en *clusters*. Posteriormente se agregan los cuatro tipos de nucleótidos marcados con un fluoróforo y bloqueados en el extremo 3'-OH (dATP, dGTP, dCTP y dTTP) que irán complementando a las cadenas molde, añadiendo un nucleótido por ciclo, seguido de un lavado de los dNTPs y el desbloqueo del extremo 3'-OH.

Cada nucleótido incorporado es detectado por el fluoróforo correspondiente de cada dNTP (Liu *et al.*, 2012). Illumina cuenta con diferentes plataformas de secuenciación (HiSeq, MiSeq) que son capaces de realizar secuenciación en configuración *paired-end* lo que permite la obtención de secuencias ensambladas más largas, de hasta 600 pares de bases (Metzker, 2010).

Estas estrategias de secuenciación nos han habilitado para desarrollar metodologías independientes de cultivo para el análisis de la diversidad de comunidades microbianas en distintos ambientes. Este es el campo de estudio de 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). Existen dos vertientes del análisis metagenómico: el análisis de amplicones y el de metagenomas totales. El uso de amplicones del gen 16S rRNA ha sido el enfoque más utilizado hasta la fecha en los estudios de microbiología ambiental. Esta molécula ha sido indicada como el estándar de oro para la identificación de especies bacterianas (Stackerbrandt & Goebel, 1994). Este enfoque 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), definidas generalmente como *clusters* o grupos de secuencias de 16S con un nivel de identidad mayor al 97%. Este criterio operativo ha sido utilizado también para la definición de especies bacterianas (Stackebrandt & Goebel, 1994; Goris *et al.*, 2007). El estudio de comunidades por medio de amplicones permite obtener la información sobre la abundancia relativa de los OTUs encontrados y hacer correcciones por el número de copias que los organismos poseen de este gen, el cual puede variar de 1 a 15 (Kembel *et al.*, 2012). La otra vertiente es la estrategia del metagenoma total, basada en la secuenciación del DNA ambiental, lo cual nos proveerá de información tanto taxonómica como del potencial funcional como se ha visto en estudios de microbiomas de plantas como la planta carnívora *Utricularia Gibba* (Alcaraz *et al.*, 2016).

## **El suelo, un ambiente heterogéneo**

El suelo se puede definir de distintas formas. De acuerdo a la Sociedad Americana de la Ciencias del Suelo (SSSA, por sus siglas en inglés) hay dos definiciones formales: i) 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 y ii) El material mineral u orgánico no consolidado en la superficie de la Tierra, condicionado por el relieve, que ha sido sujeto a factores ambientales y bióticos (clima, biota) y muestra sus efectos sobre el material parental en un periodo de tiempo (SSSA). El suelo es entonces un producto que surge de las interacciones y alteraciones que suceden en las rocas y sedimentos geológicos (Sumner, 2000).

El suelo cuenta con distintas propiedades físicas y químicas. Entre las propiedades físicas encontramos a los factores de la estructura del suelo, profundidad, disponibilidad de agua y color, entre otros. Una de las características físicas importantes del suelo es la textura que corresponde a la distribución relativa de diferentes tamaños de partículas (arena, limos y arcillas), la cual afecta directamente la aireación, movimiento del agua, crecimiento radicular entre otros. El perfil de suelo está conformado por capas bien definidas llamadas horizontes identificados por las letras H, A, B, C y R. El horizonte H es una capa superficial orgánica donde se acumula materia orgánica y se encuentra fresca o en descomposición. EL horizonte A es el superficial mineral donde la materia orgánica descompuesta se asocia con el mineral. El horizonte B es el subsuperficial mineral, en este horizonte pueden existir vestigios de la roca madre. El horizonte C es la capa profunda en las que se muestran las características de la roca madre y no presenta evidencias de desarrollo edáfico. El horizonte R es la capa más profunda, continua y coherente de roca dura que dió origen al suelo (INEGI, 2007). La disponibilidad de agua en un suelo se cuantifica como la capacidad de campo, esto

es el volúmen de agua con el cuál todas las porosidades del suelo se llenan y el aire se desplaza.

Entre las propiedades químicas encontramos al pH, el contenido de macro y micronutrientes (Carbono, Nitrógeno, Fósforo, Potasio, Calcio, Magnesio, etc.) y salinidad.

### Los principales suelos de México

México es un país megadiverso en el sentido biológico y también en el sentido geográfico. El INEGI (2014), reporta que en 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-ISRIC-ISSS, 1998). Los suelos dominantes en el territorio Mexicano son los Leptosoles (28.3% del territorio), Regosoles (13.7%), Feozems (11.7%), Calcisoles (10.4%), Luvisoles (9%) y Vertisoles (8.6%), sumando el 81.7% de los suelos que están presentes en el país (Figuras 2 y 3). El 52.4% del territorio nacional está cubierto por suelos someros y poco desarrollados (Litosoles, Regosoles y Calcisoles), los cuales tienen un bajo aprovechamiento agrícola. Los suelos más fértiles (Feozems, Luvisoles y Vertisoles) ocupan 29.4% del territorio.

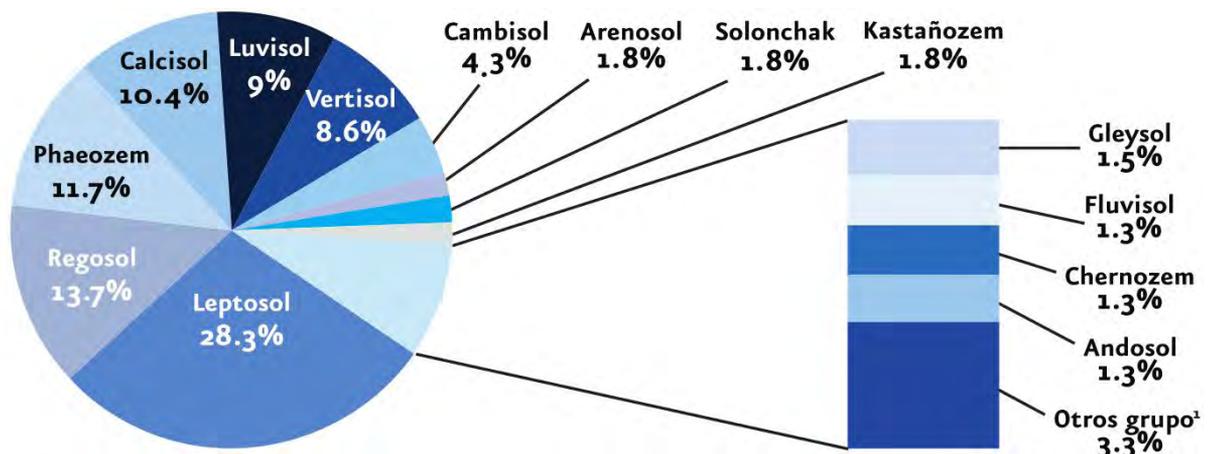


Figura 2. Proporción de tipos de suelo en México. Tomada de Semarnat, 2008.

Entre los principales tipos de suelo considerados en este trabajo, a continuación enumeramos algunas de sus características generales: 1) Los Litosoles son suelos poco desarrollados, delgados, pedregosos y pueden estar en cualquier tipo de roca; 2) 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; 3) Los Feozems se forman sobre material no consolidado y se encuentran en climas templados y húmedos, son ricos en materia orgánica y son utilizados para producción de granos y hortalizas; 4) Los Calcisoles pertenecen a zonas áridas y semiáridas, acumulan calcio y su potencial agrícola es alto bajo condiciones de riego y fertilización; 5) Los Luvisoles se encuentran comúnmente en climas templados y fríos o cálidos y húmedos con periodos de lluvia y sequía. Son suelos muy fértiles y su uso agrícola se basa en la producción de grano y caña de azúcar y finalmente; 6) Los Vertisoles pertenecen a climas semiáridos con estaciones de lluvia y sequía. Se encuentran cercanos a lechos lacustres o riberas. Tienen un alto contenido de arcilla, por lo que pueden acumular humedad. Son muy fértiles pero difíciles de trabajar y se destinan principalmente al cultivo de hortalizas y maíz (Sagarpa, 2008).

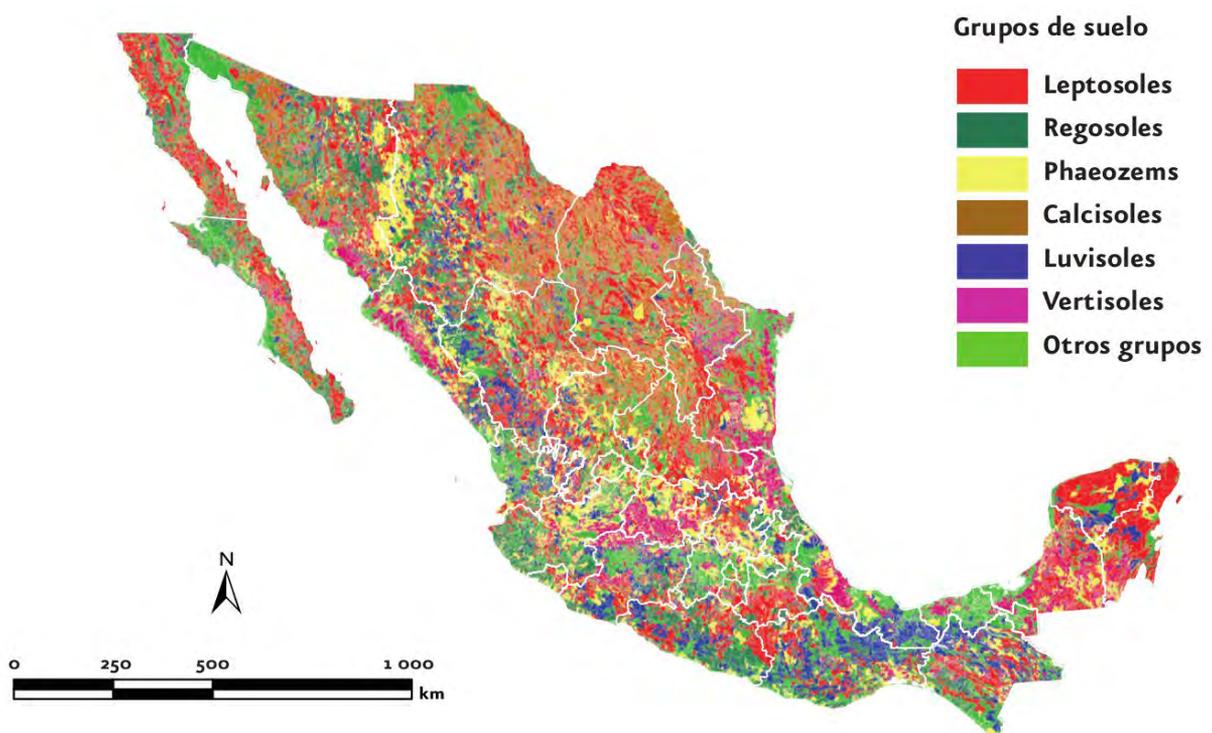


Figura 3. Distribución de los distintos tipos de suelo en el país.

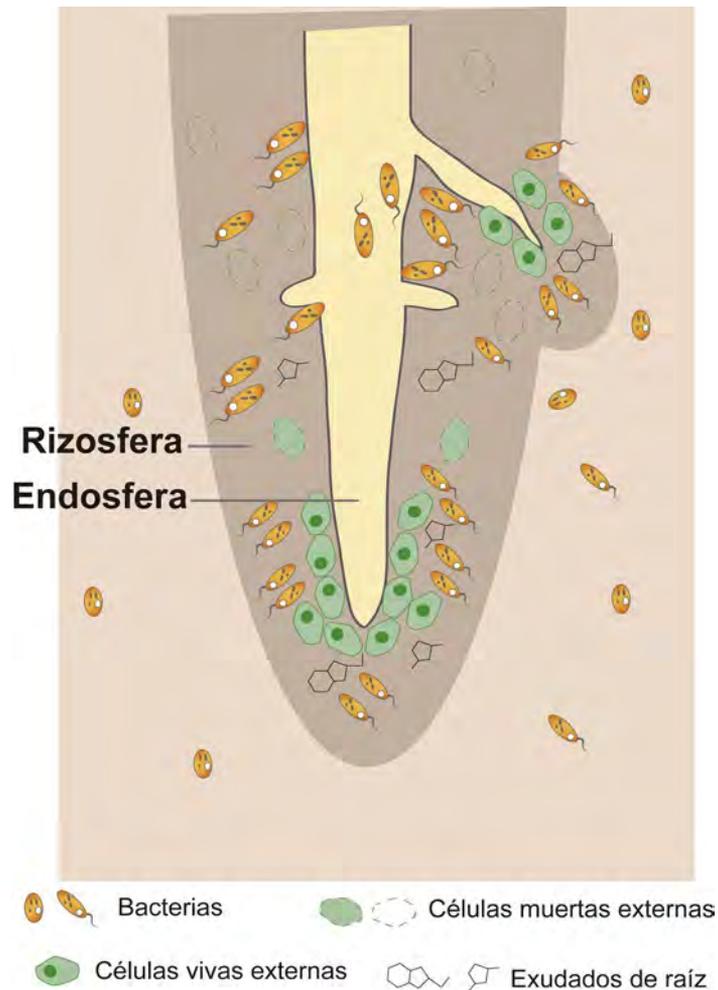
Dependiendo de la ubicación geográfica donde se localicen los suelos, éstos estarán sujetos a las condiciones climáticas del sitio. De acuerdo a Lang (1915), un ecosistema se puede clasificar en tres distintas categorías con base en el índice de aridez (IA), el cual se calcula como la temperatura anual promedio, entre la precipitación anual promedio en un sitio. Con esta métrica se definen a los climas áridos ( $IA < 40$ ), húmedos ( $40 < IA < 160$ ) e hiperhúmedos ( $IA > 160$ ).

## **Antecedentes**

La rizósfera se define como la capa de suelo estrechamente unida a las raíces de una planta (Bais *et al.*, 2006); el término fue utilizado por primera vez por Lorenz Hiltner en 1904 (Hartmann *et al.*, 2007). Ésta es la zona de la planta con el mayor reclutamiento e interacciones microbianas, ya sea con microorganismos patógenos o benéficos los cuales contribuyen en el crecimiento y salud de las plantas (Berendsen *et al.*, 2012). Además de este ambiente particular, se ha definido al compartimento endofítico como los sitios intra o intercelulares donde los microorganismos pueden habitar formando relaciones simbióticas, comensalistas o mutualistas (Ryan *et al.*, 2007).

Estos ambientes son generados por las plantas debido a los exudados radiculares liberados, los cuales son compuestos que actúan como una fuente de nutrientes para los microorganismos y también son determinantes de la actividad y densidad de sus poblaciones (Raaijmakers *et al.*, 2008). Se ha observado que las rizósferas presentan una composición taxonómica distinta de la presente en el suelo circundante representando un subconjunto de las comunidades de éste, lo que podría implicar una probable selección de los microorganismos por parte de la planta (Berendsen *et al.*, 2012). Por otro lado, se ha visto que el tipo de suelo y sus variables fisicoquímicas son un componente importante para la estructuración de las comunidades de la rizósfera. El efecto generado por estas variables ha mostrado explicar una mayor parte de la varianza entre microbiomas que parámetros

estudiados como la etapa de desarrollo de la planta o el genotipo de éstas (Lundberg *et al.*, 2012; Inceoğlu *et al.*, 2010).



**Figura 4. Los distintos compartimentos de la raíz.** Diferentes nichos pueden ser habitados por miembros de la microbiota de raíz desde el suelo en el exterior hacia la rizósfera y los espacios inter e intracelulares de las plantas. Adaptada de Bulgarelli *et al.*, 2015.

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) entre otros. En particular los estudios de Bulgarelli y Lundberg (2012) han llevado a cabo el análisis de compartimentos endofíticos, enriqueciéndolos por medio de un procedimiento de sonicación de las raíces con el fin de liberar a los

microorganismos de estos espacios. En general los estudios en *Arabidopsis* han encontrado que la rizósfera se encuentra dominado por los *phyla Proteobacteria, Acidobacteria* y *Bacteroidetes*, mientras que el compartimento endofítico lo domina el *phylum Actinobacteria*.

El tomate (*Solanum lycopersicum*), es una planta de interés agronómico y un organismo modelo en el desarrollo de frutos (*The Tomato Genome Consortium*, 2012). Esta planta se encuentra en el cuarto lugar de los vegetales producidos mundialmente. Su centro de origen no ha sido determinado con exactitud pero se tiene la hipótesis que ha sido en México o Sudamérica. Su distribución es amplia debido a su interés agronómico y su capacidad de crecer desde el nivel del mar hasta los 3,300 metros de altura (Peralta & Spooner, 2008). Diversos estudios han sido realizados para el análisis del microbioma de esta planta a nivel de filósfera (Ottesen *et al.*, 2013), rizósfera y compartimentos endofíticos (Tian *et al.*, 2015; Upreti *et al.*, 2015).

Ottesen y colaboradores (2015) llevaron a cabo un análisis metagenómico total de partes aéreas y subterráneas del tomate. En las raíces encontraron que la diversidad bacteriana era mayor que en la filósfera y encontraron particularmente a los géneros *Chryseobacterium*, *Leifsonia*, *Microbacterium*, *Arthrobacter* y *Burkholderia*. Varias cepas de éste último género han sido aisladas de compartimentos endofíticos de tomate por el grupo de Caballero Mellado (2007) y probadas para su capacidades de fijación de nitrógeno, producción de ACC deaminasa y solubilización de fósforo.

Tian y colaboradores (2015) realizaron estudios comparativos de los microorganismos endófitos de tomates sanos contra tomates infectados por nemátodos, encontrando que los compartimentos endofíticos de tomates sanos eran dominados por el orden de los *Streptomycetales, Rhizobiales* y *Micromonosporales*.

En los trabajos de Upreti y colaboradores (2015) se llevaron a cabo aislamientos de endófitos de raíces de variedades de tomates susceptibles y

resistentes al patógeno *Ralstonia solanacearum*, logrando aislar bacterias de los géneros *Pseudomonas*, *Citrobacter*, *Enterobacter*, *Rhizobium*, *Pantoea* y *Stenotrophomonas*.

En este trabajo se plantea la caracterización fisicoquímica de los suelos colectados, así como la descripción y análisis del microbioma de los distintos compartimentos de raíz de tomates, por medio de secuenciación de bibliotecas del gen rRNA 16S y la predicción de funciones metabólicas a partir del perfil filogenético descrito a partir de la comparación de los genes 16S con las bases de datos.

## Objetivo general

Caracterizar las propiedades fisicoquímicas de suelos contrastantes colectados en México de acuerdo a la carta edafológica. Caracterizar la interacción planta-rizósfera-suelo describiendo la composición taxonómica y una predicción metabólica para evaluar si existen diferencias en los microbiomas rizósfera-suelo y si efectivamente se compartimentaliza taxonómica y metabólicamente; para dicho efecto utilizaremos rizósferas y compartimentos endofíticos de *Solanum lycopersicum* cultivadas en uno de los suelos colectados.

## Objetivos particulares

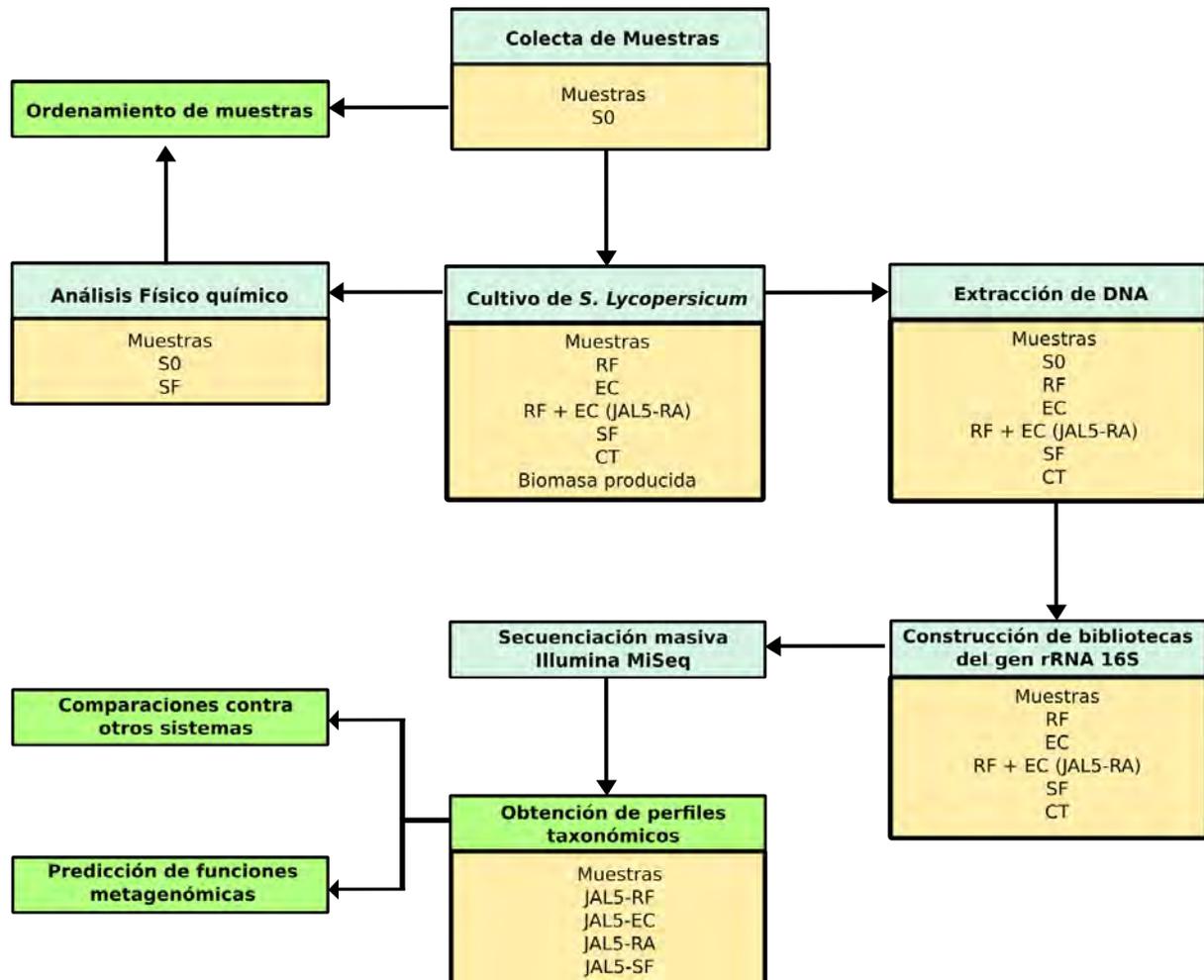
- Hacer la colecta de al menos 5 tipos de suelos contrastantes.
- Determinar las características físicoquímicas de los suelos.
- Hacer un experimento de jardín común, para contrastar el efecto del suelo en el microbioma.
- Generar el perfil filogenético del microbioma de los compartimentos de la raíz, mediante el análisis y comparación de secuencias del gen 16S rRNA.
- Hacer predicciones metabólicas a partir del perfil filogenético por cada compartimento de la raíz analizado.
- Comparar los perfiles filogenéticos con otros sistemas análogos.

## Hipótesis

El microbioma de rizósfera y compartimento endofítico de las raíces de *Solanum lycopersicum* representará un subconjunto menos diverso que el microbioma de su suelo circundante y a su vez, distinto al de compartimentos de raíces en otros sistemas.

# Metodología

## Esquema general del trabajo



**Figura 5. Esquema general del trabajo.** Utilizando distintos tipos de suelos colectados, se sembraron plantas de *S. lycopersicum* para obtener bibliotecas de amplicones del gen 16S rRNA a partir del DNA extraídos del suelo al tiempo inicial (S0), suelo al tiempo final del crecimiento (SF), rizósfera (RF), compartimento endofítico (EC) y compartimento de raíz (RA). Adicionalmente se hizo la cuantificación de nutrientes en S0 y SF.

## Colecta de distintos tipos de suelo

Se utilizó la carta edafológica de la República Mexicana, provista por la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO, 1995) para la planeación de un transecto en el cual se cubrieran los puntos para el muestreo de los principales tipos de suelo de México (Figura 6). En cada punto de colecta se tomaron los siguientes metadatos: Fecha de colecta, coordenadas, altitud, observaciones generales del sitio de muestreo y temperatura del suelo.



**Figura 6. Mapa del territorio mexicano.** Los puntos rojos señalan con el identificador de la muestra correspondiente a los sitios de colecta de suelos.

En cada localidad seleccionada se colectaron aproximadamente 2 Kg de suelo a una profundidad máxima de 10 cm. Para la toma de muestras se utilizaron

guantes estériles nuevos en cada sitio de muestreo. Las palas utilizadas se lavaron y desinfectaron con etanol al 70% al iniciar y finalizar cada colecta. Los suelos fueron conservados a 4°C en oscuridad. Un kilogramo de suelo fue transportado al IIES, UNAM, campus Morelia y utilizado para realizar el análisis de la composición fisicoquímica. El otro kilo fue utilizado para llevar a cabo la siembra de plantas en condiciones de invernadero. Adicionalmente se tomaron muestras de suelo por duplicado en tubos Falcon estériles de 50 mL, los cuales fueron congelados en nitrógeno líquido para la extracción de DNA.

### **Análisis físico-químico de los suelos.**

Los siguientes parámetros fisicoquímicos fueron evaluados en todos los suelos: pH, Nitrógeno total (NT), fósforo total (PT), carbono orgánico total (COT) y carbono inorgánico total (CIT).

Las muestras de suelos fueron preparadas previamente a los análisis de contenido de nutrientes secándose en un horno a 75°C durante 24 horas, se eliminaron tanto la materia orgánica (raíces y hojas) como las arenas grandes y finalmente se molió cada muestra en un mortero de ágata.

El carbono total y carbono inorgánico fueron determinados por el método coulométrico (Huffman, 1977) en un coulómetro UIC CM5012. El carbono orgánico total se calcula como la diferencia entre el carbono total y el carbono inorgánico total.

Para la cuantificación de Nitrógeno total (NT) y Fósforo total (PT), las muestras fueron digeridas en H<sub>2</sub>SO<sub>4</sub> concentrado, H<sub>2</sub>O<sub>2</sub> al 30%, K<sub>2</sub>SO<sub>4</sub> y CuSO<sub>4</sub> a 375°C. El NT se determinó por el método semi Kjheldahl, mientras que el PT fue determinado por el método colorimétrico del molibdato después de una reducción con ácido ascórbico (Murphy & Riley, 1962). Ambas determinaciones se realizaron en un Autoanalizador Bran-Luebbe III (Noderstedt, Germany).

El pH de los suelos en todas las condiciones fue medido utilizando un potenciómetro. Para realizar la medición se homogeneizó cada muestra de suelo, se pesaron 10g de cada uno y se agitaron en 50 mL de agua destilada durante 30 minutos en un agitador orbital. Posteriormente se realizó la medición.

### Diseño experimental para el cultivo de tomates en jardín común

Se llevó a cabo un experimento de jardín común, en el cual se sembraron dos plantas de *S. lycopersicum* en macetas duplicadas conteniendo cada uno de los suelos muestreados dentro de un invernadero, más una maceta con suelos control sin plantas pero que recibieron el mismo tratamiento de riego. Las macetas fueron codificadas numéricamente y dispuestas de manera aleatoria sobre la mesa; el riego se hizo cada tercer día hasta los 60 días de crecimiento, tiempo tras el cual se llevó a cabo la toma de muestras.



Condición	Suelo inicial	Suelo Control	Plantas Réplica 1	Plantas Réplica 2
No. de muestras	16	16	16	16
Clave/Tipo de muestra	S0	CT	SF RF EC	SF RF EC (RA)
	BGQ	BGQ	BGQ	BGQ

**Figura 7. Diseño del experimento.** Se contaban con 16 muestras de suelo inicial (S0) para extracción de DNA y análisis físico químicos (BGQ). Cada suelo se utilizó para llenar tres macetas, en dos de ellas se crecerían plantas de *S. lycopersicum* y se obtuvieron las muestras de rizósfera (RF), compartimento endofítico (EC) y suelo final (SF); la tercera sirvió como control (CT) sin planta, recibiendo el mismo tratamiento de riego. Con las plantas de la maceta réplica 2 de JAL5 se obtuvo el compartimento raíz (RA). Muestras de suelo final fueron tomadas para análisis físico químicos.

## **Germinación de plantas**

La superficie de semillas de *S. lycopersicum* fue desinfectada haciendo un lavado en etanol al 70% durante 1 minuto y posteriormente en NaOCl al 30% durante 3 minutos. Las semillas se enjuagaron en agua destilada estéril, se secaron al aire y se colocaron en cajas Petri con agar al 1% para ser germinadas en una cámara de crecimiento a 27°C durante 96 horas en oscuridad, tiempo tras el cual se eligieron 5 plántulas para transplantarlas a macetas con los 16 tipos de suelo por duplicado.

## **Crecimiento de plantas en distintos tipos de suelo**

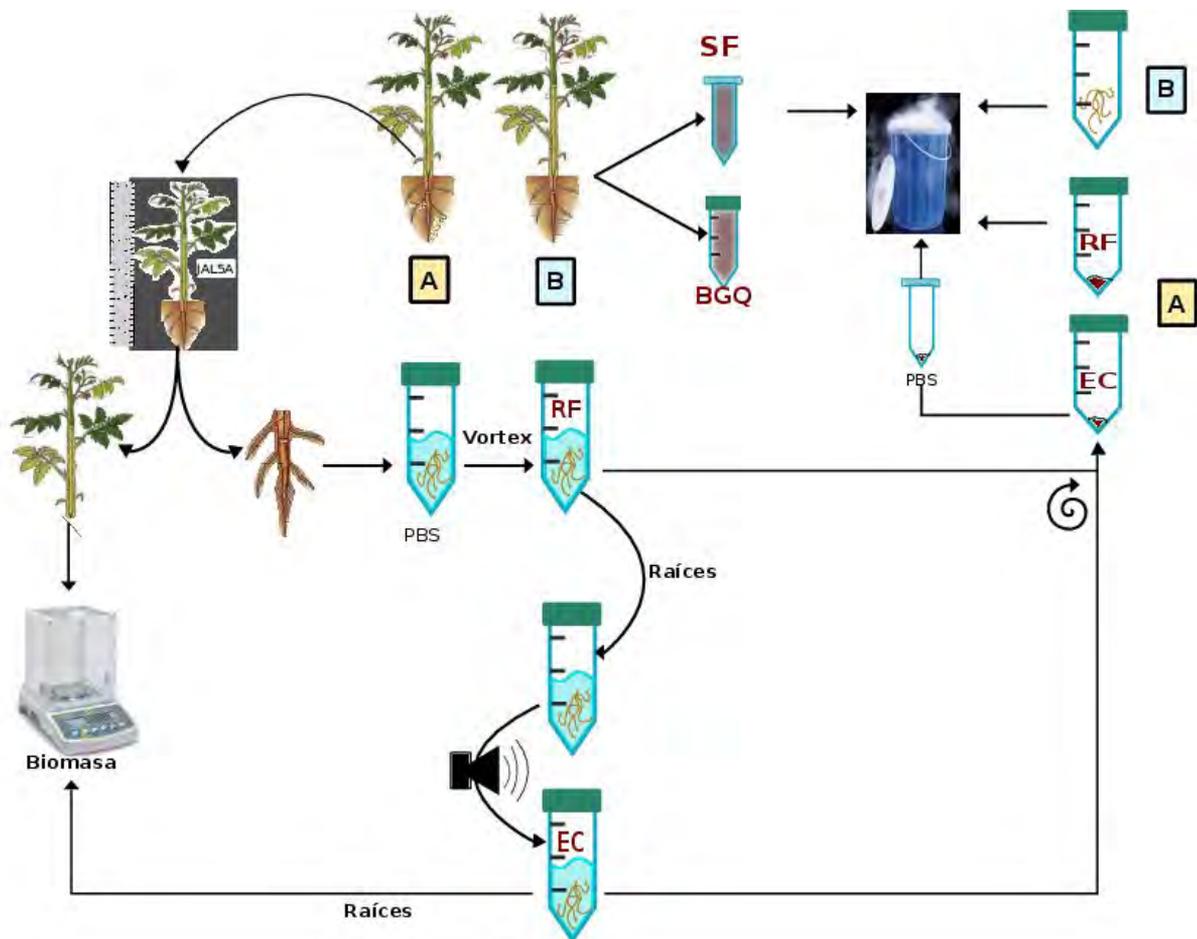
Para llevar a cabo el cultivo de *S. lycopersicum* en el experimento de jardín común, los suelos fueron homogeneizados previamente al llenado de macetas. Cinco plántulas de tomate germinadas en agar se transplantaron a macetas, las cuales fueron llenadas con cada tipo de suelo por duplicado, más una maceta control de cada tipo de suelo en donde no se sembraron plantas. Tras el establecimiento de las plantas (25 días) se seleccionaron a los 2 individuos de mayor talla y se removieron a los 3 restantes, dejando así dos réplicas biológicas por cada maceta. Las plantas fueron regadas cada tercer día, pero el esquema fue modificado a cada segundo día debido al crecimiento de las plantas y el incremento en la temperatura en el invernadero que provocaba desecación de los suelos.

Las plantas se dejaron crecer hasta los 60 días bajo condiciones de fotoperiodo natural utilizando la pantalla en el invernadero para controlar la cantidad de luz incidente en las plantas. Los datos de temperatura y humedad relativa del invernadero fueron registrados cada 2 horas a lo largo del experimento con el documentador ambiental Hobbo.

## **Toma de muestras de suelo final, rizósfera y compartimento endofítico**

Transcurridos los 60 días de crecimiento se colectaron las muestras de rizósferas (RF), compartimentos endofíticos (EC) y suelos circundantes (SF) de las plantas, así como los suelos control (CT). Adicionalmente se tomaron muestras para análisis fisicoquímicos al final del experimento.

Para hacer la toma de muestras, se sacaron las plantas de las macetas, se sacudieron las raíces sobre una superficie de papel estéril para retirar el exceso de suelo adherido a las raíces, dejando solo el suelo de la rizósfera. Las raíces fueron separadas de la parte aérea de la planta y lavadas en buffer de fosfatos (PBS) estéril para obtener la rizósfera, después se transfirieron a un tubo con buffer fresco y se sonicaron para obtener la muestra de compartimento endofítico. Una vez terminado el proceso de cada raíz, éstas se juntaron con su parte aérea correspondiente y se secaron a 70°C hasta obtener peso constante, el peso seco se registró como biomasa producida por cada planta. Todo el material de metal utilizado se desinfectó con etanol al 98% y fuego antes de comenzar con cada muestra. El protocolo utilizado para la colecta de las muestras se detalla en el Anexo 2 de procedimientos de laboratorio (<https://drive.google.com/open?id=1jxjcrXdTpU2AIKEmVCcxGYuVP08TBm7jBRHk1d9zCs>) y se presenta esquematizado en la Figura 8.



**Figura 8. Procedimiento de separación de compartimentos de raíz.** Las raíces de *S. lycopersicum* se separaron de las partes aéreas y se sometieron a tratamientos de lavado y sonicación para obtener las muestras de RF y EC. Una réplica de raíz de JAL5 fue lavada y sonicada en conjunto para obtener el compartimento de raíz.

Debido a que se contaba con un duplicado de cada suelo y dos plantas en cada maceta (identificadas como planta A y planta B), se procedió a procesar por completo ambas plantas de una de las réplicas y se congelaron las raíces completas de la segunda réplica sin procesarlas. Se muestra un ejemplo de los identificadores utilizados para cada caso:

Muestra procesada: JAL5-RF-A, JAL5-EC-A, JAL5-SF-A

Muestra sin procesar: JAL5-A2-raíces

Para comprobar la eficiencia del proceso de sonicación, una de las réplicas de plantas sembradas en JAL5 se procesó de la manera mencionada anteriormente, mientras que para la réplica restante se modificó el procedimiento, llevando a cabo la sonicación en el mismo tubo donde se obtuvo la rizósfera. Esta muestra fue denominada como el compartimento raíz (JAL5-RA).

### **Extracción de DNA de suelos.**

El DNA de suelos de la condición inicial y final (transcurrido el tiempo de crecimiento de las plantas de tomate) fue extraído con el kit PowerSoil DNA Isolate Kit (MoBio laboratories, Solana Beach, CA, USA) siguiendo el protocolo provisto por el fabricante. La extracción por este método incluye un paso de golpeo con perlas para la lisis celular y reactivos para la disolución y precipitación de ácidos húmicos que pueden inhibir las reacciones en cadena de la polimerasa (PCR) (Matheson *et al.*, 2010). Las extracciones se conservaron a  $-20^{\circ}\text{C}$  para su posterior procesamiento. La calidad del DNA se determinó por electroforesis en gel de agarosa al 0.8%. Las condiciones utilizadas para correr los geles fueron: 90V por 30 minutos a temperatura ambiente en buffer TRIS/borato/EDTA (TBE) 0.5x estéril, pH 8.3. El DNA de cada extracción se cuantificó por método espectrofotométrico con la relación  $\text{OD}_{260}/\text{OD}_{280}$  con el equipo NanoDrop lite (Nanodrop Technologies, Wilmington, DE). El protocolo se detalla en el Anexo 2 de procedimientos de laboratorio (<https://drive.google.com/open?id=1jxjcrXdTpU2AIKEmVCcxGYuVP08TBm7jBRHk1d9zCs>).

### **Extracción de DNA de rizósfera y compartimento endofítico.**

El DNA de rizósfera y compartimento endofítico colectados de las raíces de las plantas cultivadas en cada tipo de suelo en invernadero se extrajo con el kit

mencionado en el apartado anterior. Para realizar esto, se resuspendieron los pellets obtenidos de cada compartimento en 200 uL de PBS estéril. El pellet resuspendido se transfirió a los tubos de extracción del kit Power Soil y se procedió con el protocolo de manera estándar.

### **Amplificación del gen ribosomal 16S rRNA.**

Se llevó a cabo la amplificación por medio de PCR de los genes ribosomales 16S a partir del DNA genómico utilizando los primers 341F y 805R. Este par de primers amplifican las regiones hipervariables V3-V4 del gen rRNA 16S (Muyzer *et al.*, 1993). A continuación se presentan las secuencias, las cuales cuentan con una región adaptadora para secuenciación en la plataforma Illumina MiSeq. El segmento correspondiente al primer para el gen del rRNA 16S se resalta en negritas:

341F: 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGAC**CTACGGGNGGCWGCAG** 3'

805R: 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAG**GACTACHVGGGTATCTAATCC** 3'

Para llevar a cabo las reacciones se utilizó la polimerasa Pfx Platinum (Invitrogen). Esta polimerasa cuenta con buffer de reacción y un buffer *enhancer*, el cual fue utilizado para mejorar la eficiencia de amplificación. Las reacciones se realizaron por duplicado en un volumen final de 20 uL cada una y se combinaron en un solo volumen. En los casos en los que la amplificación resultó negativa con esta enzima, se usó la polimerasa Taq (Promega. Madison, WI, USA) para la amplificación, seguida de la reparación de los extremos con el fragmento Klenow (New England Biolabs). Las concentraciones finales de reactivos utilizados con la polimerasa Pfx y Taq se presentan en las tablas 1 y 2 respectivamente. Se utilizó el mismo programa de PCR (Tabla 3) en ambos casos, modificando solamente la temperatura de extensión que en el caso de Pfx es 68°C y para Taq es de 72°C.

Reactivo	Concentración final
Buffer Pfx	1X
Buffer enhancer	1X
dNTPs	200 $\mu$ M
MgSO <sub>4</sub>	1.5 mM
F primer	0.3 $\mu$ M
R primer	0.3 $\mu$ M
Pol. Pfx	0.02 U/ $\mu$ L

**Tabla 1. Concentraciones finales de reactivos para la amplificación con polimerasa Pfx.**

Reactivo	Concentración final
Taq polimerasa	0.03 U/ $\mu$ L
F primer	0.3 $\mu$ M
R primer	0.3 $\mu$ M
dNTPs	250 $\mu$ M
MgCl <sub>2</sub>	1.87 mM

**Tabla 2. Concentraciones finales de reactivos para amplificación con polimerasa Taq**

Ciclo	No. Ciclos	Temperatura (°C)	Tiempo
Desnaturalización	1X	95	3 min
Desnaturalización	5X	94	30 s
Hibridación		55	30 s
Extensión		68*/72**	30 s
Desnaturalización	25X	94	5 s
Extensión		68*/72**	30 s

**Tabla 3. Programa de PCR utilizado.** \* Temperatura para Pfx. \*\* Temperatura para Taq.

### **Purificación de amplicones y reparación con fragmento Klenow**

Los amplicones de todas las muestras fueron purificados a través de una columna de sílica con el kit Wizard SV Gel and PCR Cleanup System (Promega, Madison, WI, USA) directamente de la reacción de PCR de acuerdo al protocolo del fabricante.

Las muestras amplificadas con la polimerasa Taq fueron purificadas previamente a la reparación con el fragmento Klenow. La columna de cada muestra

se guardó y se volvió a realizar la purificación por la columna correspondiente a cada muestra después de haber realizado la reparación.

Para realizar la reparación se mezclaron los 50 uL de amplicones resultantes de la primera limpieza con 5.5 uL de buffer de reacción (1X), 0.16 uL de Klenow (5 U/uL) , y 0.7 uL de dNTPs (2.5 mM). El volumen final de reacción fue de 56.3 uL. Se realizó una incubación por 15 minutos a 25°C y una inactivación por calor durante 20 minutos a 70°C.

### **Secuenciación masiva.**

La secuenciación de amplicones de rRNA 16S de JAL5-SF, JAL5-RF, JAL5-EC y JAL5-RA se realizó en el equipo Illumina MiSeq con una configuración 2 x 300 pair-end en la Unidad de Secuenciación Masiva del Instituto de Biotecnología de la UNAM.

### **Procesamiento y análisis de datos de secuenciación.**

#### **Análisis de calidad y ensamblado**

Los datos de secuenciación fueron recibidos en el formato FASTQ, el cual contiene las secuencias de las lecturas y los valores de calidad por cada base. Se graficó la calidad de las lecturas crudas con la herramienta Fastx toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)) con la finalidad de conocer la calidad de las lecturas de la corrida de secuenciación y poder planear el ensamblado y filtrado de calidad de las secuencias ya ensambladas.

Las secuencias fueron ensambladas con el programa Pandaseq (Masella *et al.*, 2012). Este programa genera archivos de salida en formato FASTA donde se

encuentran todas las secuencias ya ensambladas. Para realizar los ensamblados se pidió que la longitud mínima y máxima de secuencias utilizadas fuera de 250 y 470 pb respectivamente, que el empalme entre lecturas fuera de al menos 15 pb y que se excediera un umbral del 98% de certidumbre de la asignación de base correcta. Las secuencias con baja calidad o bases ambiguas se descartaron para los análisis subsecuentes.

### **Asignación taxonómica**

Para asignar OTUs se hizo un agrupamiento de secuencias con CD-HIT-EST (Li *et al.*, 2006) al 97% de identidad de secuencias. Posteriormente se generó una tabla de OTUs con QIIME (Caporaso *et al.*, 2010) y los índices de diversidad alfa fueron calculados con el paquete de R phyloseq (McMurdie *et al.*, 2013). Los índices considerados son: especies observadas, Shannon, Simpson, así como el estimador no paramétrico Chao1.

Los archivos de las secuencias ensambladas de JAL5-SF, JAL5-RF, JAL5-EC y JAL5-RAIZ se cargaron en el servidor MG-RAST (Meyer *et al.*, 2008) para realizar la asignación taxonómica contra la base de datos *Greengenes* (De Santis *et al.*, 2006) usando el algoritmo de “*Best Hit Classification*” con los parámetros de longitud mínima de alineamiento de 15 pares de bases, *e-value* de  $1e-5$  y un mínimo de identidad del 97% como se ha reportado para la identificación de especies bacterianas (Stackebrandt & Goebel 1994). Al aplicar estos parámetros, las secuencias se comparan de forma pareada contra la base de datos y todas las secuencias que cumplen con el criterio se agrupan y se reportan como un OTU. Las tablas de abundancia taxonómica de cada muestra fueron descargadas en formato “*biom*”, que fue convertido a formato tabular y se calculó la abundancia relativa a nivel de *phylum* por cada compartimento.

## Predicciones funcionales

Se llevó a cabo la predicción funcional a partir de las secuencias del gen 16S rRNA de JAL5 de forma global utilizando la versión en línea del programa PICRUST (Langille *et al*, 2013) (<https://huttenhower.sph.harvard.edu/galaxy/>) contra la base de datos KEGG (Kanehisa & Goto 2000). Para realizar este análisis los archivos FASTA de los ensamblados de las muestras JAL5-SF, JAL5-RF, JAL5-EC Y JAL5-RA se concatenaron en un solo archivo, sobre el cual se realizó la asignación taxonómica contra una base de datos de organismos de referencia de *Greengenes* que solo contiene los OTUs de los cuales existen genomas completamente secuenciados, resultando en una tabla de OTUs. La tabla de OTUs fue normalizada por el número de copias del gen ribosomal en cada organismo sobre el cual se hizo la asignación taxonómica contra la base de datos *Greengenes*. Finalmente, se realizó la predicción funcional del metagenoma mediante una estrategia de inferencia de contenido génico, en donde a partir de un árbol filogenético de referencia de los OTUs, se obtienen tablas de contenido génico para los organismos de los cuales se cuente con dicha información. De los nodos terminales en los que no se cuente con la información de contenido génico, se hace una inferencia a partir de la reconstrucción de caracteres ancestrales para ese OTU, asignándole una tabla de contenido génico probable.

Las tablas de funciones predichas por compartimento se transformaron aplicando el logaritmo base 2 con el objetivo de normalizar las cuentas y se graficaron a las 50 más abundantes en forma de un mapa de calor en R, utilizando el paquete *phyloseq*. El procesamiento de estos datos se detalla en el Anexo 1 de scripts

(<https://drive.google.com/open?id=1jxjcrXdTpU2AIKEmVCcxGYuVP08TBm7jBRHk1d9zCs>).

## **Análisis de ordenamiento de muestras de suelo**

Utilizando los datos de las variables físico-químicas de los suelos y los datos del índice de aridez se hizo un ordenamiento de las muestras de suelo mediante un análisis de componentes principales en R con la librería “stats” para observar las diferencias que pueden existir entre cada uno de los suelos colectados. Los comandos utilizados para este proceso de detallan en el Anexo 1 de scripts (<https://drive.google.com/open?id=1jxjcrXdTpU2AIKEmVCcxGYuVP08TBm7jBRHk1d9zCs>).

## **Comparaciones con otros sistemas**

Utilizando las secuencias cargadas en MG-RAST que sirvieron para realizar la asignación taxonómica, se llevó a cabo un análisis comparativo del microbioma de cada compartimento de JAL5 contra proyectos públicos de amplicones de rRNA 16S de microbiomas asociados a raíces de plantas. Las comparaciones se realizaron contra el microbioma de rizósfera de *Arabidopsis thaliana* (Lundberg *et al.*, 2012) y contra el microbioma del compartimento endofítico en plantas de tomate sanas (Tian *et al.*, 2015).

Se descargó una tabla de abundancias de OTUs en formato BIOM de todas las muestras de JAL5, más las muestras de *Arabidopsis* con los identificadores ERR131835rz, ERR131834rz, ERR131826so, ERR131825so y una muestra de endófitos de tomate con el identificador 4603086.3. Esta tabla fue convertida a formato tabular y con ella se realizó un ordenamiento de las muestras con los datos de abundancia por medio de un análisis de correspondencias para observar las diferencias que existen entre los distintos microbiomas. Los comandos utilizados para este proceso se detallan en el Anexo 1 (<https://drive.google.com/open?id=1jxjcrXdTpU2AIKEmVCcxGYuVP08TBm7jBRHk1d9zCs>).

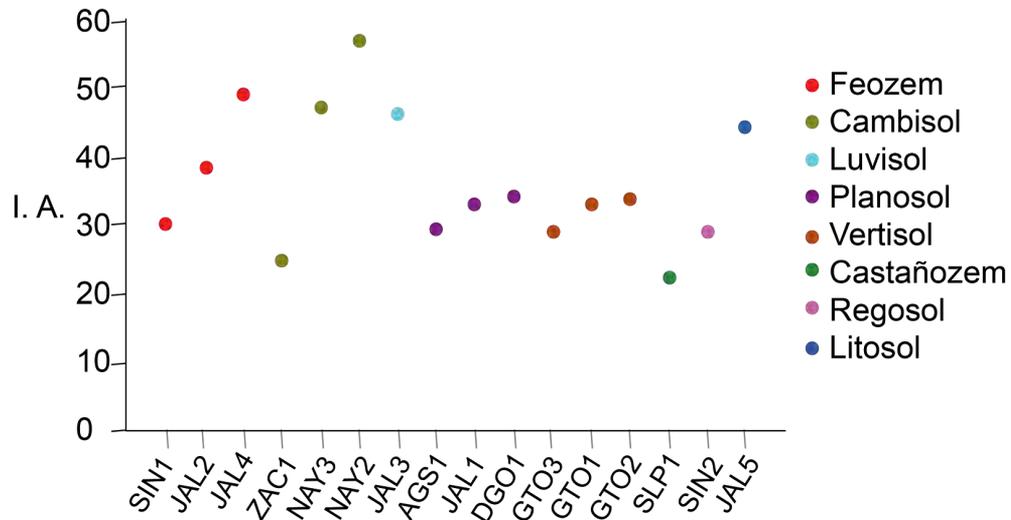
## Resultados y discusión

### Clasificación y características ambientales de los tipos de suelo colectados en este trabajo.

Se colectaron 16 diferentes tipos de suelo de México. La clasificación general por tipo de suelo se llevó mediante la búsqueda en las cartas edafológicas del territorio mexicano (CONABIO, 1995) y las coordenadas de muestreo. Las unidades principales de suelos colectados corresponden a Castañozem (1), Planosoles (3), Feozems (3), Luvisol (1), Vertisoles (3), Regosoles (1), Litosoles (1) y Cambisoles (3) (Tabla 4). De esta última unidad de suelo, la muestra ZAC1 no es un suelo *sensu stricto*, esta muestra fue tomada de un jal minero en Zacatecas, donde existe la presencia de cobertura vegetal.

Sitio	Tipo de suelo	Latitud	Longitud	Altitud (msnm)	Temperatura De muestreo (°C)	Temperatura Anual promedio(°C)	Precipitación Anual promedio(mm)	I.A.
JAL1	Planosol (W)	21.1624	-101.8583	1809	29	19	625	32.9
JAL2	Feozem (H)	21.2533	-102.2976	1791	24	19	725	38.2
JAL3	Luvisol (L)	20.8210	-102.7977	1862	19	19	875	46.1
JAL4	Feozem (H)	20.8491	-103.8546	1490	30	20	975	48.8
JAL5	Litosol (I)	21.1790	-104.5404	1328	23	21	925	44.0
NAY2	Cambisol (B)	21.8364	-105.0830	38	30	27	1525	56.5
NAY3	Cambisol (B)	22.1416	-105.2614	4	28	26	1225	47.1
GTO1	Vertisol (V)	20.5795	-100.9470	1708	17	19	625	32.9
GTO2	Vertisol (V)	20.7208	-101.3293	1710	21	20	675	33.8
GTO3	Vertisol (V)	20.8966	-100.6739	2012	22	18	525	29.2
SIN1	Feozem (H)	22.9863	-105.8795	23	28	26	775	29.8
SIN2	Regosol (R)	23.2927	-106.4787	10	36	25	725	29.0
SLP1	Castañozem (K)	21.9878	-101.2768	2130	20	17	375	22.1
AGS1	Planosol (W)	21.8276	-102.1208	2059	21	18	525	29.2
DGO1	Planosol (W)	24.0062	-104.3848	1906	15	17	575	33.8
ZAC1	Cambisol (B)	22.8181	-102.6957	2221	17	17	425	25.0

**Tabla 4. Datos ambientales y geográficos de los sitios de colecta de suelos.** Se muestra el identificador de cada muestra, tipo de suelo, coordenadas, altitud, temperatura del sitio, precipitación anual e índice de aridez de cada sitio.



**Figura 9. Suelos colectados en una variedad de condiciones climáticas presentadas como el índice de aridez.** Los suelos se disponen de acuerdo a su capacidad de retención de agua siendo SIN1 el de mejor capacidad de retención y JAL5 el menor.

En la Figura 9 se puede observar que los diferentes suelos colectados se ubican en sitios con distintas condiciones climáticas. En el eje X se encuentran cada uno de los suelos colectados, ordenados con base en su capacidad de retención de agua y en Y el índice de aridez (IA) calculado utilizando los datos históricos de precipitación y temperatura para cada sitio consultados en el sitio web del Centro de Ciencias de la Atmósfera, UNAM (<http://uniatmos.atmosfera.unam.mx/ACDM/>, 2015). Un mayor IA indica que los suelos tienen una mayor disponibilidad de agua anualmente, mientras que un IA bajo, indica una baja disponibilidad de agua por año.

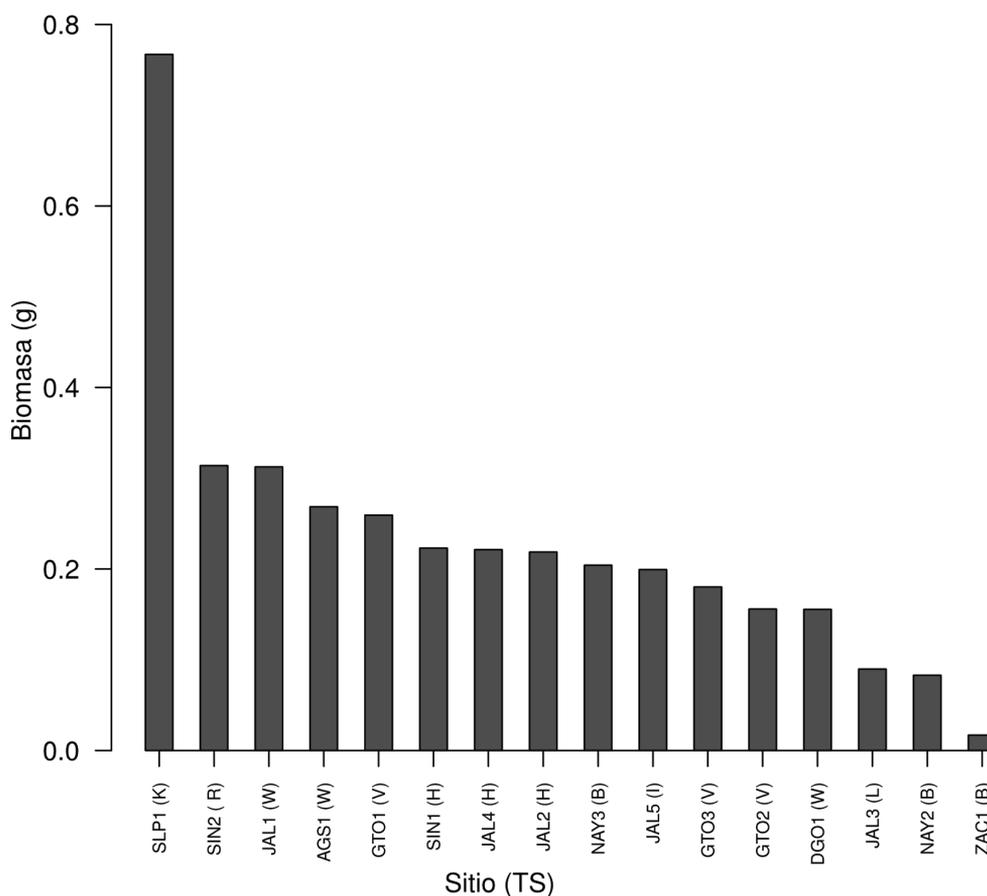
De acuerdo al índice de Lang, los ambientes se pueden clasificar en áridos ( $IA < 40$ ), húmedos ( $40 < IA < 160$ ) e hiperhúmedos ( $IA > 160$ ) (Lang et al., 1915). En este sentido, los suelos con la mayor disponibilidad de agua debido a las condiciones climáticas de los sitios (IA) son JAL3, JAL4, JAL5, NAY2 y NAY3, que corresponden a climas húmedos, mientras que el resto de las muestras de suelos pertenecen a climas considerados áridos bajo este esquema. Con base en estos parámetros, podemos esperar observar diferencias en los microbiomas de los suelos iniciales de climas húmedos contra los de climas áridos, ya que la cantidad de humedad en los suelos es un factor que puede estar determinando a las

comunidades bacterianas de suelos y rizósferas, aunque pueda ser dependiente también de la especie de planta. (Chen et al., 2007; Zhang et al., 2013).

### **Crecimiento y biomasa de las plantas utilizadas en el experimento de jardín común con los múltiples suelos colectados.**

Todas las plantas de *S. lycopersicum* utilizadas para el experimento de jardín común (64 plantas) crecieron satisfactoriamente hasta 60 días post-germinación en todos los suelos utilizados en el estudio. Durante el experimento, las plantas de ZAC1 murieron debido a desecamiento al día 35 del experimento y fueron re-sembradas, lográndose crecer hasta los 60 días también.

#### **Producción de biomasa promedio por tipo de suelo**



**Figura 10. Producción de biomasa promedio de plantas de *S. lycopersicum* por tipo de suelo.** La mayor producción de biomasa sucedió en el suelo Kastañozem SLP1.

Los datos de producción de biomasa (peso seco de las plantas) fueron documentados posteriormente a la toma de muestras de los compartimentos de raíz. En la Figura 10 se presenta un histograma con los datos de biomasa promedio por sitio, donde cada uno fue ordenado de forma decreciente de acuerdo a la producción de biomasa en el eje X. El desarrollo de las plantas varió dependiendo del suelo en el que fueron cultivadas. La mayor biomasa fue alcanzada en las plantas cultivadas en el suelo SLP1 con un promedio de 0.76 g, siendo más del doble del siguiente valor promedio (SIN2; 0.31 g) . Las plantas con el desarrollo más pobre son las correspondientes a los suelos de NAY2 y ZAC1 (Figura 11 A y B) con 0.08 y 0.06 g de peso promedio respectivamente.



**Figura 11. *S. lycopersicum* de 60 días postgerminación.** El suelo SLP1 (izquierda) tuvo las plantas con la mayor producción de biomasa. El suelo NAY2 la segunda menor (derecha).

## Análisis de fisicoquímico de los suelos.

Se determinaron las concentraciones de CT, NT, PT y pH de los suelos en el tiempo inicial y final del experimento. Los resultados de las cuantificaciones comparando el tiempo inicial y final de cada nutriente y pH se muestran en las Figura 12 (A-C).

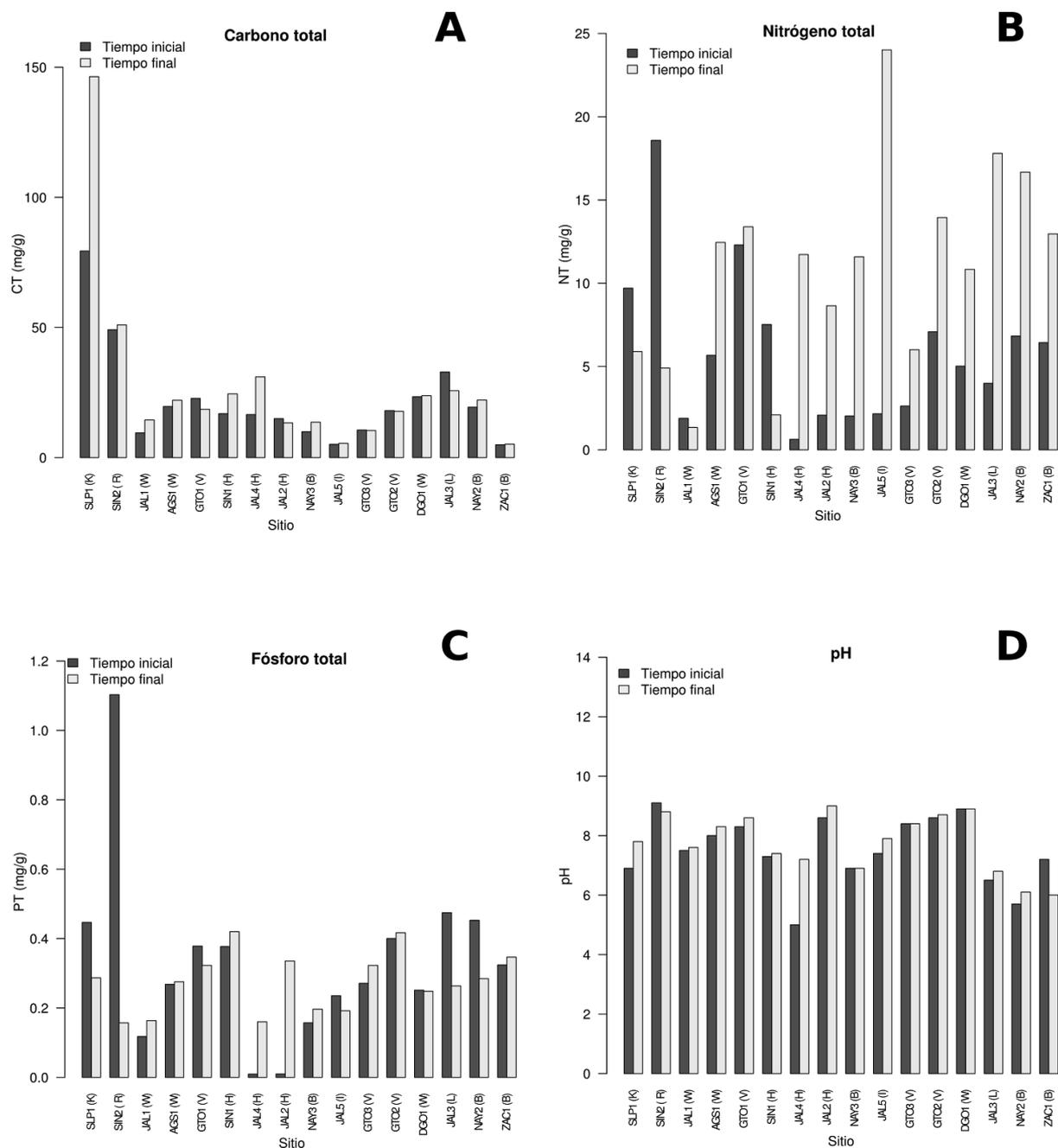
Acerca de la concentración de carbono total (Figura 12-A), se observa un incremento en el tiempo final del experimento en los suelos de los sitios JAL1 , JAL4, JAL5, NAY2, NAY3, SIN1, SIN2, SLP1, AGS1 y DGO 1. El resto de los suelos tuvieron disminuciones ligeras. En cuanto a la concentración de fósforo total (Figura 12-C), se observa un incremento en todas las muestras, salvo en JAL3, JAL5, NAY2, GTO1, SIN2, SLP1 y DGO1. Finalmente, sobre los cambios en la concentración de nitrógeno (Figura 12-B), se puede ver que la cantidad por gramo de suelo incrementó en todas las muestras menos en las muestras JAL1, SIN1, SIN2 y SLP1, las cuales presentan una disminución de concentración al tiempo final.

Sitio	Relación C:N		Relación C:P		Relación N:P	
	T0	Tf	T0	Tf	T0	Tf
JAL1 (W)	5.0	10.8	80.6	88.2	16.1	8.2
JAL2 (H)	7.2	1.5	1528.3	39.6	212.4	25.8
JAL3 (L)	8.2	1.4	69.2	97.5	8.4	67.6
JAL4 (H)	26.5	2.6	1791.3	193.6	67.6	73.2
JAL5 (I)	2.3	0.2	21.5	28.3	9.2	125.2
NAY2 (B)	2.8	1.3	42.7	77.6	15.1	58.6
NAY3 (B)	4.9	1.2	63.1	69.0	12.9	58.9
GTO1 (V)	1.8	1.4	60.1	57.5	32.5	41.5
GTO2 (V)	2.5	1.3	45.1	42.5	17.7	33.5
GTO3 (V)	4.0	1.7	39.0	32.2	9.7	18.6
SIN1 (H)	2.2	11.7	44.7	58.2	19.9	5.0
SIN2 (R)	2.6	10.4	44.5	324.9	16.8	31.3
SLP1 (K)	8.2	24.8	177.7	510.4	21.7	20.6
AGS1 (W)	3.5	1.8	73.1	79.9	21.1	45.2
DGO1 (W)	4.7	2.2	93.0	95.9	20.0	43.7
ZAC1 (B)	0.8	0.4	15.2	14.8	19.9	37.4

**Tabla 5. Relaciones entre las concentraciones de nutrientes en los distintos suelos colectados al tiempo inicial y final del experimento de cultivo de *S. lycopersicum*.**

Al observar las relaciones entre nutrientes (Tabla 5), se encontró que la relación C:N aumenta en las muestras JAL1, SIN1, SIN2 y SLP1, mientras que en las demás muestras esta relación aumenta del tiempo inicial al final. Las relaciones C:P disminuyen en las muestras GTO1 (T0=1.8; TF=1.4), GTO2 (T0=2.5; TF=1.3), GTO3 (T0=4.0; TF=1.7), ZAC1 (T0=0.8; TF=0.4), JAL2 (T0=7.2; TF=1.5) Y JAL4 (T0=26.5; TF=2.6). Las disminuciones de esta relación en los últimos dos sitios mencionados resultan ser muy drásticas debido al considerable aumento de fósforo (Figura 12-C). Las concentraciones de fósforo en el tiempo inicial de estas muestras son muy bajas (<0.01 mg/g), lo cual podría ser resultado de un error en lectura del autoanalizador.

Los incrementos de las relaciones C:N en las muestras JAL1 (T0=5.0; TF=10.8), SIN2 (T0=2.6; TF=10.4) y SLP1 (T0=8.2; TF=24.8) nos muestran que las concentraciones de nitrógeno en los suelos de estos sitios disminuyeron al final del experimento de cultivo de *S. lycopersicum*. Esta tendencia resulta inversa en los demás suelos. Los suelos en los que las concentraciones de carbono total eran bajas y no tuvieron incrementos de este nutriente (JAL2, JAL3, GTO2, GTO3, NAY2, NAY3, ZAC1, DGO1, JAL5) muestran un incremento en la cantidad de nitrógeno. Esto puede ser debido a que cuando hay una baja disponibilidad de carbono orgánico disuelto en el ambiente, las vías de nitrificación bacteriana son favorecidas, aumentando las concentraciones de nitratos en el ambiente (Vitousek, 2002; Strauss & Lamberti 2002). El incremento de nitrógeno en estas muestras nos lleva a especular sobre la presencia y probable enriquecimiento de bacterias fijadoras de nitrógeno como, por ejemplo, bacterias del orden Burkholderiales, que han sido aisladas de la rizósfera de tomates y probadas para llevar a cabo estas funciones y algunas otras de interés agronómico (Caballero-Mellado et al., 2007).



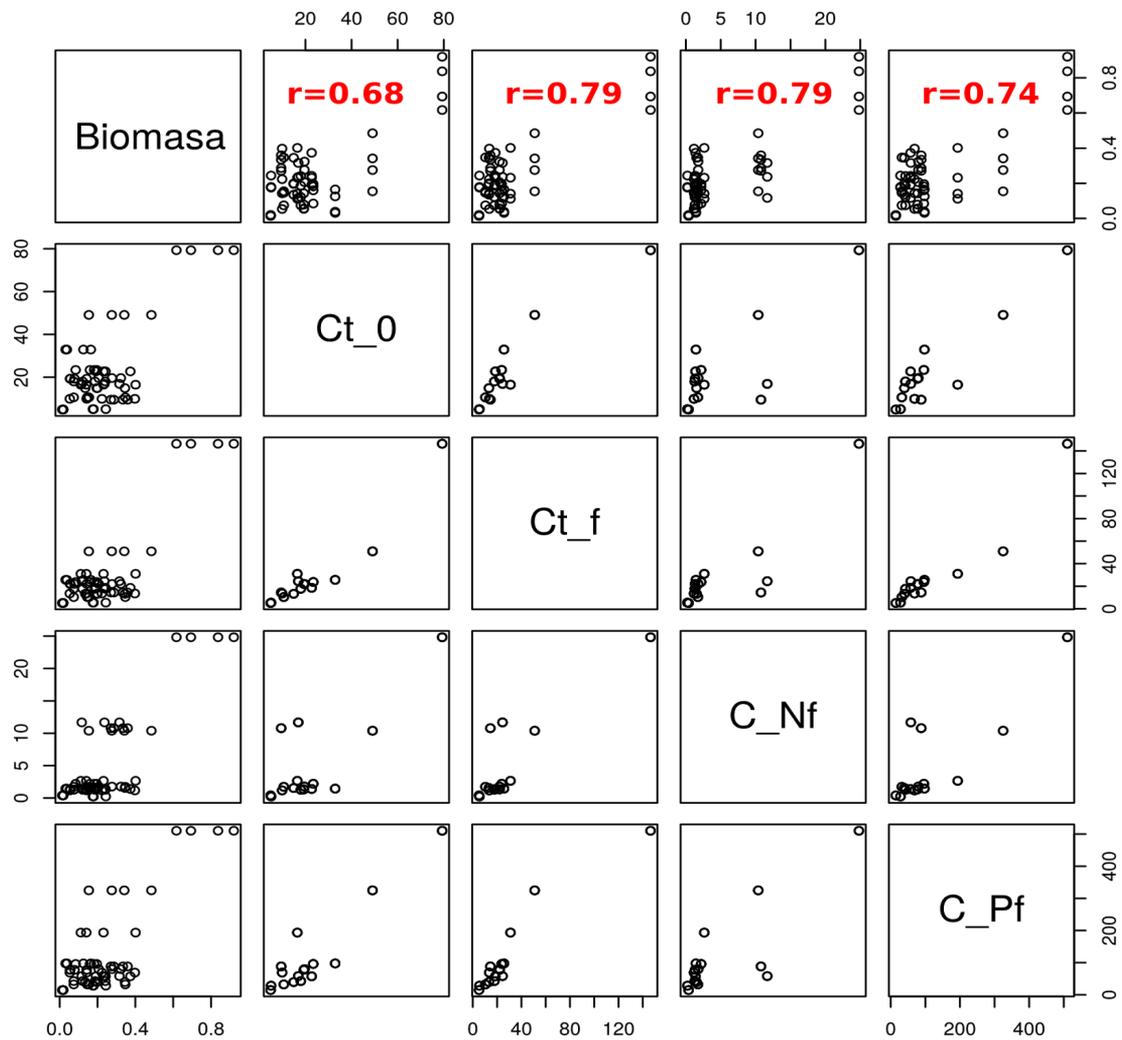
**Figura 12. Análisis fisicoquímicos realizados en los suelos al tiempo inicial y final del cultivo de *S. lycopersicum*.** A) Concentración de carbono total (mg/g suelo). B) Concentración de nitrógeno total (mg/g suelo) C) Concentración de fósforo total (mg/g suelo). D) pH medido en los suelos. Los suelos en el eje X se ordenaron de acuerdo a la producción de biomasa en orden decreciente.

La relación C:N ha sido reportada como un factor importante para los cambios en las estructuras de comunidades microbianas en suelos (Kuramae et al., 2012), por lo que se espera encontrar la mayor cantidad de diferencias entre las

comunidades bacterianas de los suelos iniciales y finales de los sitios JAL1, JAL2, JAL3, JAL4, SIN1, SIN2, y SLP1 debido a que en estos se registró el mayor cambio de esta relación.

En cuanto al pH de los suelos colectados (Figura 12-D), los Vertisoles (GTO1-3), Planosoles (JAL1, AGS1, DGO1), Feozems (JAL2, SIN1), el Litosol (JAL5), Regosol (SIN2) y Cambisol (ZAC1) tuvieron un pH alcalino ( $\text{pH} > 7$ ), mientras que el Litosol (JAL3), Feozem (JAL4), Kastañozem (SLP1) y Cambisoles (NAY2-3) presentaron un pH ácido o cercano al neutro ( $\text{pH} < 7$ ). En el tiempo final del experimento, los suelos del sitio JAL4 y SLP1 pasaron de ácidos a neutros, observándose el cambio más pronunciado en JAL4 que pasó de pH 5 a pH 7.2. El pH ha sido propuesto como uno de los factores abióticos que estructuran fuertemente a las comunidades bacterianas en los suelos (Fierer & Jansson; 2006). Se ha visto que en escalas geográficas grandes (continente americano) esta variable es responsable de explicar una buena parte de la variabilidad de las comunidades microbianas de suelos (Lauber *et al.*, 2009; Mannisto *et al.*, 2007).

Para observar qué factores fisicoquímicos del suelo podrían estar teniendo mayor influencia en la productividad de las plantas de tomate, se realizaron correlaciones múltiples de Pearson con los datos de producción de biomasa, concentración de nutrientes en suelo en el tiempo inicial y final, así como las relaciones estequiométricas de éstos (Figura 13) y se encontró que las correlaciones más altas ( $r > 0.68$ ) con la biomasa de las plantas son: CT en el tiempo inicial ( $r = 0.68$ ), CT en el tiempo final ( $r = 0.79$ ), C:N final ( $r = 0.79$ ) y C:P final ( $r = 0.79$ ). De las correlaciones mostradas, las que tienen un mayor valor de  $r$  con la biomasa son con la concentración de carbono al tiempo final y la relación C:N al tiempo final, lo cual nos sugiere la posible implicación del aporte de carbono por parte de las plantas con mayor biomasa.



**Figura 13. Correlaciones múltiples de Pearson de las variables fisicoquímicas vs. la producción de biomasa promedio.** El carbono al tiempo inicial y final, junto con la relación C:N y C:P finales son las variables más relacionadas a la producción de biomasa.

La Figura 14 muestra un gráfico de Componentes Principales (PC) con los distintos suelos colectados y de las variables fisicoquímicas del tiempo inicial. El primer componente explica el 40.9% de la varianza, mientras que el segundo explica 20.0%. Las variables que están explicando la mayor parte de esta varianza son las concentraciones de nutrientes totales (NT, CT, PT) iniciales. Se puede observar que los Vertisoles (V), Planosoles (W) y Castañozem (K) forman un grupo influenciado por la variable de pH inicial de los suelos. Las muestras de los suelos clasificados como Feozems (H) no forman un grupo discreto, sino que estos suelos presentan la mayor separación por tipo de suelo. Los Cambisoles (B) de los sitios NAY2 y NAY3

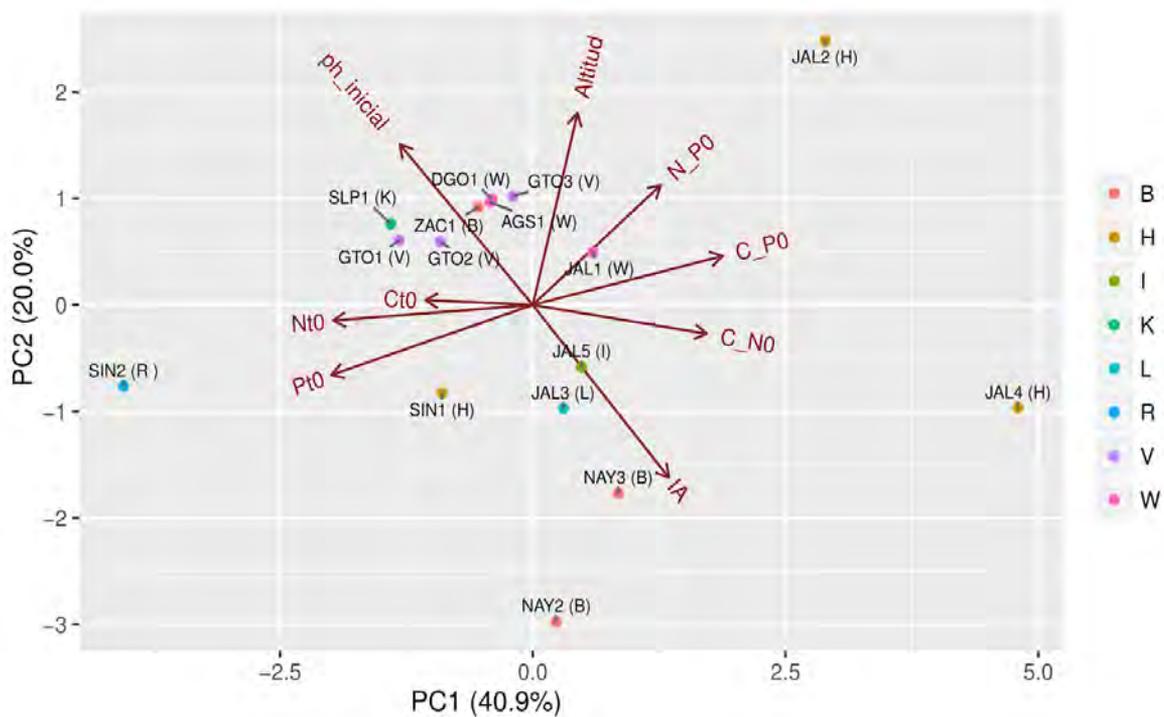
forman otro grupo distinto, guiado por el índice de aridez y no se agrupan con el tercer suelo clasificado como cambisol (ZAC1). Finalmente, la muestra de Regosol (R) se encuentra separada por las variables NT y PT.

Con base en estos resultados se puede esperar que las comunidades bacterianas de los suelos SIN2, NAY2, JAL2 y JAL4 sean las que presenten perfiles filogenéticos más distintos entre sí, como se ha mostrado en distintos estudios en los que se comparan distintos suelos iniciales y terminan estructurando comunidades bacterianas distintas (Winston *et al.*, 2014; Lundberg *et al.*, 2012; Bulgarelli *et al.*, 2015).

El mayor incremento de carbono al tiempo final es en SLP1 (T0=79.35; TF=146.37). En SLP1 también se ve la mayor diferencia entre la tasa C:N del tiempo inicial al final (T0=8.2; TF=24.8), esto nos está hablando del mayor aprovechamiento de las fuentes de nitrógeno y del incremento en el carbono en el suelo al tiempo final. Las plantas crecidas en SLP1 también tuvieron la mayor talla y un sistema radicular más desarrollado, esta condición puede estar favoreciendo una mayor deposición de carbono por parte de la raíz en el medio a comparación de las demás plantas de tomates con menos de la mitad de biomasa producida y raíces menos desarrolladas (Nielson *et al.*, 1994; Bekku *et al.*, 2012). Al observar las diferencias en el pH medido en este suelo, vemos que hay una neutralización, pasando de 6.8 en el tiempo inicial a 7.9 en el tiempo final.

El incremento del pH del medio también podría ser debido a la actividad metabólica de bacterias heterótrofas en SLP1. Se ha visto que muchas bacterias son capaces de precipitar  $\text{CaCO}_3$  en experimentos de laboratorio, esto se ha venido sugiriendo desde hace mucho tiempo (1914) y confirmado en experimentos recientes (Knorre *et al.*, 2000). El rol de las bacterias en la alcalinización, mediante la precipitación de carbonato bacteriano puede ser pasivo y activo. En el modo pasivo de precipitación de carbonatos se puede obtener por múltiples vías (amonificación, reducción de nitratos, degradación de urea/ácido úrico, sulfato reducción). Por otro lado las vías activas de precipitación de carbonatos parecen ser

una respuesta al enriquecimiento de materia orgánica en el medio que terminan dando como resultado un incremento en el pH (Castanier *et al.*, 2000). Por lo que en conjunto con los datos C:P; C:N; N:P de SLP1 pensamos que los cambios son debidos a la actividad metabólica de bacterias en un medio enriquecido por fotosintatos derivados al suelo por parte de la planta, que finalmente adquiere mayor biomasa utilizando efectivamente el N y el P disponibles en SLP1.



**Figura 14. Análisis de componentes principales de los diferentes suelos colectados.** Los puntos de cada muestra se colorean de acuerdo a su tipo de suelo. B=Cambisol, H=Feozem, I=Litosol, K=Castañozem, L=Luvisol, R=Regosol, V=Vertisol, W=Planosol. Las flechas muestran a las variables consideradas para el análisis.

Ahora bien, si el suelo y sus características fisicoquímicas fueran lo que determina el fenotipo de la planta se esperaría que las plantas sembradas en SLP1 y GTO1 fueran muy similares, esto debido a su distancia en el análisis de componentes (Figura 14). Sin embargo, la biomasa de las plantas es contrastante SLP1=0.76 g ; GTO1=0.25 g. La explicación posible a estas diferencias deben de radicar en las distintas comunidades de microorganismos que se están reclutando en la rizósfera de las plantas.

## **Construcción de bibliotecas del gen rRNA 16S**

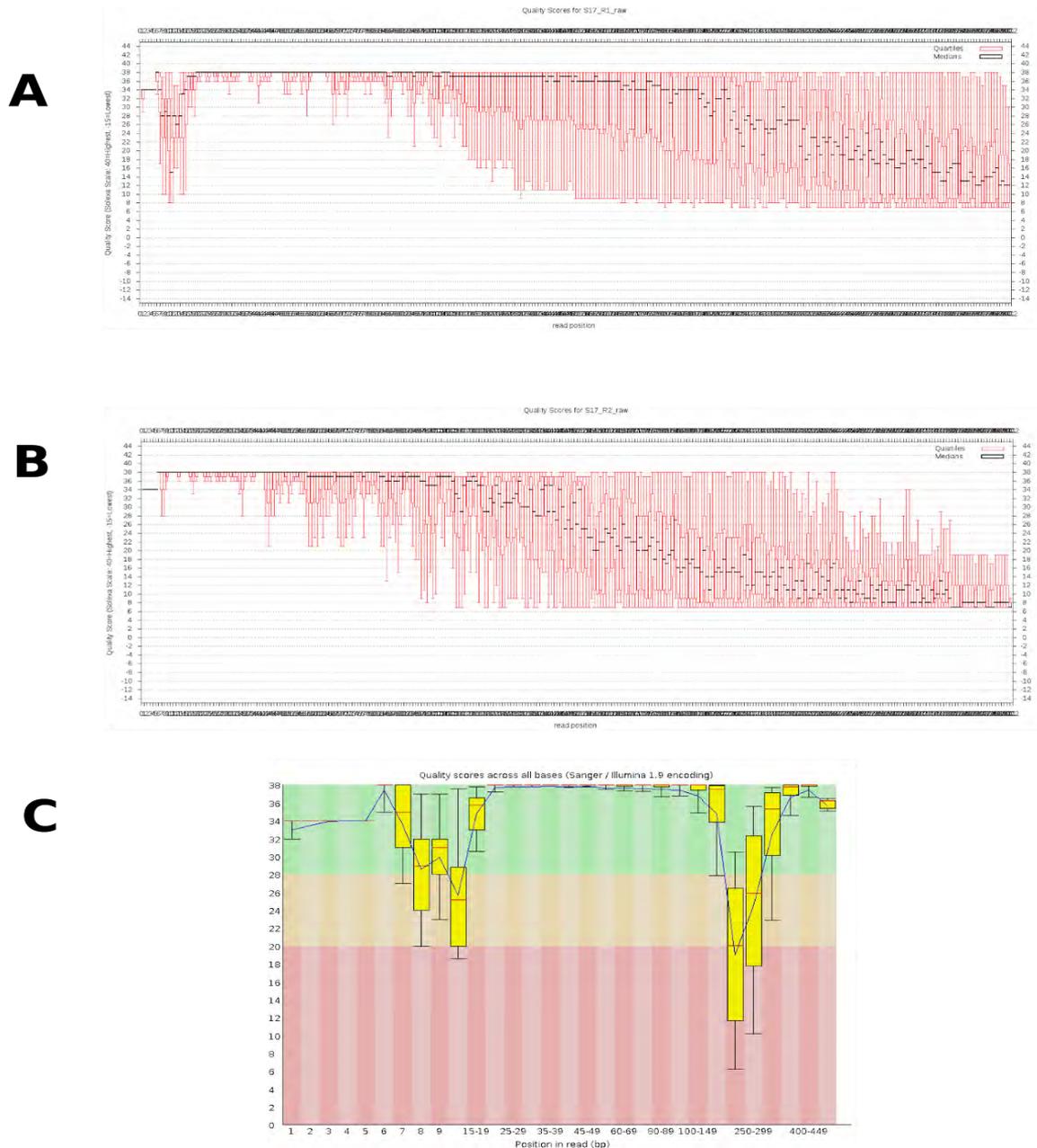
Se prepararon bibliotecas del gen rRNA 16S mediante la amplificación de la región hipervariable V3-V4 de las muestras de S0, SF, RF, EC, CT de las plantas de tomate crecidas en cada uno de los suelos y el caso particular del compartimento RA de una réplica de las plantas de JAL5 (Figura 12 y 13). En este trabajo se muestran los resultados a las muestras correspondientes a JAL5-SF, JAL5-RF, JAL5-EC Y JAL5-RA . Se eligió esta muestra para estandarizar los protocolos de extracción y separación de los componentes del suelo, rizósfera y compartimento endofítico. Todas las muestras (80) se encuentran en estos momentos en la cola del secuenciador MiSeq® en la unidad de secuenciación masiva del IBT, UNAM.

## **Control de calidad de las secuencias del gen rRNA 16S**

En la figura 15 A y B se presenta la calidad de las lecturas crudas de secuencias de la muestra JAL5-SF generadas por secuenciación Illumina MiSeq en configuración 2x300 paired end. En el eje x se presenta la posición de cada base secuenciada (1 a 300) y en el eje y la media y varianza del valor de calidad PHRED por cada posición. Se observa que las lecturas de secuenciación comienzan con calidades arriba de 20 desde el extremo 5' pero a medida que se acerca al 3', la calidad baja a un valor PHRED < 20 después de la posición 250 de las lecturas. Este efecto se ve más pronunciado en las secuencias en sentido reverso (Figura 15-B). Este comportamiento en las lecturas de Illumina ha sido documentado previamente y se ha visto que también ocurre las salidas de otras tecnologías como 454. La acumulación de estos errores hacia el final de las lecturas sucede por el decaimiento de la relación señal/ruido durante el proceso de secuenciación y se ve más pronunciado en las lecturas largas (Knief, 2014).

En la figura 15-C se puede observar el efecto sobre las calidades de secuenciación posterior al ensamblado. El ensamble fue estricto, considerando una

confianza en el llamado de las bases igual o superior al 95% de certidumbre, una longitud mínima para considerar a una secuencia de 250 bp y una máxima longitud de ensamble de 470 bp, considerando un empalme (*overlap*) de 15 pares de bases.



**Figura 15.** En esta gráfica se representa el ensamble logrado por pandaseq. A) Calidad de secuencias crudas en sentido positivo. B) Calidad de secuencias crudas en sentido reverso. C) Las primeras 9 bases corresponden a los adaptadores de secuenciación, alrededor de la base 250 la calidad cae drásticamente, pero al momento de ensamblar la calidad de la lectura más alta es tomada para el ensamblado. Las barras rojas (A y B) y amarillas (C) representan la varianza que tiene picos máximos en la zona del ensamblado.

En la tabla 6 se presentan los datos del número de secuencias crudas y ensambladas. Después del proceso de limpieza y ensamblado, se recuperaron el 17.50% de las secuencias del suelo final (SF), 16.43% de la rizósfera final(RF), 10.8% del compartimento endofítico (EC) y 10.1% de la muestra denominada raíz total (RA=RF+EC). En el proceso de ensamble se pierden miles de secuencias que no cumplen con los criterios mencionados, pero aún así se conservan entre decenas y cientos de miles de secuencias de alta calidad para el análisis. En este sentido, es preferible filtrar de esta forma y quedarse con secuencias que cubren los criterios de calidad ya que de esto dependen todos los análisis subsecuentes.

Compartimento	Secuencias crudas (PE)	No. Secuencias ensambladas	% ensamblado
SF	881504	154305	17.5
RF	753343	123844	16.44
EC	690881	74952	10.85
RA	640129	64651	10.1

**Tabla 6. Eficiencia del proceso de ensamblado de secuencias crudas.** En el peor caso se recuperó el 10% de las secuencias.

### **Diversidad microbiológica y estimación de Unidades Taxonómicas Operativas (OTUs).**

Los OTUs están definidos a un nivel de identidad de secuencia del 97%. En la Tabla 7 se enumeran los descriptores básicos de diversidad: número de especies observadas, el índice de Shannon, índice de Simpson y el estimador no paramétrico Chao1. El índice de Shannon revela la entropía de un sistema biológico estimando el total de especies presentes y su abundancia relativa, a mayores valores, mayor diversidad en términos generales (Hill, 2003). El índice de Simpson sirve para ver que tan dominada esta una comunidad, calcula la probabilidad de que en un

muestreo aleatorio dos individuos pertenezcan a diferentes especies, con valores que van de 0-1, donde 1 representa a comunidades altamente dominadas y 0 a ambientes con una equitatividad entre todos los representantes de una comunidad. Finalmente, Chao1 es un estimador que calcula la frecuencia de especies raras en la muestra y estima el número de especies no detectadas, con lo que puede entenderse como un valor esperado de la diversidad, que al ser comparado con el valor observado puede darnos idea de lo bien realizado del muestreo (Gotelli y Chao 2013).

Muestra	Observada	Shannon	Simpson	Chao1
JAL5-SF	7588	7.75	1.00	7937.15
JAL5RF	6264	6.94	1.00	7414.38
JAL5-EC	4096	6.21	0.99	5499.27
JAL5-RA	4650	6.63	0.99	6410.229

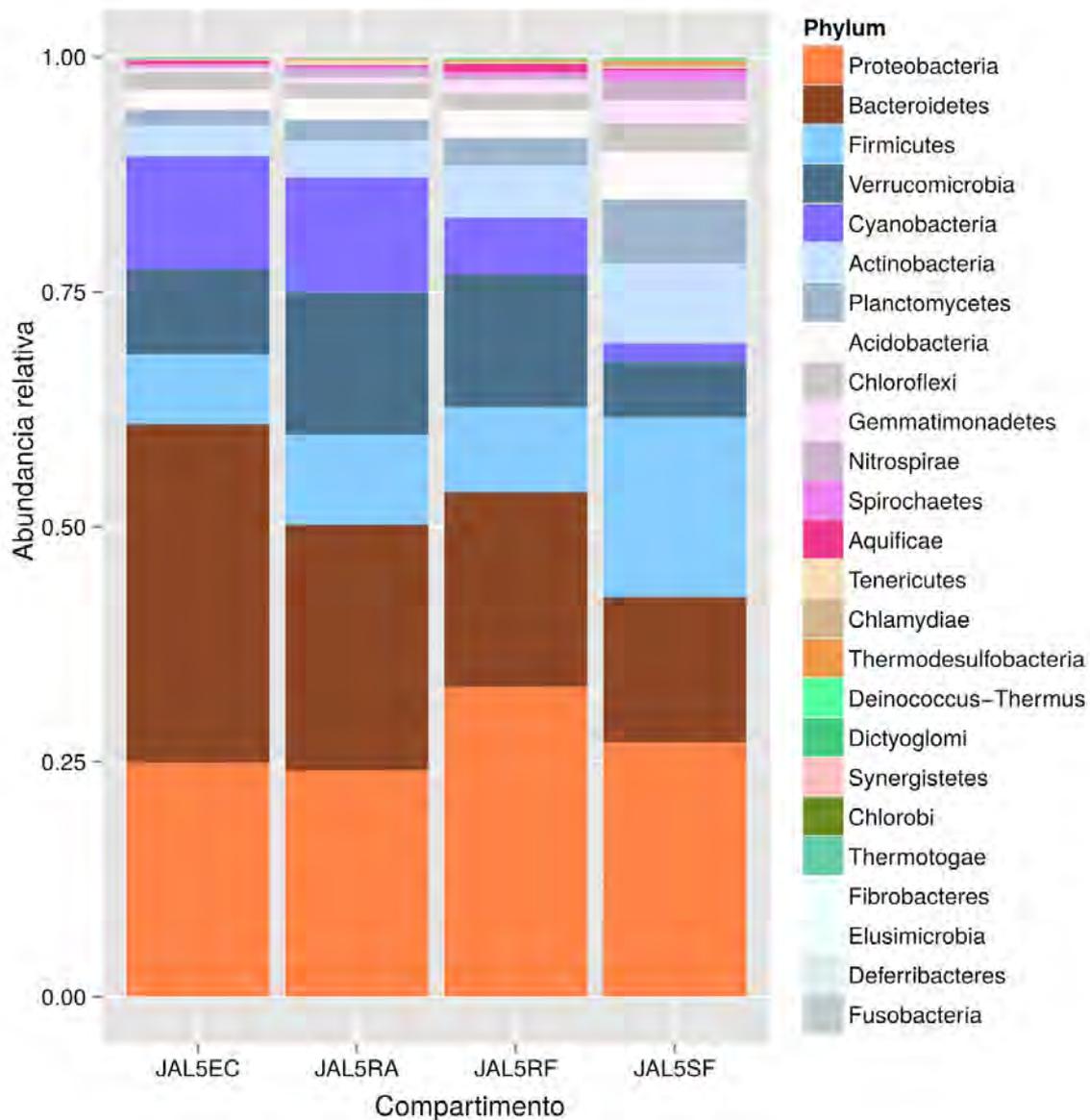
**Tabla 7. Métricas de diversidad alfa (Observada, Shannon, Simpson y Chao1) en los distintos compartimentos de raíz de JAL5.** La mayor diversidad fue obtenida en SF seguido de RF, RA y EC.

En la Tabla 7 se observa que la mayor cantidad de especies esperadas está en el suelo final (SF), disminuye en la rizósfera (RF) y se vuelve más específica en el compartimento endofítico (EC). La muestra RA comprende un proceso de extracción que mezclaba la separación del EC y RF, a través de pasos de sonicación y lavado, era esperado tener mayor diversidad que EC, pero a niveles comparables de RF. La comparación entre el número de especies observadas y las esperadas nos indica que la estrategia de extracción de DNA metagenómico, la amplificación, construcción de bibliotecas y cobertura de secuenciación fue adecuada ya que hay poca diferencia en general entre el total de especies observadas y esperadas. El índice de Shannon nos indica que estas comunidades son muy diversas, en macroorganismos se consideran ambientes muy diversos con valores de 3-4 en Shannon (Gliessmon *et al.*, 1988). En el caso del índice de

Simpson nos indica que a pesar de la gran diversidad, se trata de comunidades altamente dominadas por unas cuantas especies. Comparando estos resultados con los que se obtuvieron en *Arabidopsis* (Lundberg *et al.*, 2012) tenemos valores comparables, aunque JAL5 es más diverso que *Arabidopsis*: SF(Ath)=6.0, SF(JAL5)=7.75; RF(Ath)=5.7, RF(JAL5)=6.94 y; EC(Ath)=3.4; EC(JAL5)=6.21.

### **Perfiles filogenéticos de los compartimentos de raíz en JAL5.**

Se construyeron los perfiles filogenéticos a nivel de *Phylum* de cada compartimento de las raíces de tomates crecidos en el suelo JAL5 (Figura 16). Los grupos más abundantes (>1%) en todos los compartimentos corresponden a los phyla *Proteobacteria*, *Bacteroidetes*, *Firmicutes*, *Verrucomicrobia*, *Cyanobacteria*, *Actinobacteria*, *Planctomycetes* y *Acidobacteria*. A partir del suelo final (JAL5SF) se observan cambios en las abundancias relativas de cada phyla conforme se pasa a compartimentos específicos. El grupo de las *Proteobacterias* se enriquece en el compartimento de RF (de 27% a 33%) y disminuye ligeramente en EC y RA (24.9% y 24.1%). Los *Bacteroidetes* (15.5% en SF) se enriquecen en los compartimentos RF, RA y EC (20.7%, 26.1% y 36.1%). El patrón observado para los *Bacteroidetes* es inverso para el caso de los *Firmicutes* (19.1% en SF). La abundancia relativa de este grupo disminuye en los compartimentos RF, RA y EC (9.0%, 9.7% y 7.4%). La abundancia relativa del Phylum *Verrucomicrobia* es más del doble en los compartimentos RF y RA (14.1% y 15.15) a comparación de SF y EC (6.0% y 9.1%). Existe una tendencia similar a los *Bacteroidetes*, para los casos de *Actinobacteria* (SF=8.5%, RF=5.5%, RA=3.9%, EC=3.3%) y *Acidobacteria* (SF=5.1%, RF=2.9%, RA=2.2%, EC=2.1%). En el caso de las *Cyanobacterias*, existe un incremento en su abundancia en los compartimentos de raíz (SF=1.9%, RF=6.1%, RA=12.2%, EC=12.0%). Cabe mencionar que los datos no fueron filtrados de las secuencias de cloroplastos, lo que puede estar alterando estos conteos.



**Figura 16. Abundancia relativa a nivel de phylum en cada compartimento de raíz de JAL5.**

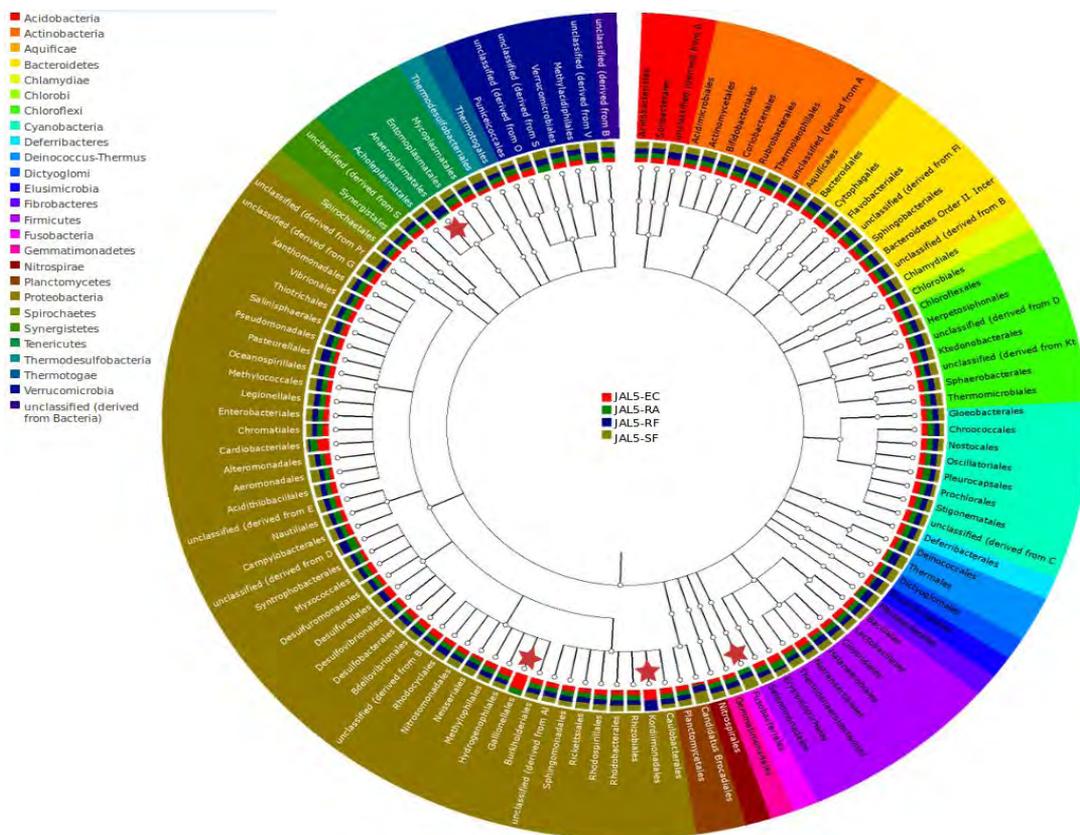
Adicionalmente, se analizaron los perfiles filogenéticos de los compartimentos, a nivel de orden. Se graficó una filogenia a partir de un árbol de referencia (Figura 17), donde se presenta la abundancia de cada filotipo por muestra con un histograma en los nodos terminales del árbol. Se puede observar que a nivel de Orden, estos se encuentran distribuidos en todos los compartimentos con diferencias en la abundancia relativa. Los grupos taxonómicos que no son ubicuos en todos los compartimentos corresponden a: 1) *Aeromonadales*, presentes

en SF, RF y RA; 2) *Gallionelales*, presentes solo en SF y EC; 3) *Fusobacteriales*, ausentes en EC y 4) *Kordiimonadales*, presentes solo en RF y EC.

Los *Anaeroplasmatales* son un orden perteneciente a los Tenericutes, dentro de este grupo se contienen a los géneros *Anaeroplasma* y *Asteroplasma*. El género *Anaeroplasma* cuenta con solamente 4 especies aisladas: *A. bactoclasticum*, *A. abactoclasticum*, *A. intermedium* y *A. varium* (Robinson & Freundt, 1987). Estos son organismos anaerobios obligados que fueron aislados de rumen de bovinos. Las diferencias entre los *Anaeroplasma* y *Asteroplasma* radican en el requerimiento de esteroides de *Anaeroplasma* para su crecimiento y en características antigénicas (Robinson & Freundt, 1987). En la muestra JAL5 fueron encontrados en suelo y rizósferas, pero no en compartimento endofítico. Se ha encontrado la presencia de esteroides en suelos y se ha visto que el tipo de esteroides presentes pueden ser indicadores de diversidad en flora y fauna (Metter *et al.*, 1970) y que además son factores estructurantes de comunidades microbianas en suelos (Guedes *et al.*, 2011).

En el orden de los *Kordiimonadales* contiene al género *Kordiimonas* con una sola especie cultivada, *Kordiimonas gwangyangensis*, aislada de sedimentos marinos (Kwon *et al.*, 2005). Es una bacteria capaz de degradar hidrocarburos aromáticos policíclicos, móvil, Gram negativa y con un contenido relativamente bajo de GC (39.3%) y un genoma secuenciado de ~4.1 Mb. Este filotipo fue identificado sólo en los compartimentos de EC y RF, lo cual puede sugerir a esta bacteria como un endófito de raíz. Existe una relación entre microorganismos endosimbiontes y el contenido bajo de GC en sus genomas, la mayoría de ellos tienen un contenido de GC < 40% (McCutcheon *et al.*, 2009), lo cual podría ser indicativo de que esta bacteria sea, efectivamente un endófito, aunque el tamaño del genoma de esta bacteria es grande para ser considerado un simbiote obligado, puede ser una especie en proceso de selección para el ámbito de la rizósfera.

Dentro del orden *Gallionales* se tiene a una sola especie cultivada del género *Gallionella*: *G. ferruginea*, siendo una bacteria quimiolitótrofa capaz de oxidar y acumular hierro, anaerobia facultativa y ha sido aislada de cuerpos de agua fresca, manantiales con minerales ferrosos y ventilas hidrotermales (Halbach *et al.*, 2001; Anderson & Pedersen, 2003). La presencia de este filotipo en SF y EC podría explicarse dado el origen de la muestra que fue tomada a un costado del volcán del Ceboruco en Jalisco, lo cual puede ser indicativo de un suelo con alto contenido de hierro. La capacidad de las bacterias de conseguir metales limitantes como el hierro puede ser clave para ser microorganismos que puedan ser seleccionados para aportar estos nutrientes a la planta, el hierro es un gran limitante de los sistemas biológicos y es clave para el establecimiento de las relaciones tanto de patógeno y simbiote con el hospedero (Pii *et al.*, 2015 ).

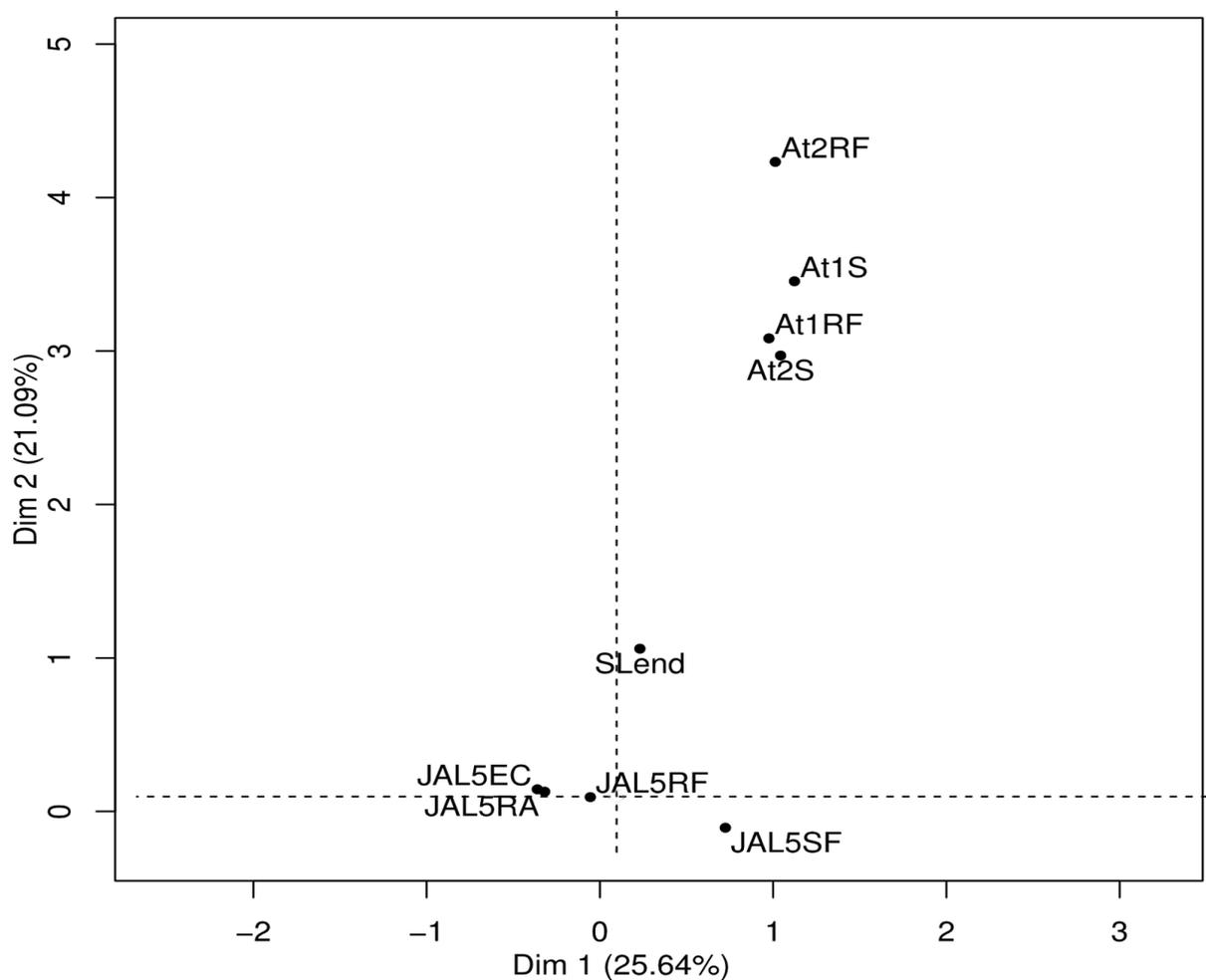


**Figura 17. Distribución taxonómica a nivel de orden en los distintos compartimentos de raíz de JAL5.** Se presentan barras superpuestas con la abundancia relativa de cada taxa por compartimento al final de cada nodo del árbol. La clave de color en la esquina superior izquierda corresponde al *phylum* de cada grupo. Las estrellas rojas marcan a los grupos que no son ubicuos en todos los compartimentos.

Para comprender la distancia y la diversidad entre las distintas muestras de JAL5 se hizo un análisis de diversidad  $\beta$  comparando las distancias entre distintos sistemas. Para tal efecto, comparamos los microbiomas de los distintos compartimentos (RF, EC, RA y SF) de la raíz de tomate crecida en el suelo del sitio JAL5 contra los microbiomas de rizósferas y suelos de *A. thaliana* (Lundberg *et al.*, 2012) y los de endófitos de tomates sanos de un estudio de inoculación con nemátodos (Tian *et al.*, 2015) por medio del análisis de correspondencia de sus perfiles filogenéticos (Figura 18). En el análisis de correspondencias explica un total de 46.6% de la varianza total. En la Figura 16 se puede observar que las muestras de rizósfera y suelos de *A. thaliana* se agrupan en la parte superior del gráfico, mientras que las muestras de raíz de JAL5 se agrupan en la parte inferior. El microbioma de endófitos de raíces de tomate se encuentra más cercano a las muestras de los tomates cultivados en el suelo JAL5 que a las de *A. thaliana*.

Estos resultados están de acuerdo con lo descrito previamente por Bulgarelli *et al.*, (2015 y 2012), Schlaeppli (2014) y Turner (2013) los cuales han realizado trabajos comparando microbiomas de rizósferas de *Arabidopsis*, cebada, trigo y avena comparando diferentes géneros, especies y genotipos de plantas. En este caso se compararon dos especies y adicionalmente dos variedades de plantas.

La separación de los compartimentos de las raíces de JAL5 que se observa en la Figura 18 muestra que la rizósfera (JAL5-RF) es un punto intermedio entre el compartimento endofítico (JAL5\_EC) y el suelo circundante (JAL5-SF). El compartimento de raíz completa (JAL5-RA) se encuentra agrupado muy cercanamente con el compartimento endofítico. Esto muestra que nuestro método para extraer el DNA metagenómico de las raíces para enriquecer a microorganismos que puedan habitar en los espacios inter o intracelulares de las plantas fue efectivo.



**Figura 18.** Análisis de correspondencias de los microbiomas de compartimentos de raíz de tomate y suelo del sitio JAL5 (JAL5EC, JAL5RF, JAL5SF, JAL5RA) contra los microbiomas de rizósfera y suelo de *A. thaliana* (AtRF y AtS) y microbioma de endófitos de tomates (Slend).

### **Predicción metabólica en los compartimentos de raíz de JAL5.**

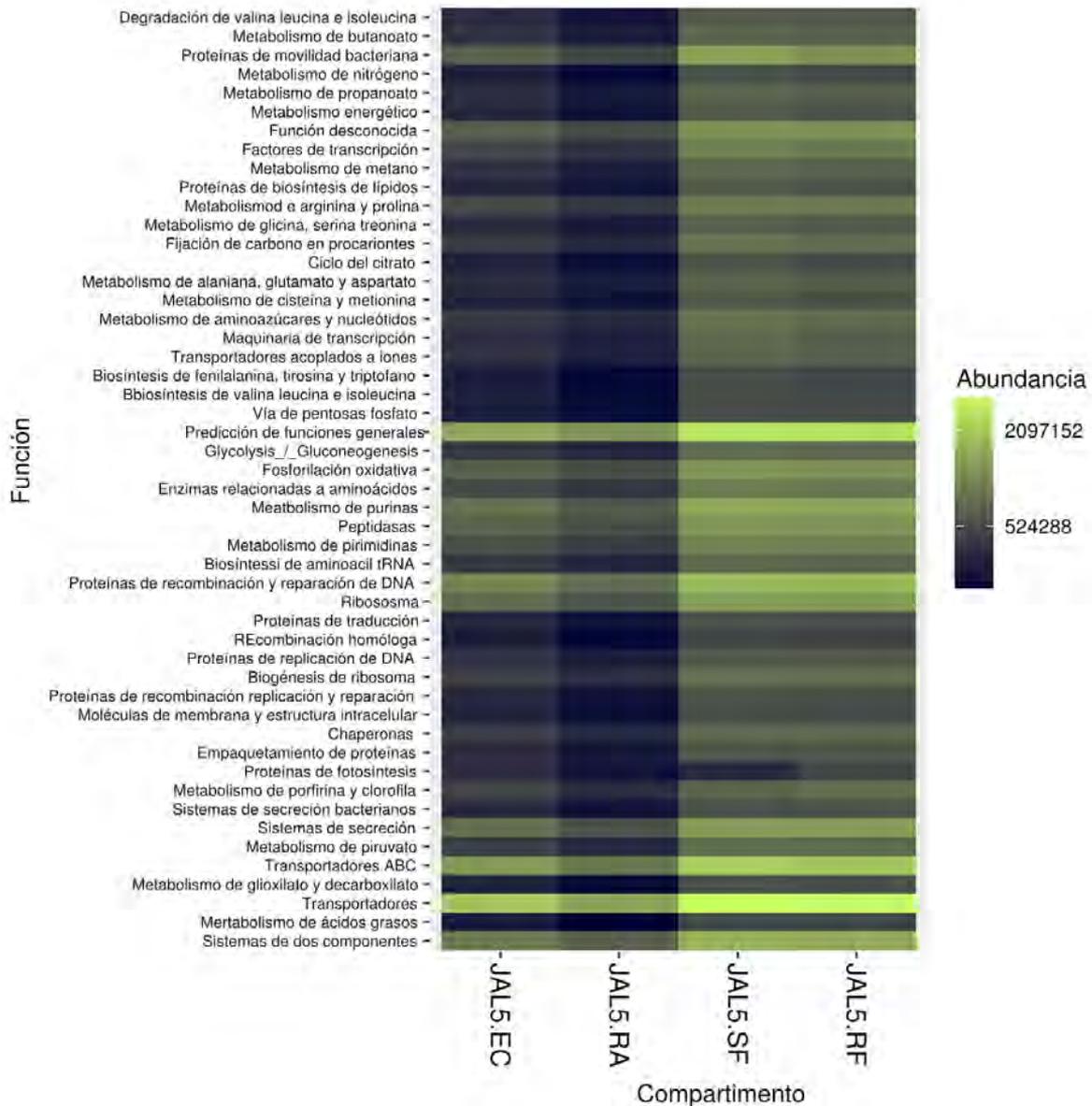
Con base en los perfiles filogenéticos de cada una de las muestras de JAL5 se realizaron las predicciones metabólicas. Para las predicciones metabólicas se utilizó la anotación del tercer nivel jerárquico de clasificación de la base de datos KEGG. Los datos se transformaron ( $\log_2$ ) y las 50 funciones que resultaron más abundantes se visualizaron como un gráfico de calor (Figura 19). La abundancia de cada función se representa mediante una clave de color, donde los colores claros

indican sobre-representación de las categorías y los colores oscuros significan baja abundancia. La primera y segunda columna corresponden a las muestras de JAL5-EC Y JAL5-RA respectivamente; la segunda y tercera son de los compartimentos JAL5-SF y JAL5-RF.

Las categorías de funciones metabólicas que se encuentran sobre-representadas en todos los casos son: transportadores, transportadores ABC y predicción de funciones generales. Las categorías medianamente representadas son: funciones de sistemas de secreción, biosíntesis de ribosomas, proteínas de reparación y recombinación de DNA, funciones desconocidas, proteínas de movilidad bacteriana, sistemas de dos componentes, metabolismo de purinas y proteínas de recombinación. En las interacciones planta bacteria, las plantas inician una comunicación con las comunidades bacterianas mediante el aporte de moléculas que pueden ser reconocidas por los microorganismos y al final del día generan fenómenos de colonización en el huésped. La capacidad de movilidad en los microorganismos resulta ser una característica importante para estas interacciones (Lugtenberg *et al.*, 2002). De igual manera la categoría de transportadores es la más representada en en todos los compartimentos. Se ha visto que algunos miembros de la familia de autotransportadores tienen funciones como adhesinas (Henderson *et al.*, 1998) y se han propuesto como factores de virulencia en patógenos. La maquinaria de sistemas de secreción tipo III, codificados por los genes *hrp* también son una forma importante de reconocimiento entre plantas y patógenos (Buttner & He, 2009).

De manera general se observa una distribución similar en las funciones metabólicas predichas para el compartimento endofítico (EC) y las rizósferas en conjunto con el compartimento endofítico (RA) . Las muestras EC y RA están menos enriquecidos en todas las funciones cuando se comparan con las columnas SF y RF. En términos generales, cuando se estudia la diversidad en suelo y rizósfera se espera que el suelo tenga una mayor diversidad en términos absolutos, dicha diversidad se ve reflejada en número de especies y esto debe de verse reflejado también en mayor diversidad de funciones metabólicas. El resultado es prometedor,

ya que nos demuestra que la metodología montada durante el transcurso de esta tesis separa efectivamente a diversas comunidades bacterianas y nos permitirá tener perfiles taxonómicos confiables con la secuenciación de las 80 muestras faltantes que se encuentran en proceso. La comparación de las muestras de amplicones del gen ribosomal 16S nos permitirá evaluar qué suelos son más diversos y tener valores esperados de los metagenomas que están en proceso de secuenciación en estos momentos.



**Figura 19. Gráfico de calor de la predicción funcional contra la base de datos KEGG en los compartimentos de raíz de tomate y suelo final de JAL5.** Cada columna representa un compartimento (EC=endofítico; RA=raíz; SF=suelo final; RF=rizósfera). Cada columna es una categoría funcional. Los colores claros indican sobrerrepresentación, los oscuros baja abundancia.

## Conclusiones

En este trabajo se lograron coleccionar 16 muestras de suelos correspondientes a 8 distintas unidades de suelos representativos del territorio mexicano, de acuerdo a los datos edafológicos. Por medio de las comparaciones de las características fisicoquímicas pudimos observar que no todos los tipos de suelo se agrupan dependiendo de su clasificación, en particular las muestras de Cambisoles se observan más distantes entre sí. La descripción de estos parámetros nos ayuda a establecer expectativas sobre los hallazgos que se puedan realizar sobre los microbiomas de cada suelo.

A su vez algunos de los parámetros fisicoquímicos son variables que sirven para explicar las diferencias observadas en el desarrollo de plantas de tomate, exitosamente cultivadas en cada tipo de suelo bajo las mismas condiciones. Al parecer existe una relación entre la producción de biomasa de las plantas y las concentraciones de nutrientes en los suelos. En particular las concentraciones de Carbono y Nitrógeno así como las relaciones entre estos dos nutrientes. El hecho de que las plantas del suelo SLP1 hayan tenido la mayor producción de biomasa nos lleva a pensar en que las comunidades bacterianas de ese suelo puedan tener un alto potencial funcional para la promoción de crecimiento vegetal.

Los métodos empleados para la separación de los microhábitats de rizósfera mostraron ser lo suficientemente efectivos para poder discriminar a cada compartimento y al ponerlos en contexto junto con sistemas de rizósferas de diferentes especies y variedades podemos justificar estos resultados, los cuales comprueban las hipótesis de la selección de microbiomas específicos está también influenciada por características intrínsecas de las plantas, como su especie o genotipo.

Las métricas de diversidad calculadas para cada compartimento de raíz de las plantas crecidas en JAL5 han comprobado las teorías sobre la reducción de

diversidad bacteriana desde el suelo hasta la rizósfera y compartimentos endofíticos, pero en general este es un sistema muy diverso.

Las predicciones metabólicas realizadas a partir de los perfiles filogenéticos nos resultan útiles para tener una aproximación a las probables funciones que las bacterias están llevando a cabo en ambientes complejos como el suelo y en particular la rizósfera.

## **Perspectivas**

En este trabajo se obtuvo el DNA metagenómico de rizósferas y compartimentos endofíticos de plantas de tomate crecidas en 8 distintos tipos de suelo colectados en 16 ubicaciones geográficas distintas. Además se obtuvo el DNA metagenómico de los suelos al tiempo de colecta (S0), de los suelos después de haber crecido las plantas (SF) y de suelos control sin plantas (CT). A partir de estas muestras, se generaron bibliotecas de amplicones del gen 16S rRNA, sumando un total de 80 bibliotecas (Tabla 8).

Las bibliotecas de amplicones se encuentran en proceso de secuenciación, por lo cual se propone continuar con este trabajo para poder realizar la caracterización completa de los microbiomas de raíces de tomate crecidos en cada tipo de suelo y poder llevar a cabo comparaciones y análisis más detallados tanto de los compartimentos como de los suelos. La continuación de este trabajo es de gran relevancia, ya que nos puede ayudar a resolver preguntas sobre la biogeografía de las bacterias, la composición de un microbioma núcleo en la raíz de *S. lycopersicum* y la relevancia de los factores bióticos y abióticos de los suelos como estructuradores de los microbiomas en los ambientes mencionados.

Adicionalmente, se plantea llevar a cabo la secuenciación de metagenomas totales de las muestras de suelos que resulten tener los perfiles filogenéticos con mayores diferencias entre sí para obtener información funcional con el objetivo de: l) Elucidar si existe un núcleo funcional operando en la rizósfera del tomate, no

importando el origen del microbioma y II) realizar comparaciones contra las predicciones funcionales realizadas en este trabajo.

1 020316 MiSeq 16S Suelo inicial JAL1S0		17 020316 MiSeq 16S Suelo f nal JAL1SF		33 020316 MiSeq 16S Rizosfera JAL1RF		49 020316 MiSeq 16S Endófitos JAL1EC		65 020316 MiSeq 16S Control JAL1CTRL
2 020316 MiSeq 16S Suelo inicial JAL2S0		18 020316 MiSeq 16S Suelo f nal JAL2SF		34 020316 MiSeq 16S Rizosfera JAL2RF		50 020316 MiSeq 16S Endófitos JAL2EC		66 020316 MiSeq 16S Control JAL2CTRL
3 020316 MiSeq 16S Suelo inicial JAL3S0		19 020316 MiSeq 16S Suelo f nal JAL3SF		35 020316 MiSeq 16S Rizosfera JAL3RF		51 020316 MiSeq 16S Endófitos JAL3EC		67 020316 MiSeq 16S Control JAL3CTRL
4 020316 MiSeq 16S Suelo inicial JAL4S0		20 020316 MiSeq 16S Suelo f nal JAL4SF		36 020316 MiSeq 16S Rizosfera JAL4RF		52 020316 MiSeq 16S Endófitos JAL4EC		68 020316 MiSeq 16S Control JAL4CTRL
5 020316 MiSeq 16S Suelo inicial JAL5S0		21 020316 MiSeq 16S Suelo f nal JAL5SF		37 020316 MiSeq 16S Rizosfera JAL5RF		53 020316 MiSeq 16S Endófitos JAL5EC		69 020316 MiSeq 16S Control JAL5CTRL
6 020316 MiSeq 16S Suelo inicial NAY2S0		22 020316 MiSeq 16S Suelo f nal NAY2SF		38 020316 MiSeq 16S Rizosfera NAY2RF		54 020316 MiSeq 16S Endófitos NAY2EC		70 020316 MiSeq 16S Control NAY2CTRL
7 020316 MiSeq 16S Suelo inicial NAY3S0		23 020316 MiSeq 16S Suelo f nal NAY3SF		39 020316 MiSeq 16S Rizosfera NAY3RF		55 020316 MiSeq 16S Endófitos NAY3EC		71 020316 MiSeq 16S Control NAY3CTRL
8 020316 MiSeq 16S Suelo inicial GTO1S0		24 020316 MiSeq 16S Suelo f nal GTO1SF		40 020316 MiSeq 16S Rizosfera GTO1RF		56 020316 MiSeq 16S Endófitos GTO1EC		72 020316 MiSeq 16S Control GTO1CTRL
9 020316 MiSeq 16S Suelo inicial GTO2S0		25 020316 MiSeq 16S Suelo f nal GTO2SF		41 020316 MiSeq 16S Rizosfera GTO2RF		57 020316 MiSeq 16S Endófitos GTO2EC		73 020316 MiSeq 16S Control GTO2CTRL
10 020316 MiSeq 16S Suelo inicial GTO3S0		26 020316 MiSeq 16S Suelo f nal GTO3SF		42 020316 MiSeq 16S Rizosfera GTO3RF		58 020316 MiSeq 16S Endófitos GTO3EC		74 020316 MiSeq 16S Control GTO3CTRL
11 020316 MiSeq 16S Suelo inicial SIN1S0		27 020316 MiSeq 16S Suelo f nal SIN1SF		43 020316 MiSeq 16S Rizosfera SIN1RF		59 020316 MiSeq 16S Endófitos SIN1EC		75 020316 MiSeq 16S Control SIN1CTRL
12 020316 MiSeq 16S Suelo inicial SIN2S0		28 020316 MiSeq 16S Suelo f nal SIN2SF		44 020316 MiSeq 16S Rizosfera SIN2RF		60 020316 MiSeq 16S Endófitos SIN2EC		76 020316 MiSeq 16S Control SIN2CTRL
13 020316 MiSeq 16S Suelo inicial SLP1S0		29 020316 MiSeq 16S Suelo f nal SLP1SF		45 020316 MiSeq 16S Rizosfera SLP1RF		61 020316 MiSeq 16S Endófitos SLP1EC		77 020316 MiSeq 16S Control SLP1CTRL
14 020316 MiSeq 16S Suelo inicial AGS1S0		30 020316 MiSeq 16S Suelo f nal AGS1SF		46 020316 MiSeq 16S Rizosfera AGS1RF		62 020316 MiSeq 16S Endófitos AGS1EC		78 020316 MiSeq 16S Control AGS1CTRL
15 020316 MiSeq 16S Suelo inicial DGO1S0		31 020316 MiSeq 16S Suelo f nal DGO1SF		47 020316 MiSeq 16S Rizosfera DGO1RF		63 020316 MiSeq 16S Endófitos DGO1EC		79 020316 MiSeq 16S Control DGO1CTRL
16 020316 MiSeq 16S Suelo inicial ZAC1S0		32 020316 MiSeq 16S Suelo f nal ZAC1SF		48 020316 MiSeq 16S Rizosfera ZAC1RF		64 020316 MiSeq 16S Endófitos ZAC1EC		80 020316 MiSeq 16S Control ZAC1CTRL

**Tabla 8. Claves de muestras enviadas a la Unidad de Secuenciación del Instituto de Biotecnología, UNAM.** Se muestra la clave utilizada por la unidad de secuenciación, la fecha de entrega, el tipo de muestra y la clave utilizada en este trabajo.

## Bibliografía

- Alcaraz, L. D., Martínez-Sánchez, S., Torres, I., Ibarra-Laclette, E., & Herrera-Estrella, L. (2016). The Metagenome of *Utricularia gibba*'s Traps: Into the Microbial Input to a Carnivorous Plant. *PLoS ONE*, 11(2), e0148979.
- Anderson, C.R. & Pedersen, K., 2003. In situ growth of *Gallionella* biofilms and partitioning of lanthanides and actinides between biological material and ferric oxyhydroxides. *Geobiology*, 1(2), pp.169–178.
- Bais, H.P. et al., 2006. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annual review of plant biology*, 57, pp.233–66.
- Barea, J. M., Navarro, E. and Montoya, E. (1976), Production of Plant Growth Regulators by Rhizosphere Phosphate-solubilizing Bacteria. *Journal of Applied Bacteriology*, 40: 129–134
- Bekku, Y. et al., 1997. Carbon input from plant to soil through root exudation in *Digitaria adscendens* and *Ambrosia artemisiifolia*. *Ecological Research*, 12, pp.305–312.
- Berendsen, R., Pieterse, C. & Bakker, P., 2012. The rhizosphere microbiome and plant health. *Trends in plant science*, 17(8), pp.478–86.
- Bulgarelli, D. et al., 2012. Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature*, 488(7409), pp.91–5.
- Bulgarelli, D. et al., 2015. Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley. *Cell Host & Microbe*, pp.392–403.
- Büttner, D. & He, S.Y., 2009. Type III protein secretion in plant pathogenic bacteria. *Plant physiology*, 150(4), pp.1656–1664.
- Caballero-Mellado, J. et al., 2007. The Tomato Rhizosphere, an Environment Rich in Nitrogen-Fixing Burkholderia Species with Capabilities of Interest for Agriculture and Bioremediation. *Applied and Environmental Microbiology*, 73(16), pp.5308–5319.
- Caporaso, J., Kuczynski, J. & Stombaugh, J., 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature ...*, 7(5), pp.335–336.
- Caporaso, J.G. et al., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME journal*, 6(8), pp.1621–4.
- Castanier S, Métayer-Levrel G Le, Perthuisot J-P. Microbial Sediments. In: Riding RE, Awramik SM, editors. Berlin, Heidelberg: Springer Berlin Heidelberg; 2000. pp. 32–39. doi:10.1007/978-3-662-04036-2\_5
- Chelius, M.K. & Triplett, E.W., 2001. The Diversity of Archaea and Bacteria in Association with the Roots of *Zea mays* L. *Microbial ecology*, 41(3), pp.252–263.
- Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), (1995). [http://www.conabio.gob.mx/informacion/metadatos/gis/eda251mgw.xml?\\_httpcache=yes&\\_xsl=/db/metadatos/xsl/fgdc\\_html.xsl&\\_indent=no](http://www.conabio.gob.mx/informacion/metadatos/gis/eda251mgw.xml?_httpcache=yes&_xsl=/db/metadatos/xsl/fgdc_html.xsl&_indent=no)
- Daniel, R., 2005. The metagenomics of soil. *Nature reviews. Microbiology*, 3(6), pp.470–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15931165>

DeSantis, T.Z. et al., 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, 72(7), pp.5069–5072.

Edwards, J. et al., 2015. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proceedings of the National Academy of Sciences*, p.201414592.

FAO-ISRIC-ITC. Lecture notes of the mayor soil of the from Sampling Data,” in Encyclopedia of Biodiversity, vol. 5, pp. 195–211, Elsevier Ltd., 2013.

Fierer, N. & Jackson, R.B., 2006. The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, 103(3), pp.626–631.

Fulthorpe, R.R., A.N. Rhodes and J.M. Tiedje. 1998. High levels of endemicity apparent in 3-chlorobenzoate degrading soil bacteria. *Applied and Environmental Microbiology*. 64:1620-1627.

Garbeva, P., van Veen, J. a & van Elsas, J.D., 2004. Microbial diversity in soil: selection microbial populations by plant and soil type and implications for disease suppressiveness. *Annual review of phytopathology*, 42(29), pp.243–70.

Girvan, M. S., Bullimore, J., Pretty, J. N., Mark, A., Ball, A. S., & Osborn, a M. (2003). Soil Type Is the Primary Determinant of the Composition of the Total and Active Bacterial Communities in Arable Soils 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.

Gliessmon S.R., *Agroecology: ecological processes in sustainable agriculture*. Lewis Publishers (CRC Press), second ed., 2007.

Grayston, S.J., Vaughan, D. & Jones, D., 1997. Rhizosphere carbon flow in trees, in comparison with annual plants: The importance of root exudation and its impact on microbial activity and nutrient availability. *Applied Soil Ecology*, 5(1), pp.29–56.

Griffiths BS, Hallett PD, Kuan HL, Gregory AS, Watts CW & Whitmore AP (2008b) Functional resilience of soil microbial communities depends on both soil structure and microbial community composition. *Biol Fertil Soils* 44: 745–754.

Gotelli NJ and A. Chao, “Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data,” in Encyclopedia of Biodiversity, vol. 5, pp. 195–211, Elsevier Ltd., 2013.

Guedes MJ1, Pereira R, Duarte K, Rocha-Santos TA, Antunes SC, Gonçalves F, Duarte AC, Freitas AC. Sterols and fatty acid biomarkers as indicators of changes in soil microbial communities in a uranium mine area. 2011. *J Environ Sci Health A Tox Hazard Subst Environ Eng*. 2011;46(6):659-68.

Hill T.C., K. A. Walsh, J. A. Harris, and B. F. Moffett, “Using ecological diversity measures with bacterial communities,” *FEMS Microbiology Ecology*, vol. 43, pp. 1–11, 2003.

Halbach, M., Koschinsky, A. & Halbach, P., 2001. Report on the Discovery of *Gallionella ferruginea* from an active Hydrothermal Field in the Deep Sea Halbach. , 10(1), pp.1–15.

Hartmann, A., Rothballer, M. & Schmid, M., 2008. Lorenz Hiltner, a pioneer in rhizosphere microbial ecology and soil bacteriology research. *Plant and Soil*, 312(1-2), pp.7–14.

Henderson, I. R., Navarro-Garcia, F. & Nataro, J. P. (1998). The greatescape: structure and function of the autotransporter proteins. *Trends Microbiol* 6, 370–378.

Huang, W.Y. & Meinschein, W.G., 1979. Sterols as ecological indicators. *Geochimica et Cosmochimica Acta*, 43(5), pp.739–745.

- Huffman, E.N., 1977. Performance of a new automatic carbon dioxide coulometer. *Microchem. J.* 2, 567e573.
- Inceoğlu, O. et al., 2010. Effects of plant genotype and growth stage on the betaproteobacterial communities associated with different potato cultivars in two fields. *Applied and environmental microbiology*, 76(11), pp.3675–84.
- INEGI 2007. Guía para la interpretación de la carta edafológica.
- Kanehisa, M. & Goto, S., 2000. KEGG: Kyoto Encyclopaedia of Genes and Genomes. *Nucl. Acids Res.*, 28(1), pp.27–30.
- Kembel, S.W. et al., 2012. Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. *PLoS computational biology*, 8(10), p.e1002743.
- 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
- Kieft, T.L., soroker, E. & firestone, M.K., 1987. Microbial biomass response to a rapid increase in water potential when dry soil is wetted. *Soil Biology and Biochemistry*, 19(2), pp.119–126.
- Knief, C. (2014). Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Frontiers in Plant Science*, 5(MAY), 216.
- Knorre H, Krumbein WE. Microbial Sediments. In: Riding RE, Awramik SM, editors. Berlin, Heidelberg: Springer Berlin Heidelberg; 2000. pp. 25–31. doi:10.1007/978-3-662-04036-2\_4
- Köberl, M. et al., 2011. Desert farming benefits from microbial potential in arid soils and promotes diversity and plant health. *PLoS ONE*, 6(9).
- Kuramae, E.E. et al., 2012. Soil characteristics more strongly influence soil bacterial communities than land-use type. *FEMS Microbiology Ecology*, 79(1), pp.12–24.
- Kwon, K.K. et al., 2005. *Kordiimonas gwangyangensis* gen. nov., sp. nov., a marine bacterium isolated from marine sediments that forms a distinct phyletic lineage (*Kordiimonadales* ord. nov.) in the "Alphaproteobacteria". *International journal of systematic and evolutionary microbiology*, 55(Pt 5), pp.2033–2037.
- Laksmanan, V., Selvaraj, G. & Bais, H., 2014. Functional soil microbiome: Belowground solutions to an aboveground problem. *Plant physiology*, 166(October), pp.689–700. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25059708>
- Lang, R., 1915. Versuch einer exakten Klassifikation der Böden in klimatischer undgeologischer Hinsicht. *Internationale Mitteilungen fuer Bodenkunde* 5, 312–346.
- Langille, M.G.I. et al., 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*, 31(9), pp.814–821.
- Lauber, C.L. et al., 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), pp.5111–20.
- Liu L., Li Y., Li S., Hu H., He Y, Pong R, Lin D., Lu L., and Law M, "Comparison of next-generation sequencing systems.," *Journal of biomedicine & biotechnology*, vol. 2012, p. 251364, Jan. 2012.
- Lochhead A.G. (1940). Qualitatives studies of soil microorganisms. III. Influence on plant growth on the character of the bacterial flora. *Can J Res.* 18, 42.

- Li, W. & Godzik, A., 2006. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22(13), pp.1658–1659.
- Lindow, S.E. & Brandl, M.T., 2003. Microbiology of the Phyllosphere MINIREVIEW Microbiology of the Phyllosphere. *Applied and Environmental Microbiology*, 69(4), pp.1875–1883.
- Lugtenberg B.J.J., Chin-A-Woeng T.F.C. & Bloemberg G.V., 2002. Microbe-plant interactions: principles and mechanisms. *Antonie Van Leeuwenhoek*, **81**, 373-383.
- Lundberg, D.S. et al., 2012. Defining the core *Arabidopsis thaliana* root microbiome. *Nature*, 488(7409), pp.86–90.
- Madsen, E.L., 2011. Microorganisms and their roles in fundamental biogeochemical cycles. *Current opinion in biotechnology*, 22(3), pp.456–64.
- Mardis E.R., “Next-generation DNA sequencing methods.,” Annual review of genomics and human genetics, vol. 9, pp. 387–402, Jan. 2008
- Männistö MK, Tirola M. Häggblom MM. Bacterial communities in Arctic fields of Finnish Lapland are stable but highly pH-dependent. *FEMS Microbiol Ecol*. 2007;59:452–465.
- Masella, A.P. et al., 2012. PANDAseq: paired-end assembler for illumina sequences. *BMC bioinformatics*, 13(1), p.31.
- McCutcheon, J.P., McDonald, B.R. & Moran, N.A., 2009. Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. *PLoS Genetics*, 5(7).
- McMurdie, P.J. & Holmes, S., 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, 8(4).
- Metzker, M.L., 2009. Sequencing technologies — the next generation. *Nature Reviews Genetics*, 11(1), pp.31–46.
- Meyer, F. et al., 2008. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics*, 9, p.386.
- Murphy J. & Riley J. P. 1962. A modified single solution method for the determination of phosphate in natural waters. *Analytica Chimica Acta* 27:31–36.
- Muyzer, G., Waal, E.C. De & Uitterlinden, A.G., 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA.
- Nielson, K.L. et al., 1994. Carbon cost of root systems : a architectural approach. , 165, pp.161–169.
- Ottesen, A.R. et al., 2013. Baseline survey of the anatomical microbial ecology of an important food plant: *Solanum lycopersicum* (tomato). *BMC microbiology*, 13, p.114.
- Peiffer, J., Spor, A. & Koren, O., 2013. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of the*
- Peralta, R.M., Ahn, C. & Gillevet, P.M., 2013. Characterization of soil bacterial community structure and physicochemical properties in created and natural wetlands. *Science of the Total Environment*, 443, pp.725–732.
- Peralta E.I. & Spooner D.M. 2006. Genetic Improvement of Solanaceous Crops. Vol. 2 Tomato. CRC Press.

- Pii, Y. et al., 2015. The interaction between iron nutrition, plant species and soil type shapes the rhizosphere microbiome. *Plant Physiology and Biochemistry*, 99, pp.39–48.
- Raaijmakers, J.M. et al., 2008. The rhizosphere: a playground and battlefield for soilborne pathogens and beneficial microorganisms. *Plant and Soil*, 321(1-2), pp.341–361.
- Roesch, L., Fulthorpe, R. & Riva, A., 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. *The ISME ...*, 1(4), pp.283–290.
- Ryan, R.P. et al., 2008. Bacterial endophytes: recent developments and applications. *FEMS microbiology letters*, 278(1), pp.1–9.
- Robinson I.M. and Freundt E.A.: Proposal for an amended classification of anaerobic mollicutes. *Int. J. Syst. Bacteriol.*, 1987, 37, 78-81.
- Schlaeppli, K. et al., 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), pp.585–92.
- Semarnat. Informe de la Situación del Medio Ambiente en Meixico. Edición 2008. Compendio de Estadísticas Ambientales. Meixico. 2008.
- 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.
- SSSA. Glossary of Soil Science Terms. Disponible en: [www.soils.org/sssagloss/index.php](http://www.soils.org/sssagloss/index.php)
- Strauss, E.A., and G.A. Lamberti. 2002. Effect of dissolved organic carbon quality on microbial decomposition and nitrification rates in stream sediments. *Freshwat. Biol.* 47: 65-74.
- Sumner, M.E. Handbook of Soil Science. CRC Press. EU.2000.
- Tian, B.-Y., Cao, Y. & Zhang, K.-Q., 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(August), p.17087.
- Technicon Industrial System (1977) Technicon publication methods no.329-74W/B. Individual/simultaneous determinations of nitrogen and or phosphorous in BD acid digest. Technicon Industrial System, New York
- Tsurumaru, H. et al., 2015. Metagenomic analysis of the bacterial community associated with the taproot of sugar beet. *Microbes and environments / JSME*, 30(1), pp.63–9.
- Upreti, R. & Thomas, P., 2015. Root-associated bacterial endophytes from *Ralstonia solanacearum* resistant and susceptible tomato cultivars and their pathogen antagonistic effects. *Frontiers in microbiology*, 6(April), p.255.
- Uksa, M., Schloter, M., Endesfelder, D., Kublik, S., Engel, M., Kautz, T., ... Fischer, D. (2015). Prokaryotes in Subsoil—Evidence for a Strong Spatial Separation of Different Phyla by Analysing Co-occurrence Networks. *Frontiers in Microbiology*, 6(November), 1–13.
- Winston, M.E. et al., 2014. Understanding Cultivar-Specificity and Soil Determinants of the Cannabis Microbiome. *PloS one*, 9(6), p.e99641.
- Vitousek, P.M., 2002. Nutrient Cycling and Limitation: Hawai'i as a Model System. Princeton University Press, Princeton, NJ.

Wooley, J.C.J., Godzik, A. & Friedberg, I., 2010. A primer on metagenomics. *PLoS computational biology*, 6(2), p.e1000667.

Zhang, X.F. et al., 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), pp.1054–1065.