



**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL
INSTITUTO POLITÉCNICO NACIONAL**

UNIDAD ZACATENCO

DEPARTAMENTO DE BIOTECNOLOGÍA Y BIOINGENIERÍA

**“Caracterización metagenómica de las comunidades microbianas de la
rizósfera de árboles de limón que expresan antimicrobianos contra la
bacteria causante del HLB (Huanglongbing)”**

TESIS

Que Presenta

MARTÍN EDUARDO SÁNCHEZ GARCÍA

Para obtener el grado de

MAESTRO EN CIENCIAS

EN LA ESPECIALIDAD DE BIOTECNOLOGIA

Directores de Tesis: DRA. GUADALUPE BEATRIZ XOCONOSTLE CÁZARES

DR. JAIME GARCÍA MENA

Ciudad de México

Diciembre, 2016

Agradecimientos

A CONACYT por la beca número 570313 que se me fue otorgada como apoyo para la realización de mis estudios a nivel de maestría en el Departamento de Biotecnología y Bioingeniería de CINVESTAV.

Contenido

Agradecimientos	- 3 -
Índice de Figuras	- 7 -
Resumen	- 9 -
Abstract.....	- 11 -
INTRODUCCIÓN.....	- 13 -
Organismos Genéticamente Modificados (OGM's).....	- 14 -
OGM's y Seguridad Alimentaria.....	- 15 -
Pérdidas de cultivo y Huanglongbing.....	- 16 -
Huanglongbing (HLB)	- 17 -
Expresión de antimicrobianos para el control de <i>Candidatus liberibacter sp.</i>	- 19 -
Plantas OGM y su interacción con microorganismos	- 22 -
Estudio de poblaciones microbianas	- 24 -
Justificación.....	- 27 -
Hipótesis	- 27 -
Objetivo General.....	- 29 -
Objetivos específicos	- 29 -
Materiales y métodos	- 31 -
Plantación Experimental.....	- 31 -
Muestreo de suelo y tejidos	- 32 -
Suelo.....	- 32 -
Raíz.....	- 33 -
Extracción de DNA genómico de muestras de suelo	- 33 -
Extracción DNA de endófitos de raíz	- 33 -
Construcción de librerías	- 34 -
Análisis de datos de secuenciación.....	- 35 -
Análisis de calidad y depuración de datos	- 35 -
Acondicionamiento de datos	- 36 -
Filtrado de homopolímeros y selección de secuencias únicas	- 37 -
Alineamiento de secuencias	- 38 -
Depuración de quimeras.....	- 38 -
Clasificación filogenética.....	- 38 -
Resultados.....	- 41 -

Generación de Librerías.....	- 41 -
Librerías 16S	- 42 -
Librerías 18S	- 42 -
Análisis de datos de secuenciación de la región V3 del gen 16S.....	- 44 -
.....	- 50 -
Análisis de datos de secuenciación del gen 18S.....	- 52 -
Discusión.....	- 59 -
Conclusiones.....	- 67 -
Bibliografía.....	- 71 -

Índice de Figuras

FIGURA 1. ESTATUS ACTUAL DEL HLB EN MÉXICO. SE MUESTRA EN ROJO LAS ZONAS BAJO CONTROL FITOSANITARIO (SINAVEF).	17 -
FIGURA 2. ÁREA DE SIEMBRA EXPERIMENTAL.	31 -
FIGURA 3. COLECTA DE MUESTRAS SIMPLES PARA LA FORMACIÓN DE MUESTRAS COMPUESTAS POR EL MÉTODO DE CUARTEO.	31 -
TABLA 1. DISTRIBUCIÓN DE ÁRBOLES DE LOS DIFERENTES TRATAMIENTOS EN LA PLANTACIÓN EXPERIMENTAL.	32 -
TABLA 2. PARÁMETROS USADOS EN TRIMMOMATICSE	35 -
FIGURA 4. FLUJO DE TRABAJO EN "GALAXY PROJECT" PARA LA ASIGNACIÓN DE IDENTIFICADOR DE SECUENCIA Y GRUPO.	36 -
FIGURA 5. FLUJO DE TRABAJO PARA EL ANÁLISIS DE DATOS CON LA PLATAFORMA DE "MOTHUR".	37 -
TABLA 3. CUANTIFICACIÓN DE DNA POR ESPECTOMETRÍA EN NANODROP 2000 (THERMO FISHER SCIENTIFIC).	41 -
FIGURA 7. ELECTROFORESIS EN GEL DE AGAROSA AL 1% DE DNA GENÓMICO DE MUESTRAS DE RAÍZ.	41 -
FIGURA 6. ELECTROFORESIS EN GEL DE AGAROSA AL 1% DE DNA GENÓMICO DE MUESTRAS DE SUELO.	41 -
FIGURA 9. ELECTROFORESIS EN GEL DE AGAROSA AL 2% DE PRODUCTOS DE PCR 18S DE MUESTRAS DE SUELO.	42 -
FIGURA 8. ELECTROFORESIS EN GEL DE AGAROSA AL 2% DE PRODUCTOS DE PCR 16SV3 DE MUESTRAS DE SUELO.	42 -
FIGURA 10. ELECTROFORESIS EN GEL DE AGAROSA AL 2% DE PRODUCTOS DE PCR 16S Y 18S PARA MUESTRAS DE RAÍZ.	43 -
TABLA 4. RELACIÓN DEL CAMBIO EN EL NÚMERO DE SECUENCIAS EN LOS DIFERENTES PASOS DEL ANÁLISIS DE AMPLICONES 16S. ...	44 -
TABLA 5. OTUS DE BACTERIAS IDENTIFICADOS A DIFERENTES NIVELES TAXONÓMICOS.	45 -
TABLA 6. ÍNDICES DE DIVERSIDAD DE BACTERIAS CALCULADOS A PARTIR DE DATOS NORMALIZADOS.	45 -
FIGURA 11. CURVA DE RAREFACCIÓN PARA DATOS 16S.	46 -
FIGURA 12. ABUNDANCIA RELATIVA DE BACTERIAS AGRUPADAS A NIVEL DE "FILO".	47 -
FIGURA 13. DISTRIBUCIÓN DE ABUNDANCIA RELATIVA EN SUELO DE OTUS DE BACTERIA A NIVEL DE "FAMILIA".	48 -
FIGURA 14. DISTRIBUCIÓN DE ABUNDANCIA RELATIVA EN RAÍZ DE OTUS DE BACTERIA A NIVEL DE "FAMILIA"	49 -
FIGURA 15. ANÁLISIS DE DISIMILITUD DE POBLACIONES DE BACTERIA A PARTIR DE LA MATRIZ DE DISTANCIA GENERADA POR EL MÉTODO DE BRAY-CURTIS.	50 -
TABLA 7. COMPARACIÓN DE GRUPOS CON EL ALGORITMO DE UNIFRAC WEIGHTED, SE MUESTRA LA DISIMILITUD (WSCORE) Y LA SIGNIFICANCIA DEL ANÁLISIS PARA CADA PAR DE DATOS DE MUESTRAS 16S.	51 -
TABLA 8. TABLA AMOVA BACTERIAS.	51 -
TABLA 9. RELACIÓN DEL NÚMERO DE SECUENCIAS A LO LARGO DE LAS DIFERENTES ETAPAS DEL ANÁLISIS DE AMPLICONES 18S.	52 -
TABLA 10. NÚMERO DE OTUS DE HONGOS DETERMINADOS A DIFERENTES NIVELES FILOGENÉTICOS.	53 -
TABLA 11. ÍNDICES DE DIVERSIDAD DE HONGOS CALCULADOS A PARTIR DE DATOS NORMALIZADOS.	53 -
FIGURA 16. CURVAS DE RAREFACCIÓN DE LIBRERÍAS 18S EN MUESTRAS DE SUELO Y RAÍZ.	54 -
FIGURA 17. ABUNDANCIA RELATIVA DE HONGOS AGRUPADOS FILOGENÉTICAMENTE A NIVEL DE CLASE.	55 -
FIGURA 18. ABUNDANCIA RELATIVA DE OTUS DE HONGOS DEFINIDOS A NIVEL DE FAMILIA.	56 -
FIGURA 19. ANÁLISIS DE DISIMILITUD DE POBLACIONES DE HONGOS A PARTIR DE DISTANCIAS CALCULADAS POR EL MÉTODO DE BRAY-CURTIS.	56 -
TABLA 12. COMPARACIÓN DE GRUPOS CON EL ALGORITMO DE UNIFRAC WEIGHTED, SE MUESTRA LA DISIMILITUD (WSCORE) Y LA SIGNIFICANCIA DEL ANÁLISIS PARA CADA PAR DE DATOS DE MUESTRAS 18S.	57 -
TABLA 13. TABLA AMOVA HONGOS.	58 -
ECUACIÓN 1. COBERTURA DE GOOD.	60 -
FIGURA 20. CURVAS DE ACUMULACIÓN DE ESPECIES PARA DATOS DE BACTERIA DE LA MUESTRA "TIEMPO 3", SE INCLUYEN LOS INTERVALOS DE CONFIANZA PARA COMPARAR LA SIMILITUD DE LAS MUESTRAS. SE GRAFICA EN "X" EL NÚMERO DE LECTURAS Y EN "Y", EL NÚMERO DE FILOTIPOS OBSERVADOS (S.OBS).	60 -
FIGURA 21. CURVAS DE ACUMULACIÓN DE ESPECIES PARA DATOS DE HONGO DEL TRATAMIENTO "DEFENSIN" Y EL CONTROL "WILDTYPE". SE INCLUYEN LOS INTERVALOS DE CONFIANZA PARA COMPARAR LA SIMILITUD DE LAS MUESTRAS. SE GRAFICA EN "X" EL NÚMERO DE LECTURAS Y EN "Y", EL NÚMERO DE FILOTIPOS OBSERVADOS (S.OBS).	61 -

FIGURA 22. COMPARACIÓN DE DIVERSIDAD A PARTIR DEL ÍNDICE INVERSA DE SIMPSON. SE MUESTRA EN A) DATOS DE BACTERIA Y EN B) DATOS DE HONGOS. - 62 -

FIGURA 23. AGRUPACIÓN CONSENSO DE MUESTRAS DE BACTERIAS. SE SEÑALA EN CADA RAMA EL COEFICIENTE DE BOOTSTRAPPING PARA ESA CONFIGURACIÓN..... - 64 -

Resumen

Se ha estudiado la expresión y actividad de antimicrobianos en planta con el fin de desarrollar estrategias para combatir fitopatógenos como bacterias y hongos; sin embargo, considerando la importancia de los microorganismos en suelo y en las raíces de plantas y la actividad inespecífica de los antimicrobianos, en el presente trabajo se estudiaron las poblaciones de microorganismos en la rizósfera de árboles de limón mexicano que expresan Lisozima o Defensina desarrollados con la finalidad de atacar la bacteria *Candidatus liberibacter* sp., agente causal del Huanglongbing en cítricos.

El estudio de las poblaciones se llevó a cabo a partir de la secuenciación de amplicones del gen 16S y 18S, de DNA extraído de suelo y raíz de la rizósfera de tres tratamientos (Expresión de Lisozima, Defensina y Lisozima/Defensina).

En muestras de suelo, la diversidad de bacterias fue muy similar entre tratamientos y el control, los grupos de mayor dominancia *Proteobacteria*, *Firmicutes*, *Acidobacteria* y *Actinobacteria*, entre otros, concuerdan con la información descrita en diversos estudios de poblaciones en suelo. En raíz la diversidad es menor, hay una alta dominancia de bacterias de la familia *Bacillaceae* y el comportamiento es muy similar entre las muestras. Con el análisis de distancia se observa que existe una relación de similitud entre muestras en función del tiempo de muestreo lo que es confirmado con el análisis de varianza el cual descarta que haya alguna influencia del tratamiento.

En hongos, las poblaciones son muy similares durante el tiempo de muestreo "0" y "1", en los siguientes periodos parecen tener un comportamiento similar pero los cambios en abundancia son más heterogéneos. La clase *Sordariomycetes* es la más abundante junto con dos grupos no clasificados (Otu01 y Otu02); es notable la presencia de hongos del género *Fusarium*, descrito como un agente patógeno, en un rango de abundancia del 20 al 80%. Finalmente el análisis de varianza señala que existe una relación entre las muestras y el tiempo de muestreo al igual que en bacterias, además que no se rechaza la hipótesis nula de que la estructura de poblaciones de hongos es distinta entre tratamientos.

Abstract

It has been studied the expression and activity of antimicrobials in plants with the aim to develop strategies to deal with phytopathogens. Recognizing the importance of soil and root microorganisms and knowing antimicrobials are non-specific, it is studied in this work the microbiome in the rhizosphere of Mexican lime tree that express lysozyme or defensin with the propose to attack *Candidatus liberibacter* sp., responsible bacteria for developing the Huanglongbing illness.

The population study was achieved with sequencing of 16S and 18S amplicons, the DNA for this assay was extracted from soil and root of three different treatments (Lysozyme, Defensin and Lysozyme/Defensin expression).

In soil samples, the diversity of bacterial communities was closed similar among treatments and wildtype control. The most dominant groups found are *Proteobacteria*, *Firmicutes*, *Acidobacteria*, *Actinobacteria*, and others; this structure matches information of several researches about soil. Diversity indexes are lower in root samples than in soil with a high dominance of bacteria from *Bacillaceae* family, besides there is a similar behavior among samples. The clustering of the matrix distance shows a relation between samples and time, this relationship is confirmed with the analysis of variance in addition that the influence of the treatment is dismissed.

Along time "0" and "1", populations of fungi in treatments and wildtype are highly similar in compare with the next sample times where the samples seems to have a similar behavior even when the changes in abundance are variable. *Sordariomycetes* class is the most abundant group found in soil and roots along with two unclassified groups; it is meaningful the presence of organisms from *Fusarium* genera in a range of abundance of 20-80% which are commonly described with pathogenic activity. Finally, the analysis of variance lets conclude that there is a relationship between the change in the structure of communities and the sample time besides that it's not reject the null hypothesis about that fungi communities among treatments are similar.

INTRODUCCIÓN

Como parte de la celebración anual de la fundación de la FAO conmemorada el 16 de octubre, se promueve una concientización mundial sobre la problemática de hambre en el mundo recalcando la importancia de la necesidad de alcanzar la seguridad alimentaria. El tema de este año “Climate is changing. Food and agricultura must too” hace notar el reto que enfrenta el campo ante las condiciones adversas que el cambio climático trae consigo para las poblaciones rurales que son las más vulnerables y para lograr la seguridad alimentaria para el 2030. Por lo que la FAO invita a los países a considerar la agricultura en sus planes de acción contra el cambio climático e invertir más en el desarrollo del campo (FAO, 2016).

Frente a estas circunstancias los organismos genéticamente modificados podrían tener una valiosa participación dado que la selección de características que incrementen la resistencia a factores bióticos (Plagas y enfermedades) y abióticos (Estrés hídrico, suelos salinos, etc.) disminuiría la pérdida de cultivos.

Los cítricos en nuestro país así como en muchos otros se encuentran amenazados por el Huanglongbing una enfermedad causada por la bacteria *Candidatus liberibacter* sp. que se aloja en el floema de plantas infectadas lo que dificulta la efectividad de las estrategias de control, por lo que el equipo de trabajo desarrolló variedades transgénicas de limón expresoras de antimicrobinos (Lisozima, Defensina, Lisozima/Defensina) para combatirla.

Aún con el potencial que los organismos genéticamente modificados tienen para combatir este tipo de situaciones adversas, hay una discusión alrededor de la seguridad en su empleo que ha limitado su aprovechamiento, por lo que las evaluaciones de inocuidad son indispensables. En su mayoría las evaluaciones se han centrado en el estudio de la biodiversidad de los agroecosistemas, y de las comunidades de microorganismos en la rizósfera debido a las relaciones que establecen con las plantas y a que son indicadores de la salud del suelo.

Organismos Genéticamente Modificados (OGM's)

Los OGM's fueron desarrollados de manera análoga a los procesos de mejoramiento tradicional practicados desde la domesticación de cultivos donde la selección y cruce de parentales con características deseables generan nuevas variedades con mayores tasas de producción, resistencia a plagas y enfermedades, entre otras (Gasman, *et al.*, 2009; Chassy, 2007).

Esta tecnología surge en los ochentas como resultado de un mayor entendimiento del funcionamiento de los genes, lo que permitió una rápida y precisa transferencia de nuevas características en cultivos de interés. En la primera generación de este tipo de cultivos se incorporaron características de resistencia a insectos, virus y herbicidas; mientras que en la segunda generación se incrementó la calidad y composición del producto, su tolerancia a factores de estrés y su eficiencia fotosintética (Gasman, *et al.*, 2009, Bennett, *et al.*, 2013).

Ocampo y colaboradores señalan que los cultivos GM destinados a la alimentación han probado ser inocuos, sin embargo hay una gran vigilancia alrededor de ellos y hasta el 2009 se habían autorizado en más de 21 países variedades transgénicas de maíz, soya, arroz, colza, jitomate, papaya y calabaza entre otros. En esta publicación reportan una evaluación llevada a cabo para conocer el beneficio y utilidad que los cultivos GM pudieran tener en pequeñas comunidades, mencionan que debido a la falta de mano de obra en campo para prácticas de deshierbe estos son abandonados; sin embargo, con el uso de OGM's y herbicidas por un periodo de dos años se podrían controlar las malezas dando lugar a una reducción en el consumo de herbicidas pasado ese periodo (Gasman, *et al.*, 2009); debe considerarse que la ingeniería genética es una herramienta biotecnológica neutra (Gutiérrez-Galeano, *et al.*, 2015) y dependerá de la creatividad del biotecnólogo.

El impacto de los OGM en el aumento de la producción y reducción de pesticidas está sujeto a variaciones dependiendo de las variedades silvestres usadas en cada lugar, del tipo de cultivo, la presencia de plagas y enfermedades, entre otros; sin embargo, algunas estimaciones sugieren que se lograría un incremento en la tasa de producción por hectárea de cultivos de maíz y algodón, a su vez que se lograría la reducción de insecticidas en cultivos como el algodón con valores del 30 al 40% en países como Argentina, EUA, China y Sudáfrica.

En 1991 se dio la primera liberación experimental de OGM's en el país, una variedad de jitomate resistente a insectos y desde entonces se han sembrado 21 cultivos diferentes en un total de 1116 casos (Gasman, *et al.*, 2009).

OGM's y Seguridad Alimentaria

Clark y colaboradores (2014) mencionan que para alcanzar la demanda creciente de alimentos se requerirá incrementar la productividad agrícola, no por el aumento del área cultivable sino por la mejora de los cultivos. Para esto los organismos genéticamente modificados representan una prometedora tecnología para incrementar la producción y aumentar las oportunidades de uso de tierras marginadas y reducir el uso de agroquímicos (Ozor, 2008).

Según la FAO, para el 2050 seremos 9.6 mil millones de habitantes por lo que se requiere que la agricultura sea más productiva y a la vez sustentable de manera que se cubran las necesidades de alimentación, el cual es un objetivo del Milenio planteado a cumplir para el 2030. Debido a que la mayoría de la población en situación de riesgo nutricional vive en países en desarrollo y el 75% en zonas agrícolas, es deseable enfocar los esfuerzos en el desarrollo de estas regiones dado que la situación de estas comunidades está relacionada con condiciones de pobreza, falta de buenas prácticas agrícolas, la presencia de climas hostiles, etc. (Chassy, 2002; Ozor, *et al.*, 2007).

En África subsahariana hay una creciente inseguridad alimentaria en países como Etiopía, Kenia, Malawi, Mozambique, entre otros, debido a situaciones como falta de semilla, además que los suelos en esta región son muy pobres y solo el 4% del suelo cultivable puede ser regado, se tienen por otro lado severas sequías y deterioro ambiental; en otros países no es ecológicamente viable extender las áreas de cultivo, lo que hace necesario desarrollar variedades y estrategias para el cultivo en estas y otras situaciones similares (Ozor, *et al.*, 2007; Clark, *et al.*, 2014).

Aun cuando variedades de maíz-Bt han sido aprobados por la EPA, la FDA y el Departamento de Agricultura de EUA para consumo humano, existen diferentes cuestionamientos sobre seguridad, ética y bienestar que han postergado su introducción en diferentes países (Ozor, 2007; Chassy, 2002).

Pérdidas de cultivo y Huanglongbing

Con la finalidad de incrementar la productividad de los cultivos es necesario trabajar en el control de plagas y enfermedades, ya que son factores que impactan fuertemente la producción como se mencionará en esta sección. Podemos ejemplificar la gravedad del problema con noticias como la publicada en el 2015 por el periódico La Jornada, donde se da a conocer la situación de campesinos en Chiapas y Oaxaca frente a las afectaciones en un 70% de las plantaciones de café por roya. Según reportan, esto ha ocasionado una reducción de 500 mil quintales a 150 mil (Un quintal es equivalente a 100 Kg), mencionan también que en Veracruz se tiene una pérdida en el 70% de su producción y que la renovación de estas zonas cafetaleras requerirá una inversión inicial de 100 millones de pesos con resultados que se reflejarán en un plazo de 5 a 6 años (Pérez, 2015).

García-Lara y Bergvinson (2007) describen el problema de la pérdida post-cosecha de maíz en México, haciendo una revisión sobre los factores involucrados, las plagas relacionadas y las alternativas de solución. En este trabajo citan un estudio realizado por CYMMIT donde señalan la infertilidad del suelo y los ataques por plagas como los factores más importantes en la producción, estos generan pérdidas del 70 al 95% bajo condiciones deficientes de manejo. Mencionan a su vez que la temperatura y humedad juegan un papel importante en los niveles de daño y pérdida, además que la carencia de buenas prácticas acentúa el problema.

Un informe del 2005 de Argentina sobre el control de plagas en agricultura señala que las pérdidas en este país son del 10% para soya y girasol y del 15% para trigo y maíz. Recordemos que la magnitud de la pérdida está relacionada con la condición socioeconómica del productor. Por ello, a nivel mundial se tienen pérdidas del 40% de la producción, reduciéndose al 30% por la aplicación de medidas de control, sin embargo las pérdidas son diferentes para cada región. En países como Japón, EUA y la Comunidad Europea Occidental esta cifra ronda del 10 al 30% y en países más pobres se incrementa en un rango del 40 al 75% (Huerga y San-Juan, 2005).

La presencia de enfermedades y plagas cierra el mercado para la exportación debido a las regulaciones que los países imponen para la importación, ya que como ejemplifica Huerga y San-Juan (2005), la Unión Europea, Japón y Brasil cierran sus fronteras a productos de países con presencia de Carpocapsa, Mosca de los Frutos, entre otros.

En este informe se señala también la preocupación por el uso de plaguicidas debido al daño potencial en la salud de los trabajadores y del medio ambiente, se afirma que aun cuando juegan un papel importante en la producción de cultivos, el manejo integral de plagas debe garantizar a su vez la salud de los agricultores y del medio ambiente (Hueriga y San-Juan, 2005).

Huanglongbing (HLB)

El Huanglongbing es un padecimiento que afecta severamente la producción de cítricos a nivel mundial y está presente en nuestro país en 391 municipios de 22 entidades, 314 de los cuales son aéreas citrícolas equivalentes al 43% del total en el país según datos reportados por SENASICA (2016).



Figura 1. Estatus actual del HLB en México. Se muestra en rojo las zonas bajo control fitosanitario (SINAVEF).

Fue descrito por primera vez en 1919 en China y en los años posteriores se identificaron dos formas del padecimiento, una correspondiente a la variedad asiática y la otra a una variedad africana, se reporta que ambas muestran comportamientos similares pero se desarrollan bajo condiciones de temperatura diferente. En esta revisión de 1992, el autor reporta que hasta esta fecha se estima que en Asia y África se habían perdido más de 60 millones de plantas.

Debido a la expansión de la enfermedad en los 50's a provincias de China y Taiwan, se comenzaron a buscar medidas para su erradicación a partir de la eliminación de árboles enfermos y el establecimiento de plantaciones nuevas en zonas libres de HLB, sin embargo su expansión continuó en Indonesia y Filipinas reduciendo las plantaciones en un 60% en estos países con reportes posteriores en otras áreas como Malasia, la Península Arábiga e India (Aubert, 1992).

Este padecimiento es considerado la enfermedad más devastadora a nivel mundial en cítricos debido a los daños que ocasiona en plantas y frutos, la dificultad en el diagnóstico temprano y los elevados costos económicos y ambientales para su manejo. El agente causal es una bacteria gram - transmitida por insectos que se limita al floema de las plantas lo que ocasiona una oclusión del sistema vascular. La infección es resultado de una interacción entre el hospedero, el patógeno y el vector; se dispersa a través de *Diaphorina citri* y *Trioza eritreae* (Santivañez, et al., 2013).

Santivañez y colaboradores (2013), mencionan que los síntomas son muy similares a la deficiencia de minerales, principalmente Zinc, por lo que suele confundirse y debido a su aparición lenta se dispersa con rapidez antes de mostrar síntomas severos. La aparición de los síntomas comienza en unas cuantas hojas con coloración amarilla contrastando con el resto de la planta, se esparce por todo el árbol generando defoliación y muerte progresiva de ramas afectadas. Su confirmación solo es posible por PCR.

Además del impacto económico generado por los costos de manejo y las pérdidas en el rendimiento, la eliminación y quema de millones de árboles, así como el uso intensivo de plaguicidas generan un gran impacto ambiental en suelo, agua y aire (Santivañez, et al., 2013); por lo que en México y en otros países se ha incluido el control biológico dentro de las estrategias de control del vector. En el caso de nuestro país se ha llevado a cabo la producción masiva de la avispa *Tamarixiasp.* y su liberación en áreas específicas (Arredondo-Bernal, 2013).

En la publicación de Arredondo-Bernal y colaboradores (2013), se recauda información de las estrategias desarrolladas en países de América Latina, que incluye la eliminación de árboles positivos a la infección, la generación y monitoreo de plantas certificadas, el establecimiento de comisiones y políticas para el control del HLB y su vector.

En México en 2010 se publicó en el Diario Oficial de la Federación el "Acuerdo por el que se dan a conocer las medidas fitosanitarias que deberían aplicarse para el control del Huanglongbing (*Candidatus liberibactersp.*) y su vector". Establece que cuando haya evidencia de la presencia de HLB, se deberá controlar la movilización de material hospedante y su vector, establecer puntos de

verificación y retener y disponer o destruir material infectado. Enlista también los requisitos que se deben cubrir para obtener certificación para la producción de material vegetal propagativo y los lineamientos que deben seguir los huertos. Las medidas de control que contempla, son la aplicación de producto químico, eliminar plantas detectadas con la enfermedad y la eliminación de la totalidad de árboles de una plantación cuando se detecte HLB en 28% de las plantas.

Un estudio de análisis de riesgo generado en el 2011 por Salcedo y colaboradores, indica que se prevén pérdidas de empleos en nuestro país principalmente en el sector primario, con pérdidas económicas de mayor repercusión en el sector de bebidas, alimentos y tabaco. El impacto en la reducción de empleos será de 4000 a 27000 considerando situaciones de bajo y alto impacto, con una pérdida de 26000 a 126000 empleos directos e indirectos después de 5 años y una repercusión económica de hasta 7000 millones de pesos, sin embargo este solo es un modelo estadístico generado a partir de diferentes indicadores.

Aunque las medidas de control tomadas hasta ahora están enfocadas en el control del vector y la erradicación de árboles positivos al HLB, Aubert (1992) reporta la aplicación de químicos combinado con inyecciones de antibióticos para atacar a la bacteria causante de HLB en Sudáfrica. A su vez Gómez-Jaimes, *et al.*, (2011) describen un estudio para el control del HLB mediante la aplicación de bactericidas (Kasugamicina, sulfato de cobre pentahidratado y tetraciclina+oxitetraciclina) aplicado a follaje y o suelo, al final de este estudio no encontraron una diferencia significativa entre los diferentes ensayos por lo que los autores sugieren continuar las evaluaciones con diferentes dosis, ingredientes activos e intervalos de aplicaciones.

En estos últimos reportes se observa la intención de atacar la problemática desde otro punto, enfocando las estrategias en el control de la bacteria, sin embargo es importante hacer notar que en árboles infectados, la bacteria se localiza en el floema por lo que no es de fácil acceso para la aplicación de bactericidas.

Expresión de antimicrobianos para el control de *Candidatus liberibacter* sp.

Se ha hecho uso de antimicrobianos para el control de fitopatógenos en diversos trabajos, entre los cuales se citan enzimas hidrolíticas como quitinasas y glucanasas que pueden degradar la pared celular de hongos (Nakajima *et al.*, 1997). Se han también evaluado el uso de lisozima y defensina

humana para conferir resistencia contra diversos patógenos; entre ellos, Zecua-Nájera (2013) desarrolló una estrategia para el control de HLB mediante la expresión de lisozima humana.

La lisozima, es definida como una enzima con actividad hidrolítica capaz de romper la unión entre N-acetilglucosamina y el ácido N-acetilmurámico en la pared celular de bacterias (Düring, 1993; Huang, *et al.*, 2002), específicamente la lisozima humana tiene un amplio espectro de acción y alta actividad catalítica por lo que se ha promovido su uso (Nakajima, *et al.*, 1997). Huang y colaboradores (2002) expresaron y evaluaron la actividad de lisozima humana en arroz obteniendo resultados similares a la proteína expresada en humanos y en gallina. Por otro lado Nakajima *et al.*, (1997) retaron plantas de tabaco que expresan lisozima humana contra el hongo fitopatógeno *Eryshiphe cichoracearum* el cual mostró una reducción en el área colonizada por el hongo de 0.06 y 0.01% en comparación con el 8.02% en plantas silvestres y al retarse con *Pseudomonas syringae* el desarrollo de halos cloróticos en plantas transgénicas se redujo al 17 y hasta el 50% en comparación con las plantas silvestres. Estos resultados permiten concluir que la expresión de lisozima humana puede incrementar la resistencia a hongos y bacterias fitopatógenas.

Las defensinas por otro lado, están conservadas en plantas, insectos, así como invertebrados y vertebrados (Schröder y Harder, 1999) y así como la lisozima humana, la B-defensina humana es muy potente al interactuar y desestabilizar las membranas de bacterias gram negativas y levaduras. Las defensinas poseen propiedades anfipáticas ya que los dominios catiónicos se acercan a las porciones negativas de los fosfolípidos y la porción hidrofóbica interacciona con las cadenas de lípidos en la membrana lo que la desestabiliza y causa su permeabilización (Aerts, *et al.*, 2007; Ganz, 2004, Schröder y Harder, 1999). Aerts y colaboradores expresaron la proteína en *Arabidopsis thaliana* para lo cual reportaron fue necesario añadir una etiqueta a la proteína para que esta pasará por el sistema de secreción y lograr la generación de enlaces disulfuros necesarios para su actividad. Al evaluar la actividad antifúngica de poblaciones con alta y baja expresión y plantas silvestres, se observó que la población con alta expresión tiene una notable reducción en la susceptibilidad a *Botrytis cinerea* (Hongo necrotrófico) en comparación con la población control y la línea de baja expresión que exhibe a su vez una reducción en la susceptibilidad observable solo en los primeros 3 días post-inoculación.

Como una estrategia para combatir el HLB, Zecua-Nájera (2013) desarrolló plantas de limón mexicano que expresan lisozima humana con el fin de conferirles resistencia a la enfermedad. El objetivo de Zecua-Nájera consistió en expresar lisozima fusionada a *CsPP16*, proteína capaz de

entrar en el sistema vascular de las plantas y llevar el antimicrobiano para atacar así la bacteria causante. La secuencia de lisozima la obtuvo del GenBank, optimizó el uso de codones para expresión en cítricos y la fusionó a *CsPP16* por una bisagra que permite el plegamiento independiente. Para la transformación de tejido de limón utilizó cepas de *Agrobacterium* con un plásmido binario con la secuencia de lisozima entre los bordes izquierdo y derecho del T-DNA de *Agrobacterium*.

Zecua-Nájera realizó diferentes ensayos para evaluar la efectividad de la estrategia, una de ellas consistió en generar una transformación transitoria en árboles adultos con síntomas de HLB, sesenta días posteriores a la transformación logró apreciar la ausencia de síntomas de HLB en brotes transformados mientras que en hojas maduras los síntomas persistieron, menciona que es debido posiblemente a la oclusión del sistema vascular en hojas maduras que impiden el movimiento de la proteína. Explica que, la medición de la tasa fotosintética muestra que las plantas enfermas tratadas presentaron valores similares a las plantas sanas, caso contrario al observado en plantas enfermas que no fueron tratadas o que se transformaron para la expresión de *GUS*, las cuales presentaron un abatimiento en la tasa fotosintética.

Transformó también árboles certificados libres de infección y los retó en un ambiente hostil es decir con presencia de infección en cuyo caso fue Tecomán, Colima. En este ensayo identificó una reducción en el número de bacterias vivas en árboles tratados.

Al final concluye que no constituye una medida de control, pero sí de mitigación de los síntomas y del proceso patogénico; el tratamiento mejora las condiciones fisiológicas de las plantas además que disminuye notablemente el contenido de bacterias vivas, lo que podría reducir los efectos negativos en la producción.

En general, estos reportes muestran la efectividad que los antimicrobianos humanos defensina y lisozima pueden conferir a las plantas contra el ataque de hongos y bacterias fitopatógenas y ya que ambas moléculas tienen un rango de acción amplio es necesario evaluar el efecto de este tipo de estrategias en las poblaciones microbianas que interactúan negativa y positivamente con las plantas.

Plantas OGM y su interacción con microorganismos

Lee et al., (2011) describe las bacterias y hongos como componentes críticos de los ciclos biogeoquímicos y como indicadores de la salud del suelo, por tanto menciona, que la estabilidad ecológica de los cultivos GM podría depender de su relación con las comunidades microbianas del suelo.

Es citado por diferentes autores que las plantas tienen la capacidad de influir en el establecimiento de microorganismos en el suelo mediante la excreción de compuestos a través de la raíz, estos incluyen tanto productos del metabolismo primario como secundario, lo que lleva al enriquecimiento de diferentes microorganismos en la rizósfera en comparación con el suelo circundante (Bais, *et al.*, 2006; Guttman, *et al.*, 2014). Este proceso se denomina efecto rizosférico, Bais, *et al.* (2006) cita las observaciones de Hilner sobre el aumento en el número y actividad de microorganismos en la vecindad de las raíces.

En un estudio realizado por Gundale y colaboradores (2014) se observa la influencia de la microbiota del suelo en el desarrollo de pinos dado que realizaron ensayos con suelo extraído de Canadá y Suecia en el que hacen crecer pinos canadienses. Usando ambos tipos de suelo esterilizados como controles cuantificaron un 43% más de biomasa en árboles crecidos en el suelo sueco, por otro lado al inocular suelo estéril con extracto de suelo sueco observaron un incremento en la biomasa de los pinos esto hace notar que el desarrollo de plantas está fuertemente influenciado por las comunidades bióticas de la rizósfera.

Otros estudios evalúan la capacidad de establecer relaciones planta-microorganismo en función de la etapa de desarrollo y genoma de la planta. Chaparro y colaboradores (2014) evalúan las diferencias de las comunidades de la rizósfera a lo largo del desarrollo de *Arabidopsis*, en este estudio notan cambios significativos en la etapa de floración y germinación de la planta por procesos relacionados posiblemente con la respuesta inmune y que parece ser también común bajo condiciones de estrés para la planta como condiciones limitantes de nutrientes. Haney y colaboradores (2015) evalúan la capacidad de diferentes accesos de *Arabidopsis* en el establecimiento de relaciones. En este estudio además de observar que la interacción de plantas con bacterias está ligada al acceso a nutrientes, las diferencias en el genotipo de una misma especie producen diferencias en la capacidad de las plantas para soportar el crecimiento de microorganismos en la raíz alterando con ello la capacidad de las plantas para responder a diferentes factores.

Dado que el genotipo de las plantas genéticamente modificadas es diferente al de sus parentales silvestres, se han realizado estudios sobre el efecto del transgen en el establecimiento de relaciones en la rizósfera de las plantas.

Una evaluación de una línea transgénica de arroz que contiene un inserto de la fusión de trehalosa-6-fosfato sintasa/fosfatasa (Incrementa la resistencia a sequía y frío) muestra que no hay diferencias significativas en la diversidad de bacterias y hongos en comparación con la línea silvestre, a su vez los resultados indican una mayor asociación con las variables ambientales (Lee, *et al.*, 2011).

Cotta y colaboradores (2013) evalúan por otro lado el efecto de la expresión de la proteína CRY en maíz-Bt (utilizado para combatir plagas de insectos) sobre las poblaciones de microorganismos en la rizósfera. Las evaluaciones a partir de amplificación del gen 16S y de la región ITS muestran que la estructura de las poblaciones de microorganismos en el suelo se vio afectada por el tipo de suelo y la etapa de desarrollo de la planta sin efecto aparente por el transgen.

Icoz y Stotzky (2008) determinan la presencia de la proteína CRY a partir de cultivos hidropónicos de maíz-Bt en los exudados de las raíces de maíz. Posteriormente evalúan la persistencia de esta proteína en suelo, siendo la proteína CRY más estable en caolinita y derivados enriquecidos con caolinita o montmorillonita (Fracciones minerales del suelo que corresponden a arcilla). Los resultados muestran que la persistencia de la proteína cambio en función de la cantidad de arcilla, aumentando a mayor cantidad de caolinita y menor de montmorillonita, por otro lado el aumento en el pH del suelo promovió la degradación de la proteína. La persistencia varió según los diferentes parámetros de 10 a 50 días y ya que no se detectó una acumulación de la proteína se sugiere un riesgo bajo para el medio ambiente.

En conjunto estos estudios sugieren en concordancia con diversos autores que, son deseables las interacciones planta-microorganismo ya que pueden influenciar positivamente el crecimiento de plantas a través de diferentes mecanismos como movilización de nutrientes, protección contra patógenos e incremento en la tolerancia a estrés biótico y abiótico (Bais, *et al.*, 2006; Guttman, *et al.*, 2014), por otro lado estas interacciones están influenciadas principalmente por los factores edáficos, por el genotipo del hospedero y por la concentración y estabilidad de los compuestos excretados por la planta (Bais, *et al.*, 2006; Zarraonaindia, *et al.*, 2015).

Estudio de poblaciones microbianas

Dado que la mayoría de los microorganismos no son cultivables, el entendimiento de la diversidad microbiana se ha mejorado mediante el uso de DNA como marcador molecular (Nocker, *et al.*, 2007). Existen diversos trabajos que ejemplifican la necesidad de evaluar tanto los microorganismos cultivables como los no cultivables; García-Armisen, *et al.*, (2010) estudiaron la diversidad de microorganismos en procesos de fermentación de cacao, señalan que aun cuando se han desarrollado estudios previos de diversidad en este proceso, éstos han omitido las poblaciones no cultivables, por lo que emplearon el gen 16SrDNA para obtener una imagen cualitativa y semicuantitativa de la comunidad microbiana. De forma análoga Abassian, *et al.*, (2015), aplica metodologías no dependientes de cultivo para el estudio de columnas de Winogradsky ya que menciona esto permitiría disponer de un análisis más completo del ambiente.

Entre las estrategias para el estudio de comunidades de microorganismos se tiene el “Análisis de restricción de amplicones de DNA ribosomal”, el “Análisis de regiones intergénicas”, la “Electroforesis en geles desnaturalizantes de gradiente/Electroforesis en geles de gradiente de temperatura”, entre otros; este tipo de metodologías se fundamentan en las diferencias en la secuencia de nucleótidos; sin embargo, en su mayoría solo dan un perfil de la estructura de la comunidad y son útiles para detectar cambios pero no funcionan como un método para medir la diversidad o detectar grupos filogenéticos específicos (Nocker, *et al.*, 2007).

A diferencia de los métodos previos, la estrategia de amplificación y secuenciación de marcadores genéticos informativos de la taxonomía del DNA microbiano de una comunidad, denominado como “Análisis de marcadores genéticos o secuenciación de amplicones”, permite caracterizar la diversidad de las comunidades microbianas en términos de identidad y abundancia relativa (Guttman, *et al.*, 2014).

Los genes del RNA ribosomal de la subunidad pequeña y grande han sido usados como marcadores taxonómicos para la identificación de microorganismos (Kim, *et al.*, 2013). Se caracterizan principalmente por tener regiones altamente conservadas y variables, las regiones conservadas son empleadas para la generación de alineamientos y diseño de oligonucleótidos para la amplificación de fragmentos por PCR. La identificación de los microorganismos se lleva a cabo a partir de la comparación de las regiones variables que se hipotetiza son regiones privativas de cada organismo, y que permiten además encontrar relaciones evolutivas entre ellos (Kim, *et al.*, 2013; Van de Peer, *et al.*, 1996).

La subunidad pequeña ribosomal en bacterias es codificada por el gen 16S rDNA que está conservado y posee 9 regiones hipervariables que pueden ser usadas para distinguir entre diferentes microorganismos (Nguyen, *et al.*, 2002; Baker, *et al.*, 2003) mientras en eucariontes, el gen 18S rDNA contiene 8 regiones hipervariables y se ha usado como un elemento genético identificador. Aun cuando la resolución taxonómica de estos genes puede no ser siempre suficiente para identificar especie o cepa, sí pueden proveer información de la diversidad y dinámica de los microorganismos relacionados (Kim, *et al.*, 2013; Smith, *et al.*, 1999; Van de Peer, *et al.*, 1993).

El empleo de este tipo de técnicas en combinación con tecnologías de secuenciación masiva han permitido mejorar el conocimiento de la estructura y función de las comunidades de microorganismos además que permiten desarrollar múltiples análisis a través del uso de barcodes por lo que se incrementa la capacidad de análisis (Guttman, *et al.*, 2014; Whiteley, *et al.*, 2012).

Justificación

La expresión en injertos de limón de los antimicrobianos lisozima y defensina con la capacidad de transportarse a través del floema, es una estrategia desarrollada con anterioridad por el equipo de trabajo para combatir el Huanglongbing (HLB) en cítricos. Debido a la actividad inespecífica de estos antimicrobianos, es necesario evaluar la diversidad de las comunidades de microorganismos que interactúan en la rizósfera de los árboles, ya que los microorganismos son importantes para el desarrollo de las plantas además de ser indicadores de la salud del suelo como un elemento esencial de los ciclos biogeoquímicos.

Hipótesis

La presencia de antimicrobianos Lisozima y Defensina en el tejido vascular de los árboles de limón no modificará la diversidad de poblaciones microbianas de la rizósfera.

Objetivo General

Estudiar la dinámica de la población microbiana en la rizósfera de limón genéticamente modificado que expresa lisozima, defensina o ambos, comparado con plantas control.

Objetivos específicos

1. Construcción de las librerías 16S y 18S
2. Análisis de poblaciones a partir de los datos de secuenciación.

Materiales y métodos

Plantación Experimental

Como parte de las evaluaciones de riesgo establecidas en la LBOGM (Ley de Bioseguridad de Organismos Genéticamente Modificados), se estableció una plantación experimental en el Rancho el Pandelo y las Ánimas en el municipio de Tecomán Colima. Esta área agrícola es la principal productora de limón mexicano y tiene una extensión de 50,000 hectáreas de este cultivo.



Figura 2. Área de siembra experimental. Se muestra en verde el área de siembra con coordenadas 18°56'5.02"N y 103°55'34.88"O. Se puede apreciar hacia el norte, árboles de limón pertenecientes a otro predio, hacia el sureste árboles cultivados en el rancho y un área de invernaderos.

La plantación experimental se conforma por 80 árboles (20 árboles por tratamiento) distribuidos aleatoriamente en 5 filas de 16, con una separación de 5 metros entre columnas de árboles y de 2.5 m entre cada uno de ellos, como se muestra en la Tabla 1.



Figura 3. Colecta de muestras simples para la formación de muestras compuestas por el método de cuarteo.

Se dispone de tres tratamientos, los cuales son árboles de limón mexicano con espinas transformados para la expresión de lisozima o defensina fusionadas a *CsPP16*, injertados por reglamento en patrones de *Citrus volkameriana* (confiere resistencia al virus de la tristeza de los cítricos). Además de plantas control, se evaluaron tres tratamientos, correspondientes a plantas expresando lisozima, defensina y

lisozima-defensina.

La duración de la plantación experimental fue de 3 meses, de acuerdo a la autorización de siembra experimental expedida por un dictamen vinculante por la autoridad federal competente. En el periodo se realizaron 4 muestreos, uno al inicio de la plantación y el resto en los meses consecutivos denominados como t0, t1, t2 y t3.

Muestreo de suelo y tejidos

Suelo

Las muestras corresponden a muestras compuestas que se formaron por el suelo tomado de la rizósfera de cuatro diferentes árboles de un mismo tratamiento localizados en diferentes puntos del área de la plantación a partir de las muestras simples por el método de cuarteo, se mezclan partes iguales y se submuestra una porción del total (Figura 3. Toma de muestra). En total se disponen de 4 muestras correspondientes a cada uno de los tratamientos por evento de muestreo.

Tabla 1. Distribución de árboles de los diferentes tratamientos en la plantación experimental. Los números corresponden a los diferentes tratamientos Liozima(T1), Defensina(T2), Lis-Def(T3) y control(T4).

T 4	T 3	T 2	T 2	T 1
T 3	T 1	T 2	T 2	T 3
T 2	T 3	T 1	T 1	T 3
T 2	T 2	T 2	T 3	T 1
T 4	T 3	T 2	T 2	T 2
T 3	T 1	T 4	T 2	T 3
T 1	T 2	T 3	T 1	T 4
T 3	T 3	T 4	T 4	T 4
T 4	T 1	T 4	T 4	T 3
T 2	T 4	T 3	T 3	T 1
T 1	T 4	T 1	T 4	T 4
T 4	T 2	T 4	T 1	T 2
T 1	T 1	T 1	T 4	T 4
T 2	T 2	T 3	T 2	T 4
T 3	T 3	T 4	T 3	T 3
T 1	T 1	T 2	T 1	T 1

45 metros

2.5 m

5 m

35 metros

Las muestras de aproximadamente 500g se conservaron en bolsas de plástico y temperatura ambiente hasta su arribo al laboratorio.

En laboratorio, una parte de la muestra se conserva en tubos Falcon de 50 ml (Rotulados con la fecha, el lugar de muestreo, el tratamiento y el lugar de muestreo) a -70 °C y el resto se mantiene a 4 °C en cuarto frío.

Raíz

De los árboles muestreados se tomó raíz que se transportó bajo las mismas condiciones que las muestras de hoja, en laboratorio se realizaron lavados para reducir el material adherido a la raíz con agua destilada y se almacenó a -70° C.

Extracción de DNA genómico de muestras de suelo

La extracción de DNA de suelo y raíz para la construcción de las librerías se llevó a cabo con el kit de Mobio acorde a las indicaciones del proveedor con dos ligeras modificaciones:

- Agitación por 15 min posterior a la adición de la solución C1 con equipo TissueLyser de Qiagen a una frecuencia de 50Hz.
- Elución con 50 µl de H₂O_{mq} estéril en lugar de 100 µl.

Posterior a la extracción se evaluó la calidad del DNA por electroforesis en gel de agarosa al 1% y se cuantificó concentración por espectrometría en el equipo NANODROP (Thermo Fisher Scientific).

Extracción DNA de endófitos de raíz

Se siguió el protocolo de Kumar, *et al.*, (2015) para el análisis de endófitos, para lo cual se esterilizó la superficie de la raíz por lavados consecutivos de etanol al 70% por 3 min, hipoclorito de sodio 0.5% por 3 min y etanol al 70% por 30s. Se tomó una muestra de tejido y se evaluó el proceso en cajas Petri conteniendo medio LB agar, comparando el crecimiento entre muestras tratadas y no tratadas.

Para la extracción de DNA se trituró el tejido en mortero con nitrógeno líquido para congelar la muestra y romperla más fácilmente, las muestras trituradas se colocaron en tubos de 2 ml junto con un balón metálico y se aplicaron ciclos de 2 min en TissueLyser a 50 Hz congelando la muestra entre cada ciclo. Las muestras se mantuvieron en todo momento en hielo seco para evitar su descongelación. Se procede a realizar la extracción de DNA con el kit "PowerSoil DNA Isolation", de muestras compuestas de dos árboles por tratamiento, se cuantificó la concentración de DNA y se evaluó la calidad del DNA extraído en gel.

Construcción de librerías

Para el desarrollo de las librerías 16S y 18S se evaluó inicialmente la amplificación con oligos sin barcode para cada caso 16S (oligos CGO465F, CGO465R) y 18S. Una vez confirmada la amplificación descartando contaminación y ausencia de inhibidores de la reacción se procedió a trabajar con los oligos para secuenciación en Ion torrent.

Las muestras se amplificaron con los diferentes oligos con barcode en volúmenes de 50 µl por cuadruplicado de manera que al purificarse con Wizard® SV Gel en PCR Clean-Up System, se obtuviera un extracto puro en concentración y volumen adecuado para continuar con el proceso de secuenciación. Se purificaron los 200 µl concentrándolos a la mitad de su volumen y se mide su concentración en Nanodrop 2000.

A partir de la cuantificación de DNA de los productos de PCR purificados, se calculó el volumen necesario para tomar 1 µg de cada una de las muestras, este volumen se colocó en tubos limpios y se dispusieron para ser secuenciados por el Laboratorio de Referencias y apoyo a la Caracterización de Genomas, Transcriptomas, y Microbiomas del Departamento de Genética y Biología Molecular.

En el caso de la construcción de librerías 18S, previo a la amplificación de las muestras con oligos con barcode, se llevó a cabo una amplificación con oligos 18S (sin barcode) con el objetivo de obtener un molde de DNA inicial.

Análisis de datos de secuenciación

Análisis de calidad y depuración de datos

El tipo de archivo generado por el proceso de secuenciación es “fastq” y contiene además de las secuencias, información de la calidad con la que cada una de las bases fue secuenciada. Para procesar esta información los archivos se examinaron con el software FastQC de Babraham Bioinformatics (www.bioinformatics.babraham.ac.uk) que genera reportes a través de los cuales es posible evaluar la calidad y establecer parámetros para los procesos de pretratamiento de la información.

Tabla 2. Parámetros usados en TrimmomaticSE

Parámetro	Descripción
LEADING	Calidad mínima al inicio de la secuencia (<i>startq</i>).
TRAILING	Calidad mínima al final de la secuencia (<i>endq</i>)
SLIDINGWINDOW	Calidad mínima promedio (mean) de un determinado número de bases (<i>nbases</i>).
MINLEN	Longitud mínima (<i>xlen</i>)
HEADCROP	Elimina un número determinado de bases al inicio de las secuencias (<i>hcnbases</i>)
CROP	Corta las secuencias a una longitud dada (<i>cnbases</i>)

Dentro de los parámetros que evalúa el sistema se consideró el número de secuencias y longitud y la calidad de secuenciación por base al inicio y posterior a la depuración. Para depurar secuencias con bases de baja calidad, los archivos “fastq” se procesaron con TrimmomaticSE, a los parámetros **LEADING** y **TRAILING** se les asignó un valor de 20 en la escala de “phred33”, el valor de **SLIDINGWINDOW** se estableció en función del número de secuencias conservadas al final del análisis y al impacto en la mejora de la calidad. **MINLEN** para las secuencias 16S se estableció en 150 y en 200 para secuencias 18S. Los valores de **HEADCROP** y **CROP** fueron dependientes de las condiciones particulares de los datos dependiendo de la presencia de bases con mala calidad de secuenciación al inicio o al final de la secuencia.

Acondicionamiento de datos

Es necesario concatenar los archivos con las secuencias de las diferentes muestras en uno solo y generar además un archivo "Group", con el cual "Mothur" pueda reconocer las secuencias correspondientes a cada muestra. Para esto los archivos "fasta" se transfirieron a la plataforma en línea de **Galaxy** y se manipularon a manera de tablas como se muestra en la siguiente figura.

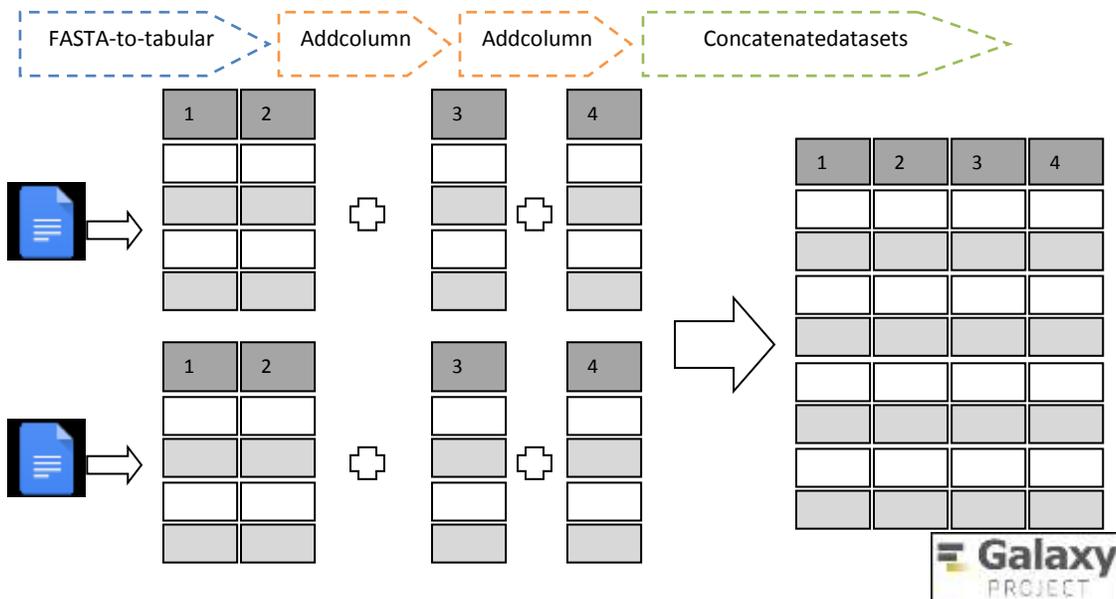


Figura 4. Flujo de trabajo en "Galaxy PROJECT" para la asignación de identificador de secuencia y grupo.

Las primeras dos columnas de la tabla corresponden al identificador original y la secuencia del archivo fasta, en pasos posteriores se insertaron dos columnas con la función "Addcolumn", en la primera columna insertada se creó un identificador para la secuencia y en la segunda un identificador de la muestra a la que pertenece. Con la columna creada como identificador de las lecturas y la columna de las secuencias se creó un archivo "fasta" y con el identificador de las lecturas y las muestras se creó un archivo de texto con formato tabular correspondiente al "Groupfile".

El análisis posterior se llevó a cabo de acuerdo a la metodología de Kozich JJ, *et al.*, (2013) en la plataforma de "Mothur", como se muestra en la Figura 5.

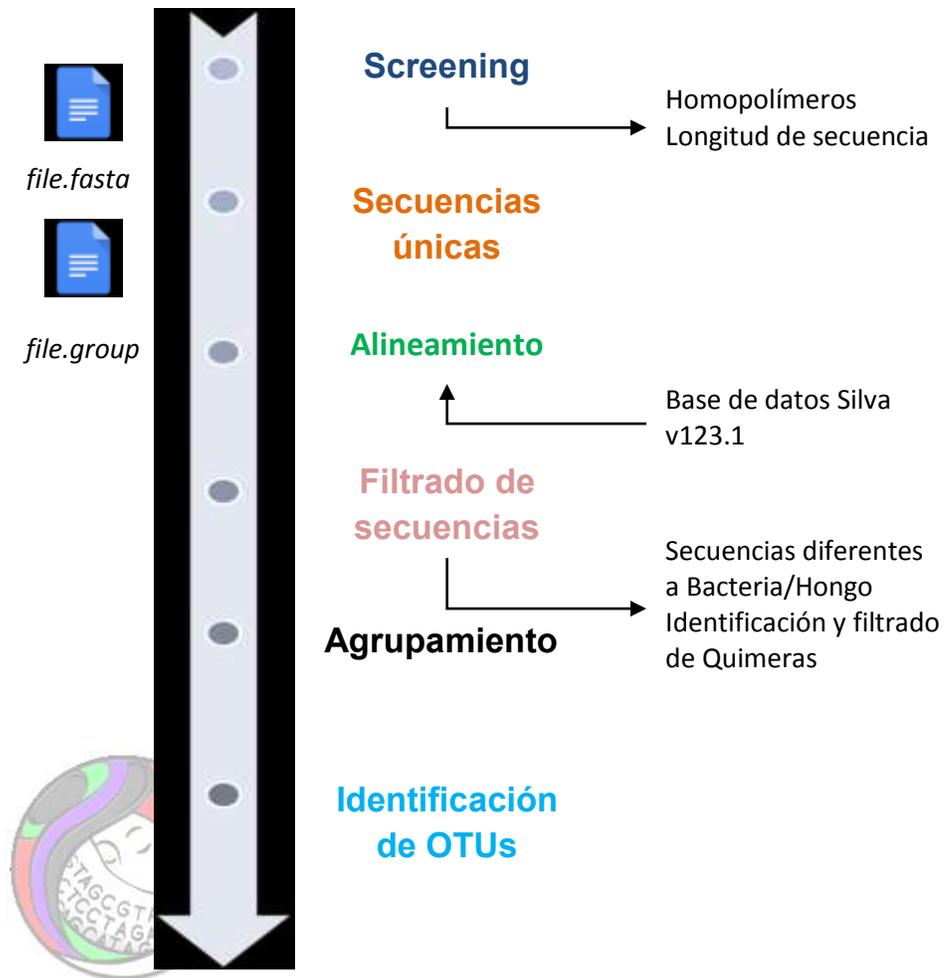


Figura 5. Flujo de trabajo para el análisis de datos con la plataforma de "Mothur".

Filtrado de homopolímeros y selección de secuencias únicas

Partiendo de los archivos generados en "Galaxy Project" se generó en "Mothur" un informe de las secuencias con la distribución de la longitud de las secuencias, la presencia de homopolímeros de bases y ambigüedades, para después eliminar homopolímeros con longitud mayor a 8 y eliminar secuencias que tuvieran ambigüedades.

Para reducir la cantidad de información a procesar se seleccionaron las secuencias únicas del archivo "fasta" y se generó un archivo "Count file" que mantiene la relación del número de lecturas para cada una de las secuencias identificadas.

Alineamiento de secuencias

El alineamiento se realizó con la base de datos de Silva en su versión 123.1, la cual contiene secuencias de la subunidad pequeña ribosomal y tiene un total de 172412 secuencias diferentes con una longitud de 50 000 columnas, el algoritmo de alineamiento empleado es “needleman”, con método de búsqueda por “kmeros”.

Para reducir el área de búsqueda y por tanto el tiempo de procesamiento, se seleccionaron 1000 secuencias problema, las cuales se alinearon con la base de datos completa, posteriormente el informe del alineamiento nos señaló las posiciones en las cuales se alineaba con la base de datos de referencia.

Para el caso de las secuencias 16S, se seleccionaron las posiciones 5000-15000 y se guardaron como un archivo nuevo y para las secuencias 18S de la posición 18000 -40000.

Una vez alineadas las secuencias se filtraron las columnas con gaps al inicio y al final, se optimizaron de manera que las secuencias compartan la misma posición de inicio y final y se determinaron secuencias únicas nuevamente.

De acuerdo al protocolo, considerando la presencia de errores de PCR y secuenciación se agruparon las secuencias por diferencia de dos bases.

Depuración de quimeras

Se eliminaron secuencias formadas durante la amplificación por polimerización de amplicones incompletos. Con el algoritmo de Uchime se identificaron secuencias quiméricas y posteriormente se filtraron.

Clasificación filogenética

Los archivos de referencia incluidos en la base de datos de Silva, constan de un archivo “fasta” con secuencias alineadas y un archivo de texto que incluye la taxonomía de las secuencias incluidas en el archivo fasta. Con estos dos archivos de referencia se asignó taxonomía a las secuencias

problema 16S y 18S. El archivo con las clasificaciones taxonómicas esta truncado a 6 niveles, lo cual es funcional para el análisis de las secuencias 16S pero no para el caso de las secuencias 18S por lo cual se requirió crear este archivo a partir del formato “ARB” de la base de datos de “Silva”, de manera que se dispusiera de la información taxonómica completa.

Para el caso del análisis de secuencias 16S, ya que este gen esta conservado en DNA de cloroplasto, mitocondria y Arqueas principalmente, se removieron todas aquellas secuencias diferentes a “Bacteria” y en el caso del análisis 18S aquellas con clasificación diferente a “Fungi”.

Las secuencias restantes se agruparon por el método de Phylotype, que genera tantas agrupaciones como niveles taxonómicos se dispongan, con lo que posteriormente se crearon las tablas de OTUs al nivel taxonómico “2” que corresponde a “Familia”.

Análisis posteriores como índices de diversidad y curvas de acumulación de especies requieren que las muestras tengan un número igual de secuencias por lo que se determinó el número de lecturas en cada uno de los grupos y se normalizaron al tamaño de la muestra más pequeña. Se determinaron entonces el índice de diversidad de Chao y Simpson y las tablas de acumulación de especies.

Se hizo uso de la plataforma de “R” para la manipulación de datos y generación de gráficos y del editor de imágenes “GIMP” para dar formato a las imágenes de esta tesis.

Resultados

Generación de Librerías

Posterior a la extracción de DNA de muestras de suelo se verificó su calidad en geles de agarosa como se muestra en la Figura 6 y 7. Se puede observar la presencia de una única banda por carril, un buen indicativo de la integridad del DNA extraído, adicionalmente la relación 260/280 está en la mayoría de los casos alrededor de 1.8 lo que muestra un buen nivel de pureza. En caso de presentarse impurezas como proteínas, esta relación da valores menores a 1.8.

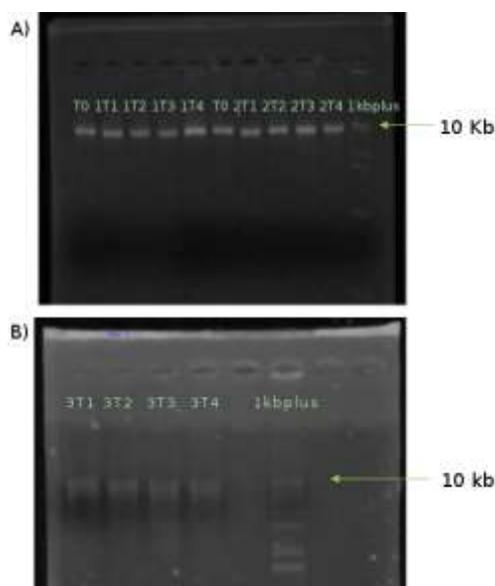


Figura 6. Electroforesis en gel de agarosa al 1% de DNA genómico de muestras de suelo. Los tratamientos se identifican como T1, T2, T3 y T4 anteceditos de un número que indica el tiempo de muestreo.



Figura 7. Electroforesis en gel de agarosa al 1% de DNA genómico de muestras de raíz. En el gel se analizan las extracciones de dos muestras distintas de raíz por tratamiento (T1-T2-T3-T4), la notación del tratamiento es antecedita de un número (1-2) usado para identificar entre ambas muestras.

Tabla 3. Cuantificación de DNA por espectrometría en Nanodrop 2000 (Thermo Fisher Scientific).

Muestra	[DNA] ng/μl	260/280
T0	4.6	1.65
1T1	20.3	1.6
1T2	7.9	1.73
1T3	10.1	1.72
1T4	13.1	1.79
2T1	14.4	1.86
2T2	9.6	1.86
2T3	11.9	1.84
2T4	10.4	1.84
3T1	13.8	1.82
3T2	15.1	1.73
3T3	12.4	1.87
3T4	11.8	1.8
1-T1	42.4	1.88
1_T2	51.4	1.87
1_T3	14.3	1.91
1_T4	14.3	1.97
2_T1	15.0	1.94
2_T2	54.1	1.89
2_T3	23.3	2.03
2_T4	14.6	1.95

Librerías 16S

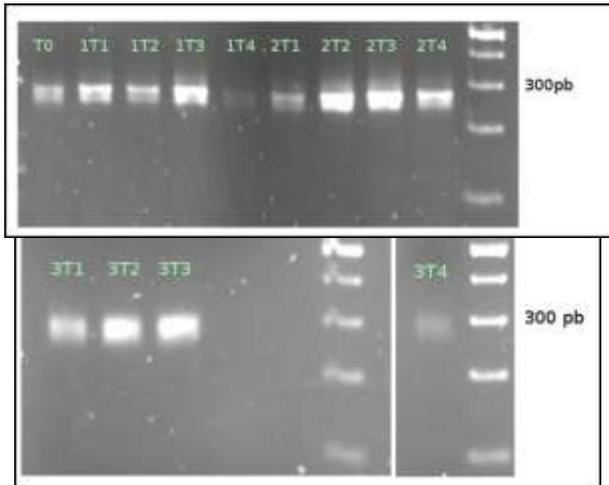


Figura 8. Electroforesis en gel de agarosa al 2% de productos de PCR 16SV3 de muestras de suelo. Se observa la banda esperada a 300pb aproximadamente.

Para la amplificación de la región V3 del gen 16S se usaron 10 ng de DNA en reacciones de 50 μ l. La banda de interés se visualizó en geles de agarosa al 2%, se observa en la Figura la banda deseada con un tamaño menor a 300pb. La región amplificada es un poco menor a 200; sin embargo, el adaptador y barcode de los oligos para secuenciación aumentan 100 pb el tamaño del fragmento.

Librerías 18S

De igual forma para la amplificación del fragmento 18S se usaron 10 ng de DNA en reacciones de 50 μ l, estos amplicones se utilizaron como plantillas para la generación de la librería con los oligos para secuenciación, en este caso la banda resultante es un poco mayor a 400 pb mientras que el tamaño de la región amplificada es cercano a 300 pb, en la Figura 9 se muestra la amplificación de la región con oligos 18S para secuenciación.



Figura 9. Electroforesis en gel de agarosa al 2% de productos de PCR 18S de muestras de suelo. Se observa la banda esperada por arriba de 400 pb.

Las muestras de raíz se procesaron de la misma manera que las de suelo, el resultado al igual que los descritos previamente consisten en una banda menor a 300pb para el fragmento 16S y mayor a 400pb para el 18S, se muestra en la Figura 10 las bandas obtenidas para cada uno de los casos.

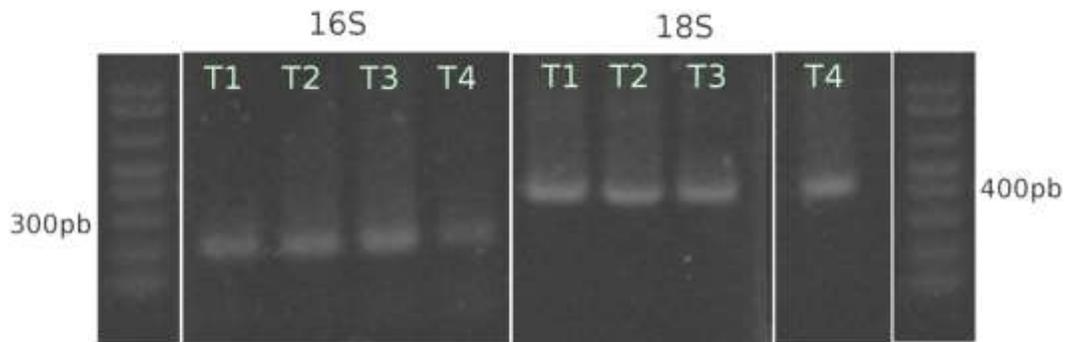


Figura 10. Electroforesis en gel de agarosa al 2% de productos de PCR 16S y 18S para muestras de raíz. Se observa la banda menor a 300 pb y mayor a 400 pb en reacciones 16S y 18S correspondientemente.

Estos productos de amplificación se acondicionaron y secuenciaron en Ion torrent por el Laboratorio de Referencias y apoyo a la Caracterización de Genomas, Transcriptomas, y Microbiomas del Departamento de Genética y Biología Molecular.

Análisis de datos de secuenciación de la región V3 del gen 16S

Las 13 muestras correspondientes a suelo cuentan con un total de 2'020,689 secuencias crudas a las cuales se les evaluó la calidad de secuenciación. Inicialmente los datos se procesaron con TrimmomaticSE usando los parámetros sugeridos en literatura de un promedio de calidad de 25 por cada 5 bases; sin embargo, esto hace que se pierda hasta el 90% de los datos. Evaluando diferentes condiciones la opción más viable resultó usar un SLIDINGWINDOW de 5:20 con una longitud mínima de 150 y una máxima de 200, esto permitió mantener más del 50% de los datos con valores de calidad buenos.

Tabla 4. Relación del cambio en el número de secuencias en los diferentes pasos del análisis de amplicones 16S.

Proceso	Suelo		Raíz	
	Secuencias únicas	Secuencias totales	Secuencias únicas	Secuencias totales
Alineamiento/Filtrado	367637	1070313	39758	827843
Precluster	81064	1004588	9211	802748
Eliminación de quimeras	75506	990436	8965	801902
Filtrado de linajes	75421	990077	7046	57891

Un reporte generado por “Mothur” de los datos procesados muestra un total de 1'087,341 secuencias totales, con un máximo de longitud de homopolímeros encontrados de 12, una cantidad no reportada en literatura para este tipo de secuencias por lo que al conservar aquellas con un máximo de 8 resulta en 1'070,313 secuencias de las cuales se identifican 367,637 secuencias únicas.

Referente a las muestras de raíz, se recibieron 1'297,093 secuencias correspondientes a 4 muestras, las cuales se procesaron en TrimmomaticSE con los mismos parámetros considerados para las muestras de suelo conservándose 835,246 secuencias. De ellas se redujeron a una cantidad de 827,843 secuencias debido a la eliminación de aquellas con homopolímeros mayores a una longitud de 8. En comparación con las muestras de suelo, solo se identificaron 39,758 secuencias únicas.

Estos datos se procesaron conjuntamente siguiendo la secuencia de pasos reportada en la metodología. Lo más notable del proceso resulta en la eliminación de secuencias no clasificadas como “Bacterias”; para ambos tipos de muestras la mayoría de estas secuencias filtradas

corresponden a cloroplasto y mitocondria; en el caso de las muestras de raíz este análisis reduce de 801,902 a 57,891 secuencias. Evaluando las causas de la reducción tan notable de secuencias en raíz, se remueven todas aquellas no clasificadas como “Bacteria” conservando además aquellas clasificadas como “Cloroplasto”, lo que resulta en 800,677 secuencias conservadas al final del proceso, esto da pie a concluir que el proceso se ha desarrollado correctamente y que en su mayoría las secuencias filtradas corresponden a cloroplasto.

Tabla 5. OTUs de bacterias identificados a diferentes niveles taxonómicos.

Label	1	2	3	4	5	6	7
numOtus	1432	565	261	114	42	2	1

Considerando que el archivo taxonómico dispone de 15 niveles, previo al agrupamiento filogenético las taxonomías resultantes para las secuencias se modificaron para contener solo 7 niveles.

Los niveles corresponden a 1-

Genero, 2-Familia,

3-Orden, 4-Clase,

5-Filo, 6-Dominio y

7-Root. Se

esperaría que el

nivel 6

correspondiente al

“Dominio” tuviera

un número de

OTUs igual a 1 ya

que solo deberían

de tenerse

“Bacterias”; sin

Tabla 6. Índices de diversidad de bacterias calculados a partir de datos normalizados. El índice S.obs muestra el número de OTUs observado en las submuestras y S.chao1 muestra una estimación de riqueza, con estos índices se calculó un porcentaje de cobertura y adicionalmente con S.obs se determinó cobertura por el método de Good.

	S.obs	S.chao1 (Estimado)	% Cobertura	Cobertura de Good	Simpson	Shannon
Liozima_root	173	241.077	71.761	0.9616	0.379	2.237
Defensina_root	229	303.286	75.506	0.9491	0.219	3.097
Lis-Def_root	178	235.417	75.611	0.9604	0.269	2.595
Control_root	171	256.800	66.589	0.9620	0.455	1.909
T0	328	349.577	93.828	0.9869	0.025	4.455
Liozima_t1	333	378.000	88.095	0.9867	0.033	4.408
Defensina_t1	332	430.077	77.195	0.9867	0.020	4.600
Lis/Def_t1	324	373.500	86.747	0.9870	0.021	4.574
Control_t1	330	383.040	86.153	0.9868	0.027	4.408
Liozima_t2	356	421.207	84.519	0.9858	0.021	4.609
Defensina_t2	324	369.000	87.805	0.9870	0.024	4.482
Lis/Def_t2	334	364.100	91.733	0.9866	0.025	4.417
Control_t2	349	383.500	91.004	0.9860	0.019	4.659
Liozima_t3	346	388.000	89.175	0.9862	0.019	4.616
Defensina_t3	334	373.808	89.351	0.9866	0.019	4.592
Lis/Def_t3	345	375.886	91.783	0.9862	0.019	4.603
Control_t3	325	355.875	91.324	0.9870	0.018	4.621

embargo, aún con el filtrado permanecieron dos secuencias clasificadas como “unclassified”.

Debido al alto número de OTUs identificados a nivel de género, la tabla de OTUs se generó a partir de la información del nivel 2 (Familia), datos con los cuales se generaron las curvas de rarefacción, b-diversidad, se graficó la abundancia relativa y se evaluó la similitud entre los tratamientos.

Para construir las curvas de rarefacción se normalizó el tamaño de las muestras de suelo a 25,000 que es el tamaño de la muestra más pequeña, en el caso de los datos de raíz al tener una menor cantidad de secuencias solo fue posible muestrear 4,500 lecturas. Este tipo de gráficos nos da información de la cobertura del muestreo y de la diversidad en las diferentes muestras.

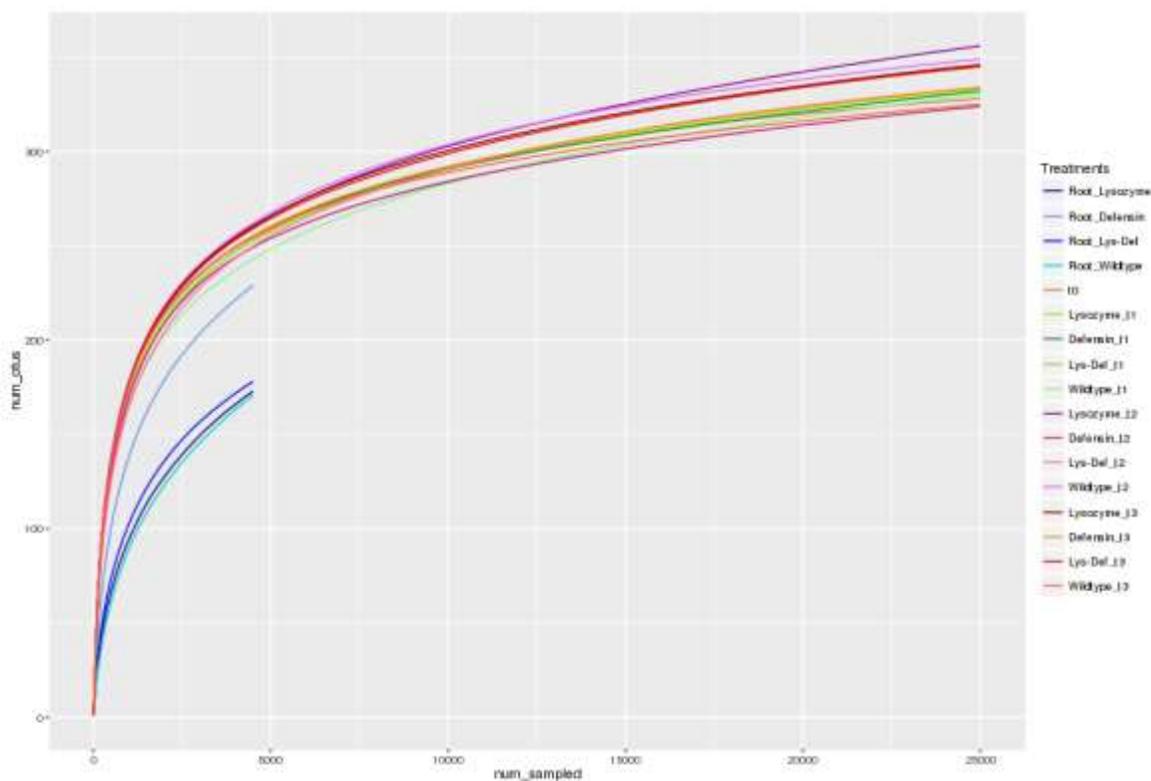


Figura 11. Curva de rarefacción para datos 16S. Se submuestrearon 25000 lecturas de muestras de suelo y 4500 de raíz debido a la diferencia en el número de lecturas de ambas poblaciones.

En general, se puede observar que las muestras de raíz tienen una menor riqueza comparada con las muestras de suelo, por otro lado las curvas en la mayoría de los casos tienden a acercarse a la asintota. Analizando las estimaciones de diversidad de población Chao1 y comparándolo con la observada podemos notar que se cubre aproximadamente un 90 % de la población del suelo y un 75% de la población de raíz, esto posiblemente debido a la baja cantidad de información que se dispone para estas muestras. En diversos trabajos se hace uso del índice de “Cobertura de Good” el cual tiene valores arriba del 90% en todos los casos como se puede observar en la Tabla 3.

Las principales diferencias que muestran los índices de diversidad (Simpson y Shannon) se pueden observar entre muestras de suelo y raíz siendo más bajo en los tejidos de raíz posiblemente atribuible al nivel de cobertura de éstas, aunque la estimación de Chao1 en estas muestras sigue siendo más baja que la estimada para suelos. Tanto en muestras de raíz como de suelo no se observa una reducción en la diversidad de los tratamientos los cuales tienen un comportamiento muy similar a los controles.

A nivel de filo o división se identificaron 42 grupos como se muestra en la Figura 12, de estos los 10 primeros (Otu0001-Otu0010) muestran ser los más abundantes. La distribución de abundancia es semejante dentro de cada tipo de muestra (Raíz, suelo), podemos observar que el Otu0002 correspondiente a "*Firmicutes*", es el más abundante en muestras de raíz y las proporciones de este y otros "Otus" son similares entre los diferentes tratamientos; por otro lado la dominancia del Otu0002 en raíz, no es comparable con las muestras de suelo donde su proporción se reduce considerablemente. El Otu0001 clasificado como "*Proteobacteria*" presenta una mayor abundancia en suelo, en comparación con las muestras de raíz además que otros grupos como el Otu0003, Otu0004, Otu0005, Otu0006, Otu0007 y Otu0008 tienen una mayor abundancia relativa respecto a las de raíz.

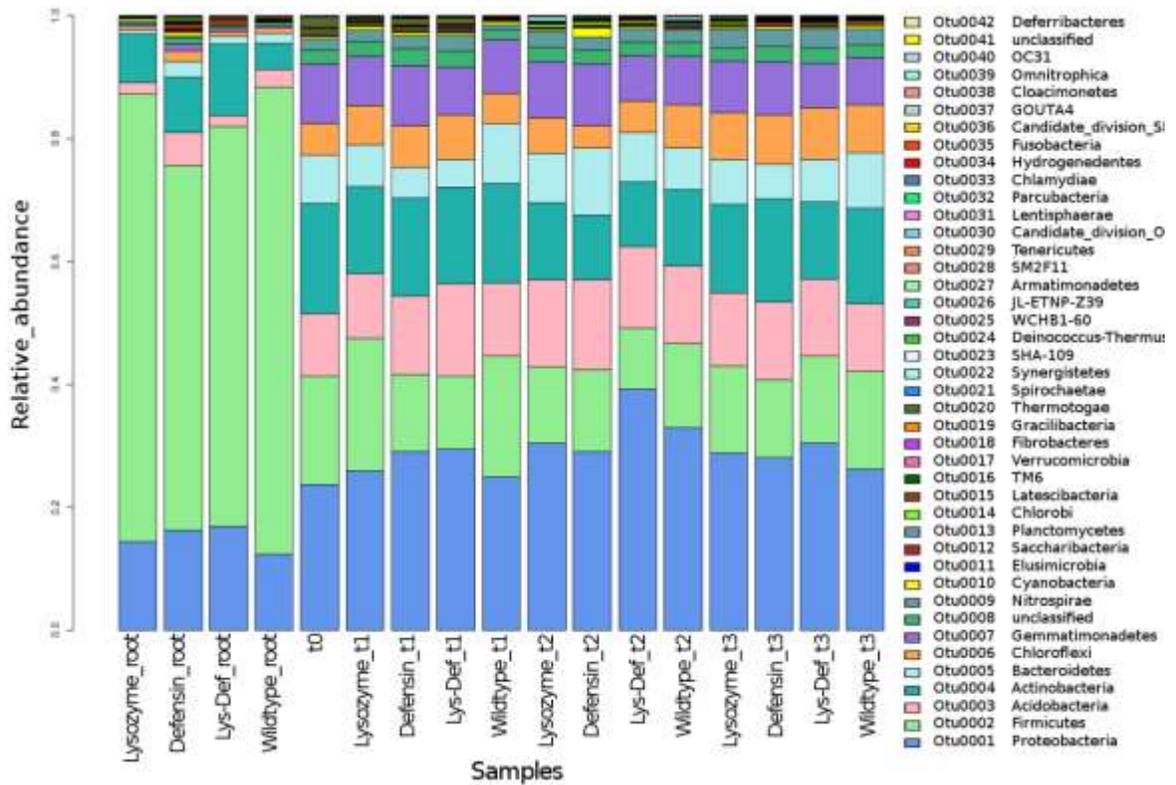


Figura 12. Abundancia relativa de bacterias agrupadas a nivel de "filo". Se observa como cambia la dominancia del Otu0001 y Otu0002 en muestras de suelo y raíz.

A nivel de familia se determinaron 565 Otus de los cuales en suelo 48 tienen una abundancia mayor al 1% en al menos una de las muestras, los más abundantes son el Otu0002 (*Unclassified*) y el Otu0003 (*Lachnospiraceae*) en rangos de 5 al 10% como se muestra en la Figura 5. En su mayoría los Otus tienen una abundancia menor al 5% y se pueden observar ligeras diferencias relacionadas con el nivel de abundancia.

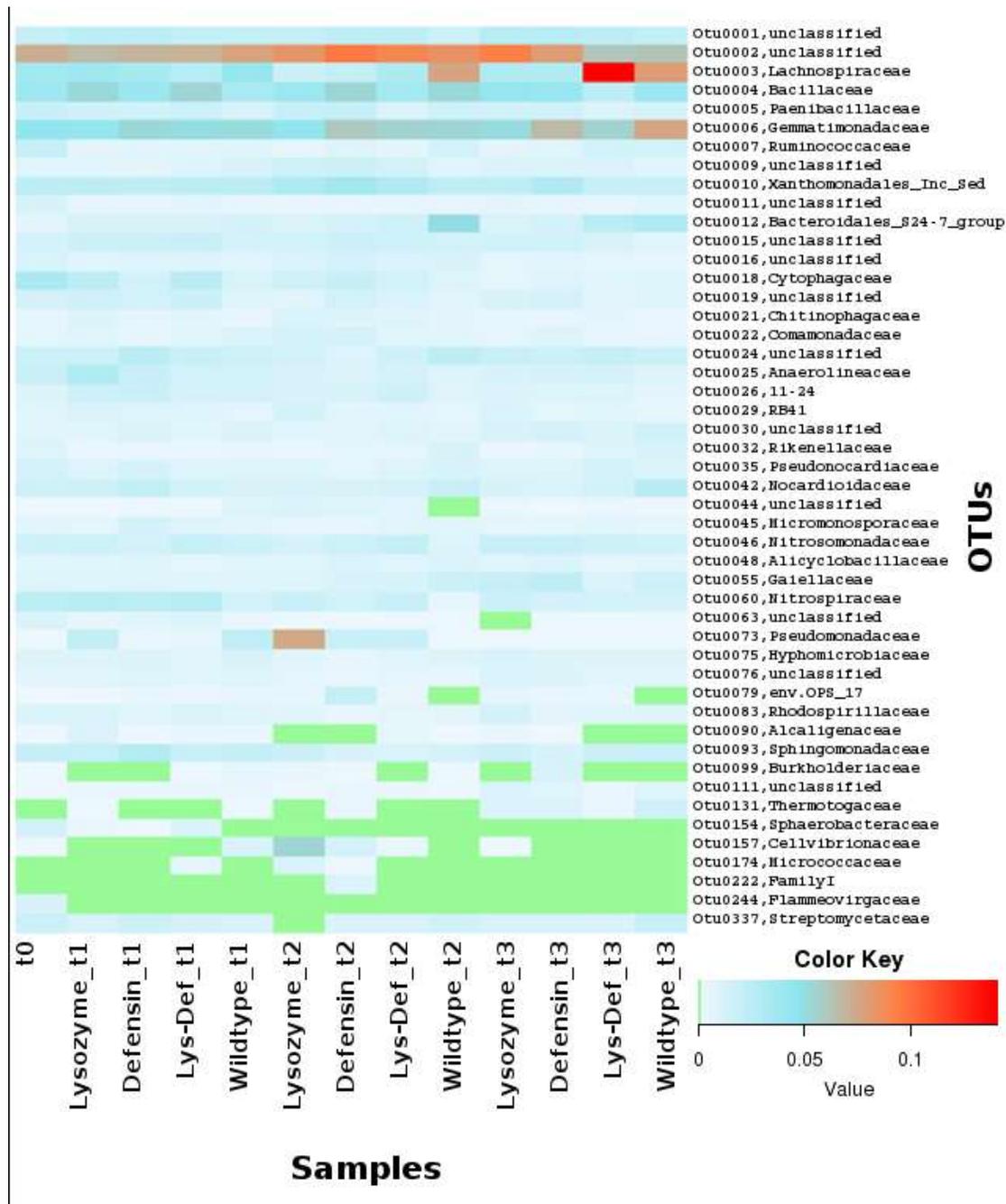


Figura 13. Distribución de abundancia relativa en suelo de OTUs de bacteria a nivel de "Familia". Los OTUs con mayor abundancia se encuentran en rangos del 5 al 10%.

En muestras de raíz el OTU más abundante es el Otu004 clasificado como "*Bacillaceae*" con una abundancia del 40 al 60%, seguido del Otu005 (*Paenibacillaceae*) con abundancia alrededor del 10%, ambas familias tienen una abundancia menor al 5% en muestras de suelo por lo que es notable su enriquecimiento en raíz. En comparación con muestras de suelo, los Otu0002 y Otu0003 parecen tener una abundancia menor en estas librerías.

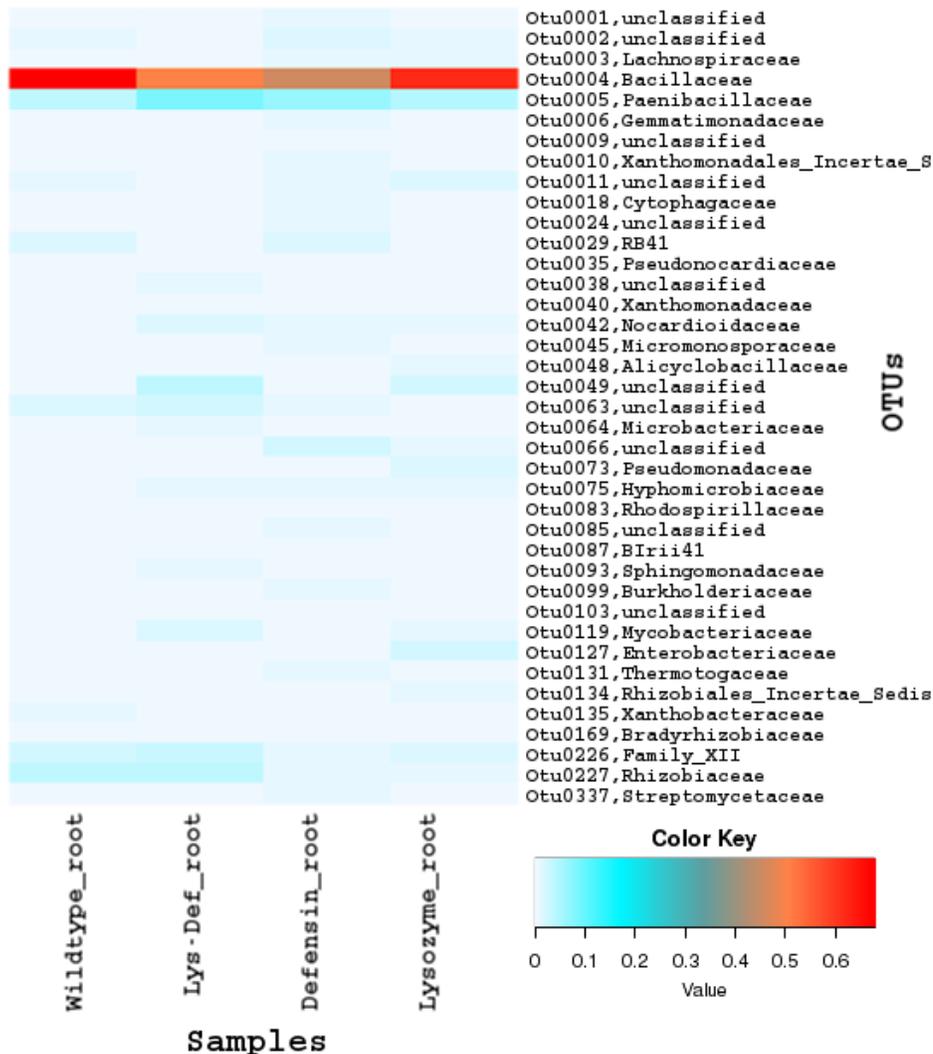


Figura 14. Distribución de abundancia relativa en raíz de OTUs de bacteria a nivel de "Familia". Hay una mayor representatividad del Otu004 y Otu005 con abundancias del 60 y el 20%.

Comparando la abundancia relativa de Otus de muestras cuyos tratamientos corresponden a árboles transgénicos con árboles silvestres, no se observa una diferencia en la distribución de los Otus además que los índices de diversidad calculados no muestran una disminución en la diversidad de las poblaciones, por lo que para evaluar la similitud entre las poblaciones y descartar diferencias asociadas a los antimicrobianos, se determinó la distancia entre las muestras a partir del método de Bray-Curtis.

El agrupamiento de las muestras a partir de la matriz de distancia muestra dos diferentes grupos a un nivel de disimilitud del 70% que corresponden a muestras de raíz y de suelo. Dentro de ambos grupos las diferencias entre tratamientos van de un rango del 10 al 25%. A un nivel de disimilitud del 20% aproximadamente se observa la formación de 6 diferentes grupos, 4 de los cuales agrupan muestras de suelo (C, D, E, F). El nodo “F” es el grupo de mayor tamaño y se conforma por muestras de los tiempos t1, t2 y t3, que se subdividen a un nivel de disimilitud cercano al 18% en un grupo con la totalidad de muestras del t3 y en otro grupo con muestras del t1 y t2. El grupo “D” por otro lado relaciona muestras del “t1” con el “t0”, y los nodos “C” y “E” corresponden a muestras que se separan de los grupos (F y D). Esto da indicios de una relación entre el tiempo y el cambio en la estructura de las poblaciones, además se observa que la magnitud de la diferencia que limita la formación de diferentes grupos es pequeña. Comparando las plantas control (Wildtype) con los diferentes tratamientos, estos parecen tener un comportamiento similar por lo que las diferencias que pudieran existir podrían estar relacionadas con la heterogeneidad del suelo.

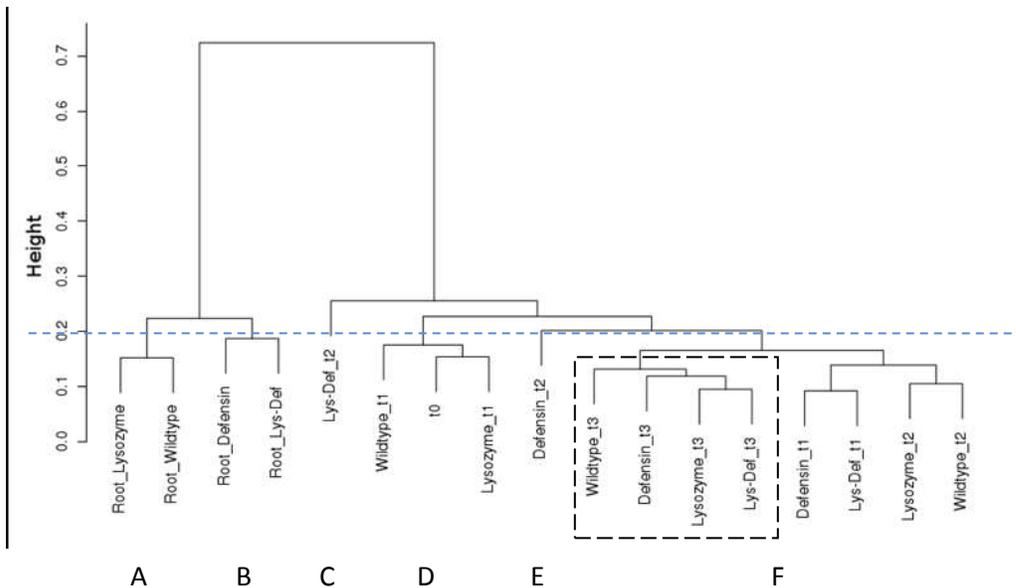


Figura 15. Análisis de disimilitud de poblaciones de bacteria a partir de la matriz de distancia generada por el método de Bray-Curtis. En el eje “y” se muestra el coeficiente de disimilitud y en el eje “x” se identifican con letras los grupos generados a un coeficiente de 0.2. Se señala en el rectángulo, el subgrupo generado en el nodo “F” que agrupa las muestras del “t3”, un posible indicativo de la relación del tiempo con el cambio en las poblaciones.

Tabla 7. Comparación de grupos con el algoritmo de UniFrac weighted, se muestra la disimilitud (WScore) y la significancia del análisis para cada par de datos de muestras 16S.

Groups	WScore	WSig
T1-T2	0.660108	<0.0010
T1-T3	0.629924	<0.0010
T2-T3	0.526869	<0.0010
T1-T4	0.637588	<0.0010
T2-T4	0.677818	<0.0010
T3-T4	0.677818	<0.0010
T1-root	1	0.006
T2-root	1	<0.0010
T3-root	1	<0.0010
T4-root	1	<0.0010
T1-t0	0.904102	<0.0010
T2-t0	0.881247	<0.0010
T3-t0	0.881247	<0.0010
T4-t0	0.937673	<0.0010
root-t0	1	0.274
t0-t1	0.881331	0.047
t0-t2	0.910935	0.049
t1-t2	0.764401	0.004
t0-t3	1	<0.0010
t1-t3	0.901275	<0.0010
t2-t3	0.829796	<0.0010

Para evaluar el efecto de las variables tiempo y tratamiento se asignaron las muestras a diferentes grupos de estudio y con el algoritmo de UniFrac weighted se obtuvo el grado de disimilitud así como el parámetro de correlación. Se observa que cuando las muestras se agrupan en función del tratamiento (T1-T2-T3-T4), las diferencias son significativas con un coeficiente de disimilitud alrededor de 0.6 entre los grupos, por otro lado cuando se establecen en función del tiempo (t0-t1-t2-t3), el coeficiente entre los grupos es en la mayoría de los casos mayor a 0.8 y significativos considerando un “ α ” de 0.05. Las mayores diferencias se observan entre el tiempo t0 y el t3 y comparando muestras de raíz con tratamientos (T1-T2-T3-T4-root).

Las diferencias parecen tener mayor relación con el tiempo dado que tienen un coeficiente mayor a los grupos formados

en relación al tratamiento, adicionalmente el análisis de varianza (amova, análisis de varianza molecular) señala que las poblaciones son significativamente diferentes (P-value < α) cuando las muestras son agrupadas en función del tiempo y que las poblaciones no son diferentes al relacionar las muestras por tratamiento.

Tabla 8. Tabla AMOVA Bacterias. Se evalúan los grupos de estudio “Tratamiento” y “tiempo”, siendo significativamente diferentes las agrupaciones definidas en función del tiempo utilizando un error de 0.008 (La corrección de Bonferoni para un “alpha” de 0.05).

T1-T2-T3-T4	Among	Within	Total	t0-t1-t2-t3	Among	Within	Total
SS	0.046105	0.171895	0.218	SS	0.128564	0.123011	0.25157
Df	3	8	11	df	3	9	12
MS	0.0153683	0.0214868		MS	0.0428547	0.0136679	
Fs:	0.715244			Fs:	3.13543		
P-value	0.843			P-value:	< 0.001*		

Análisis de datos de secuenciación del gen 18S

Se obtuvieron 1'360,141 secuencias resultantes del proceso de secuenciación que corresponden a trece muestras de suelo y cuatro de raíz. Para evaluar la calidad de las secuencias, estos archivos se procesaron en FastQc antes y después de ser depuradas en TrimmomaticSE para monitorear el proceso. Tomando en consideración las condiciones evaluadas en bacterias, se usaron los mismos parámetros (SLIDINGWINDOW:5:20) aumentando para este caso la longitud mínima a 280 (MINLEN:280). El resultado es una reducción en el 70% de la información debido a que el proceso de depuración corta aquellas secciones de baja calidad dejando secuencias muy cortas que posteriormente son filtradas al no cumplir el mínimo de longitud. Conforme se reduce el rango de calidad en el "SLIDINGWINDOW" se incrementa la longitud de las secuencias sin embargo aun con un mínimo de 200 únicamente se conserva el 50% de las secuencias. De los diferentes parámetros evaluados se optó trabajar con "SLIDINGWINDOW:4:17" y "MINLEN:200" con lo que se conserva del 40 al 60% de la información, adicionalmente se eliminaron 15 posiciones al inicio de las secuencias ya que la posición 7 y 15 presentan valores de calidad muy bajos.

Tabla 9. Relación del número de secuencias a lo largo de las diferentes etapas del análisis de amplicones 18S.

Suelo y raíz		
Proceso	Secuencias únicas	Secuencias totales
Alineamiento/Filtrado	720312	62978
Precluster	720312	21282
Eliminación de quimeras	720004	21162
Filtrado de linajes	621568	16764

En total se procesan 720,689 secuencias resultantes del proceso previo, las cuales se reducen a 720,312 posterior a la depuración de lecturas con homopolímeros mayores a 8, del total de lecturas 276,466 secuencias son únicas y con ellas se procede a continuar el análisis. Como se observa en la Tabla 9, el filtrado de secuencias no clasificadas como "Fungi" descarta aproximadamente 100,000 secuencias que corresponden a Eucariontes diversos.

Posterior a la clasificación se disponen de 15 niveles filogenéticos por lo que se acondiciona para contener solo 12 niveles (Ya que es la información disponible para este taxa). En su mayoría los hongos están clasificados dentro del subreino "*Dikaria*"; sin embargo, la división "*Glomeromycota*" aunque es una división polifilética al no pertenecer a este u otro subreino, erróneamente en el archivo taxonómico la división se encuentra en la posición correspondiente a subreino. Para

alinearse correctamente las clasificaciones en el nivel correspondiente, se insertó una posición en hongos de esta división (*Glomeromycota*) de manera que a nivel de subreino se clasificaran como “*Unclassified*”. Ésta corrección es importante ya que el paso posterior involucra un agrupamiento filogenético por lo que los niveles deben corresponder con la clasificación.

Tabla 10. Número de OTUs de hongos determinados a diferentes niveles filogenéticos.

Label	1	2	3	4	5	6	7	8	9	10	11	12
numOtus	96	72	48	26	14	10	4	1	1	1	1	1

En la Tabla 8, se muestra el resultado del agrupamiento filogenético para el cual los diferentes niveles corresponden a 1-Género, 2-Familia, 3-Orden, 4-Clase, 5-Subdivisión, 6-División, 7-Subreino, 8-Reino, 9-Subgrupo, 10-Clado, 11-Dominio y 12-Root. Los análisis de diversidad, similitud, cobertura, entre otros se llevaron a cabo a partir de la tabla de OTUs generada del nivel 2 correspondiente a “Familia”.

La Tabla 11, muestra diferentes índices de diversidad calculados; S.obs y S.chao1,

son indicativos de la riqueza de Otus y a partir de ellos se puede evaluar la cobertura del muestreo. En la mayoría de los casos se observa que el porcentaje de cobertura esta por arriba del 90%, mientras que el índice “Cobertura

Tabla 11. Índices de diversidad de hongos calculados a partir de datos normalizados. El índice S.obs muestra el número de OTUs observado en las submuestras y S.chao1 muestra una estimación de diversidad, con estos índices se calculó un porcentaje de cobertura y adicionalmente con S.obs se determinó cobertura por el método de Good.

	S.obs	S.chao1 (Estimado)	% Cobertura	Cobertura de Good	Simpson	Shannon
Liozima_root	24	26	92.308	0.9976	0.840	0.474
Defensina_root	18	23.25	77.419	0.9982	0.921	0.223
Lis-Def_root	27	27	100.000	0.9973	0.508	1.178
Control_root	34	35	97.143	0.9966	0.187	2.095
T0	33	33.25	99.248	0.9967	0.200	2.011
Liozima_t1	38	41	92.683	0.9962	0.181	2.120
Defensina_t1	32	35	91.429	0.9968	0.228	1.966
Lis/Def_t1	37	55	67.273	0.9963	0.193	2.058
Control_t1	38	45.5	83.516	0.9962	0.182	2.141
Liozima_t2	40	42.143	94.915	0.996	0.217	2.013
Defensina_t2	39	42	92.857	0.9961	0.172	2.161
Lis/Def_t2	40	47	85.106	0.996	0.146	2.258
Control_t2	33	33	100.000	0.9967	0.138	2.260
Liozima_t3	40	54	74.074	0.996	0.153	2.231
Defensina_t3	24	26.5	90.566	0.9976	0.602	0.867
Lis/Def_t3	33	40	82.500	0.9967	0.346	1.470
Control_t3	32	32.6	98.160	0.9968	0.229	1.795

de Good” asciende a 0.99. Los valores de porcentaje de Cobertura coinciden con el comportamiento de las “Curvas de rarefacción”, como ejemplo tenemos la muestra Root_defensin que tiene una cobertura del 77%, donde se puede notar que la curva sigue creciendo en

comparación a muestras como Lis-Def_root y Wildtype_t3 que tienen valores de 100 y 98% respectivamente y cuyas curvas muestran una saturación congruente con los valores calculados.

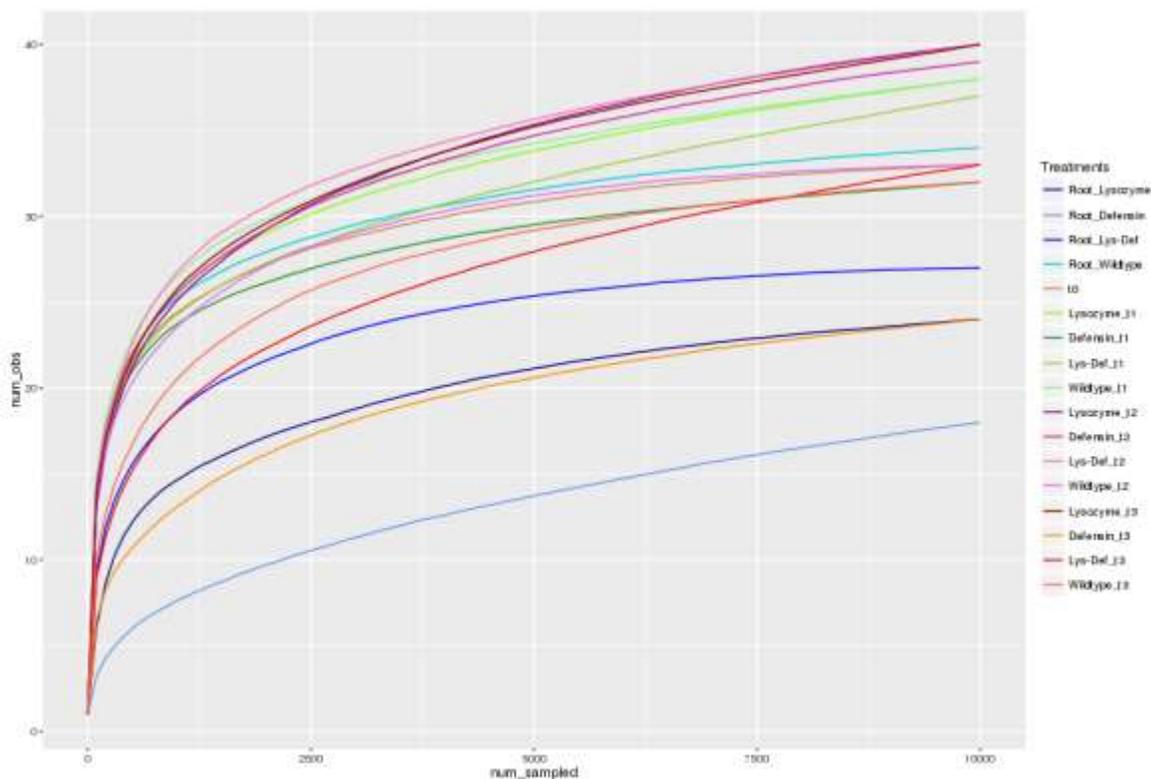


Figura 16. Curvas de rarefacción de librerías 18S en muestras de suelo y raíz.

En general, se puede observar menor riqueza en muestras de raíz que de suelo, a excepción de “Control_root” que presenta un valor muy alto similar al de suelo y “Defensina_t3” que por otro lado presenta una reducción considerable. Los índices de diversidad muestran de igual manera un bajo nivel en muestras de raíz “Lysozima_Root” y “Defensin_root” así como una reducción en los valores de muestras del tiempo 3.

Se identificaron 26 grupos filogenéticos a nivel de clase, el más dominante es el Otu001 clasificado como “*Sordariomycetes*”, en raíz tiene una abundancia relativa mayor al 95% en tratamientos de “Lisozima”, “Defensina” y “Lis/Def” mientras en el control tiene una abundancia cercana al 90% lo que hace que se incremente la abundancia de otros grupos y a su vez probablemente el nivel de diversidad de la muestra. Las muestras de suelo del tiempo 0 y tiempo 1 denominadas como “t0” y “t1” respectivamente, tienen un comportamiento muy similar al descrito en el control de raíz (Wildtype_root); sin embargo, la proporción de “*Sordariomycetes*” parece reducirse en el tiempo 2 y tiempo 3 de muestreo (t2 y t3 respectivamente), aumentando la abundancia del Otu02 (*Unclassified*) en el tiempo 2, el cual se reduce en el tiempo 3 donde se incrementa el Otu03 (*Unclassified*) como se observa en la Figura 17.

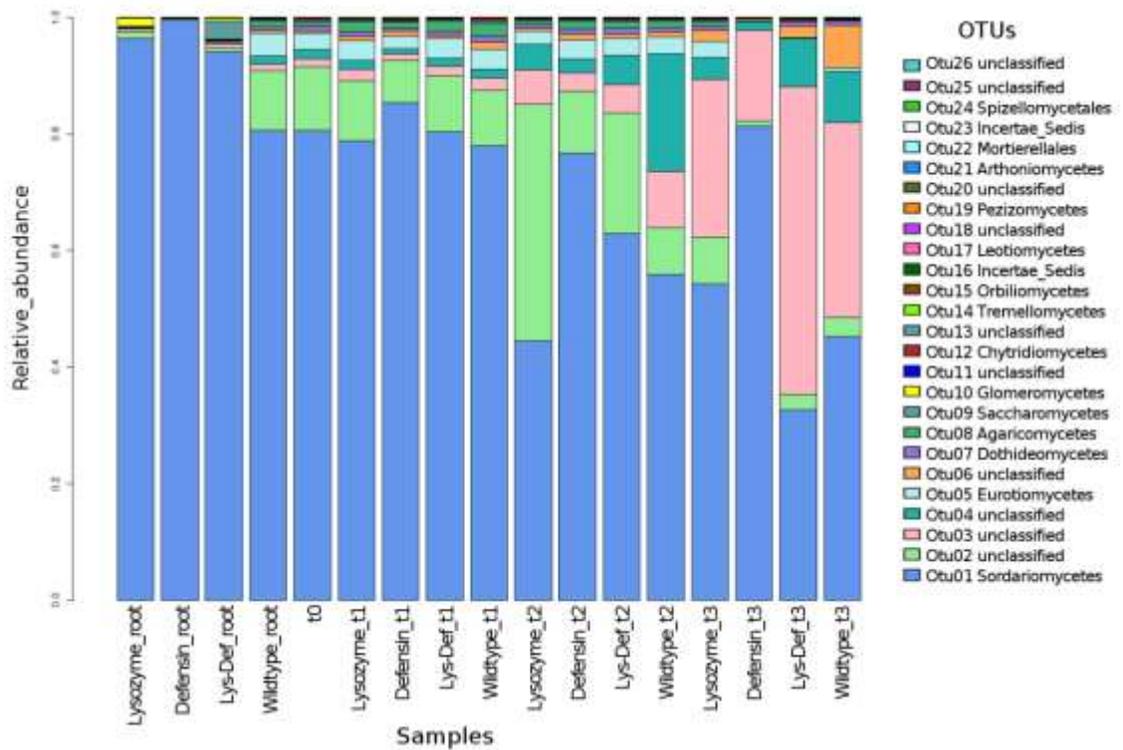


Figura 17. Abundancia relativa de hongos agrupados filogenéticamente a nivel de clase. En el eje “x” se muestran los diferentes tratamientos en muestras de raíz (root) y suelo a lo largo de los tiempos de muestreo (Tiempo 0, “t0”; Tiempo 1, “t1”; Tiempo 2, “t2” y Tiempo 3, “t3”).

Con los datos de la tabla de Otus definida a partir del nivel taxonómico de “Familia” se grafican los grupos que se encuentran presentes con una abundancia relativa de al menos 0.5%. El Otu10 clasificado como “*Nectriaceae*” es el que muestra una mayor abundancia en rangos del 70% al 80% en los tratamientos en tejido de raíz (Lysozyme_root, Defensin_root, Lys-Def_root) y Defensin_t3 de suelo, en el resto de las muestras incluido el control de raíz (Wildtype_root) su abundancia esta en un rango del 20 al 30%, adicionalmente el Otu01, Otu02, Otu06 y Otu12 están mayor representados con valores alrededor del 20% en muestras de suelo (Excepto Defensin_t3) y en el control de raíz (Wildtype_root) como se muestra en la Figura 18.

De forma similar al análisis de bacterias, las muestras de raíz a excepción del control (Wildtype_root) se separan del grupo de muestras de suelo con un coeficiente cercano a 0.7 mientras las diferencias entre muestras de suelo van de un rango de 0.1 a 0.6. Dentro del grupo “F”, se forma un subgrupo a un coeficiente de 0.1 aproximadamente que corresponde a muestras del “t1” y “t0” principalmente que como se observaba en la Figura 19 son muy similares unas de otras.

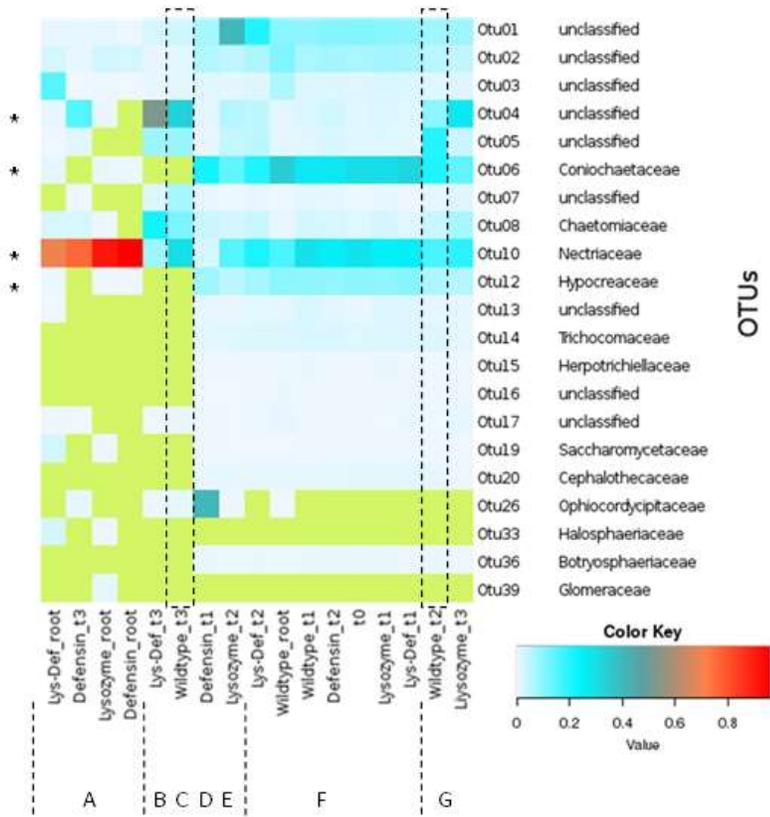


Figura 18. Abundancia relativa de OTUs de hongos definidos a nivel de Familia. En el eje X se señalan las muestras graficadas y la agrupación generada a partir del análisis de distancia. Los asteriscos muestran los OTUs que parecen tener mayor participación en la segregación de los grupos. Se marcan las muestras Wildtype_t2 y Wildtype_t3 como ejemplos del cambio de las poblaciones en el muestreo t3. El color verde señala valores de abundancia relativa cercanos a cero.



Figura 19. Análisis de disimilitud de poblaciones de hongos a partir de distancias calculadas por el método de Bray-Curtis. Se muestran los grupos generados con un índice de disimilitud menor a 0.3. El cuadro punteado señala un subgrupo en "F" con un índice cercano a 0.1.

La muestra Wildtype_root (Control de raíz) por otro lado se asemeja a muestras de suelo del “tiempo 1” como se observaba en la Figura 17, éste mismo gráfico permite señalar que posterior al “tiempo 1”, el cambio en las poblaciones sigue un comportamiento similar en los muestreos posteriores; sin embargo, las diferencias entre las muestras son más notorias, lo cual se refleja en la magnitud del coeficiente de disimilitud entre muestras del “tiempo 3” en los grupos “B” y “C”, del “G” y de la muestra de Defensina del tiempo 3 (Defensin_t3). De igual forma en la Figura 18 no se observan diferencias muy notables entre muestras del “tiempo 0”, “tiempo 1” y “tiempo 2”, pero en muestras del “tiempo 3” éstas se acentúan; los asteriscos en el diagrama señalan los OTUs donde se observa los cambios en abundancia más relevantes y que pueden por tanto definir la segregación de grupos. Se señalan también las muestras Wildtype_t2 y Wildtype_t3 para ejemplificar el cambio entre el “tiempo 2” y “tiempo 3”, en ellas se observa un ligero enriquecimiento del Otu10 y Otu04 mientras el Otu06 se reduce del 20% a valores cercanos a cero, este comportamiento es similar al de las muestras “Lys-Def” y “Defensin” en el “tiempo 2” y “tiempo 3” sin embargo, en la muestra de defensina (Defensin) el enriquecimiento del Otu10 es mucho más notable que en las anteriores.

Tabla 12. Comparación de grupos con el algoritmo de UniFrac weighted, se muestra la disimilitud (WScore) y la significancia del análisis para cada par de datos de muestras 18S.

Groups	WScore	WSig
T1-T2	0.585864	0.007
T1-T3	0.489128	<0.0010
T2-T3	0.562099	<0.0010
T1-T4	0.499547	<0.0010
T2-T4	0.557856	<0.0010
T3-T4	0.247099	<0.0010
T1-root	0.801968	<0.0010
T2-root	0.560197	<0.0010
T3-root	0.778203	<0.0010
T4-root	0.778203	<0.0010
T1-t0	0.700424	<0.0010
T2-t0	0.585864	0.042
T3-t0	0.471447	0.01
T4-t0	0.47544	<0.0010
root-t0	0.801968	<0.0010
t0-t1	0.329869	0.117
t0-t2	0.623719	<0.0010
t1-t2	0.482515	0.002
t0-t3	0.844605	<0.0010
t1-t3	0.844605	<0.0010
t2-t3	0.844605	<0.0010

En la Figura 19 se observan las relaciones que existen entre las muestras, como se analizaba tienen mayor segregación a un coeficiente un poco mayor que en el análisis de bacterias; sin embargo, la información de la Figura 18 muestra que hay similitud en el comportamiento con cambios de diferente magnitud. Al analizar la matriz de distancias con el comando “unifrac.weighted” de “mothur” se observa que las diferencias entre los grupos definidos en función del tratamiento son del alrededor de 0.5, cabe señalar que las diferencias entre tratamientos y raíz no son muy distintas como en el caso de bacterias (WScore = 1). En el caso de los grupos formados en base al tiempo de muestreo (t0-t1-t2-t3), el coeficiente de disimilitud es menor entre t0-t1-t2, sin embargo este se incrementa para el “t3”.

Para determinar el efecto de las diferentes variables, se establecieron grupos de estudio definidos por las variables “Tratamiento” y “tiempo” y se evaluaron con el análisis de varianza “amova”,

identificándose que hay diferencias significativas entre los grupos generados en función de la variable tiempo ($P\text{-value} < \alpha$) de manera que las poblaciones cambian a través del tiempo de manera igual sin una influencia aparente por la expresión de los antimicrobianos.

Tabla 13. Tabla AMOVA Hongos. Evalúa los grupos de estudio “Tratamiento” y “Tiempo”, siendo significativamente diferentes las agrupaciones definidas en función del tiempo utilizando un error de 0.008 (La corrección de Bonferoni para un “alpha” de 0.05).

T1-T2-T3-T4	Among	Within	Total	t0-t1-t2-t3	Among	Within	Total
SS	0.178879	1.22851	1.4073 9	SS	0.853116	0.604021	1.45714
df	3	8	11	df	3	9	12
MS	0.0596265	0.153564		MS	0.284372	0.0671134	
Fs:	0.388285			Fs:	4.23719		
P-value:	0.938			P-value:	0.003*		

Discusión

En relación a los resultados obtenidos de la secuenciación de fragmentos del gen que codifica la subunidad ribosomal 16S y 18S, las secuencias con valores de calidad de 20 o mayor en la escala de Phred33, fueron considerados de alta calidad, en concordancia con lo reportado por Quail, *et al.*, 2012. Por lo que aun cuando el valor se redujo a 25, que era el recomendado por el protocolo utilizado, las secuencias conservadas posteriores al filtrado en este estudio, siguen manteniendo una buena calidad. Contrario a plataformas de secuenciación como Illumina que obtiene niveles de calidad igual o mayor a 30, en IonTorrent se reportan alrededor de 20 (Quail, *et al.*, 2012), esto hace que bajo los mismos parámetros de evaluación, se conserve un mayor número de secuencias en Illumina, por lo que considerando el desempeño de las plataformas se sugiere usar parámetros de filtrado más laxos para IonTorrent con el objeto de no excluir organismos (Salipante, *et al.*, 2014). Tomando en cuenta que el filtrado conserva del 50 al 60% de secuencias 16S bajo un valor de calidad de 20, podría considerarse entonces reducir el valor para conservar un mayor número de lecturas como fue el caso en lecturas del gen 18S, donde el valor de calidad utilizado fue de 17.

En los diferentes niveles taxonómicos se tienen cierto número de OTUs no clasificados los cuales se incrementan conforme el nivel es más bajo como "Género". Esto es resultado de diferentes condiciones como la base de datos de referencia, Spain, *et al.*, (2009), estudia la microbiota de suelo proveniente de una granja en Oklahoma, y entre sus resultados observan que del 50 al 65% de los OTUs encontrados no están caracterizados a nivel de orden y familia, por lo que consideran que la diversidad es mucho más amplia que la representada en las bases de datos. Adicionalmente a estas observaciones Spain, *et al.*, (2009), cita un reporte de Janssen donde señala las diferencias en la asignación taxonómica en función del tamaño de las secuencias, donde la proporción de *Proteobacterias* cambiaba de 25-40% para secuencias mayores a 300pb a 42-50% para secuencias de tamaño aproximado a 100, por lo que Janssen, *et al.*, (2006) filtra secuencias menores a 300pb debido a que las asignaciones para este tipo de secuencias no son muy validas; el análisis del presente trabajo considero secuencias alrededor de 200 pb por lo que las taxonomías descritas así como sus proporciones pueden solo ser una aproximación de la composición del medio.

Para evaluar la calidad del muestreo, se llevó a cabo el análisis de rarefacción donde se normaliza el tamaño de las muestras, con lo que se construyen curvas de acumulación bajo el fundamento de que conforme el esfuerzo de muestreo es mayor, la curva tenderá a una asíntota como indicativo de que todos los organismos han sido observados (Gotelli, *et al.*, 2011; Hughes, *et al.*, 2001), esta representación permite observar cualitativamente el nivel de cobertura alcanzado en

el muestreo y compararlo además con el del resto. Como se describe en la Figura 11, se identificó una menor riqueza de bacterias en muestras de raíz dado que las curvas en este caso tienen una altura menor y un crecimiento más lento lo que Gotelli, *et al.*, (2011), asocia con un menor contenido de filotipos y una distribución menos uniforme contrario al crecimiento rápido de las curvas de muestras de suelo.

Para tener una imagen más cuantitativa de la cobertura, mediante la función “estimateR” del paquete “vegan” en R, se determinaron las especies observadas (S.obs), las estimadas (S.chao1), así como el indicador de “cobertura de Good”. Al calcular el porcentaje de cobertura con los datos de S.obs y S.chao1, se observan valores congruentes con las curvas de acumulación de especies a diferencia de la estimada por “cobertura de Good”; sin embargo, este índice está influenciado por la presencia de grupos muy abundantes dado que supone que si la mayoría de OTUs solo se han visto una vez, el número de OTUs no detectados podría ser alto por lo que “C” tendería a cero (Hughes, *et al.*, 2004), por tanto este indicador estaría sobreestimando la cobertura de muestreo.

$$C = 1 - \left(\frac{S}{N}\right)$$

Ecuación 1. Cobertura de Good. La ecuación muestra en “s” el número de OTUs o filotipos y como N el número total de lecturas.

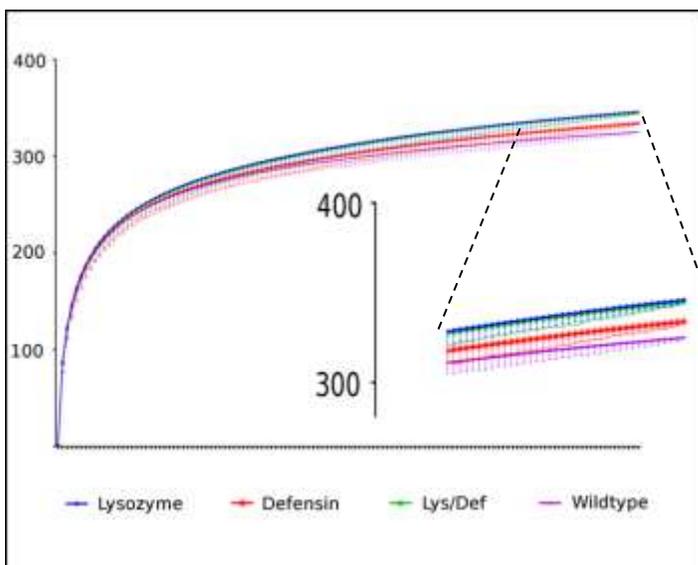


Figura 20. Curvas de acumulación de especies para datos de bacteria de la muestra “tiempo 3”, se incluyen los intervalos de confianza para comparar la similitud de las muestras. Se grafica en “x” el número de lecturas y en “y”, el número de filotipos observados (S.obs).

S.obs y S.chao1, son estimadores no paramétricos basados en Mark-release-recapture (MRR) que consideran la proporción de OTUs que fueron observados múltiples veces (Recapturados) en comparación con los que fueron observados solo una vez, esto se lleva a cabo por repeticiones aleatorias lo que permite definir intervalos de

confianza que pueden compararse con el resto de las muestras y definir si existen diferencias

significativas (Hughes, *et al.*, 2001). El análisis de muestras de bacterias en la Figura 20, permite observar que aún cuando las diferencias entre las muestras parecen ser muy pequeñas, el intervalo de confianza solo asocia la muestra de “Lysozyme” y “Lys-Def”, mientras que el IC del control y la muestra de defensina no se sobrepone con alguna otra muestra lo que denota bajo la afirmación de Hughes, *et al.*, (2001), que existen diferencias significativas entre las muestras; sin embargo, se asume que la distribución actual de la muestra está relacionada con la distribución real (Hughes, *et al.*, 2004). Considerando la heterogeneidad del suelo, las muestras compuestas pretenden disminuir la variabilidad, dado que muestras individuales pueden presentar mayor variación además de consumir tiempo y recursos (DOF, 2003), con ello se pretende entonces disminuir la variabilidad aleatoria relacionada con las características biológicas de las muestras (Sánchez-Pla & Ruíz de Villa, 2010.); sin embargo, sin replicas técnicas no se puede evaluar la precisión del análisis por lo que las diferencias observadas podrían ser producto del error experimental (Montgomery, 2004).

Analizando la evolución de los diferentes tratamientos con el tiempo, se observa un comportamiento similar entre datos de bacterias y hongos, a lo largo del estudio se presentan fluctuaciones en el número de S.obs de diferente magnitud (Figura 21) sin embargo el patrón entre muestras diferentes es muy similar.

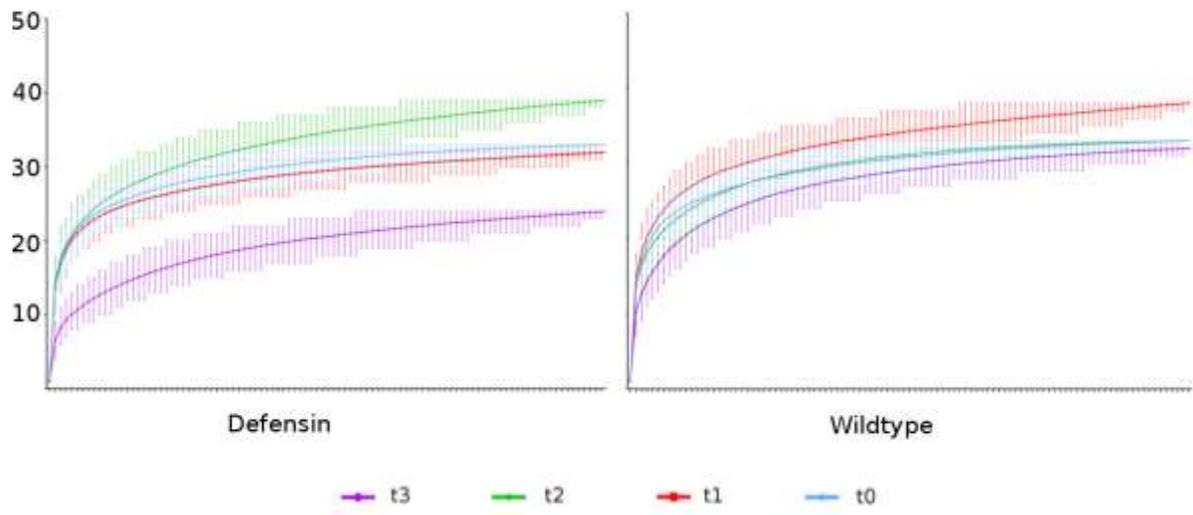


Figura 21. Curvas de acumulación de especies para datos de hongo del tratamiento "Defensin" y el control "Wildtype". Se incluyen los intervalos de confianza para comparar la similitud de las muestras. Se grafica en "x" el número de lecturas y en "y", el número de filotipos observados (S.obs).

De forma similar los estimadores de diversidad son no paramétricos y las múltiples repeticiones aleatorias permiten definir intervalos de confianza, existe también similitud entre algunas muestras considerando que los IC's se sobreponen, el análisis de varianza factorial rechaza las hipótesis nulas por lo que hay un efecto derivado de la variable "tiempo" y "tratamiento". De manera semejante al caso previo, la varianza de estos datos corresponde únicamente a la distribución de la muestra por lo que podrían no representar la población de la cual se extrajo. Los análisis de hongos muestran mayor heterogeneidad así como un nivel de diversidad más bajo.

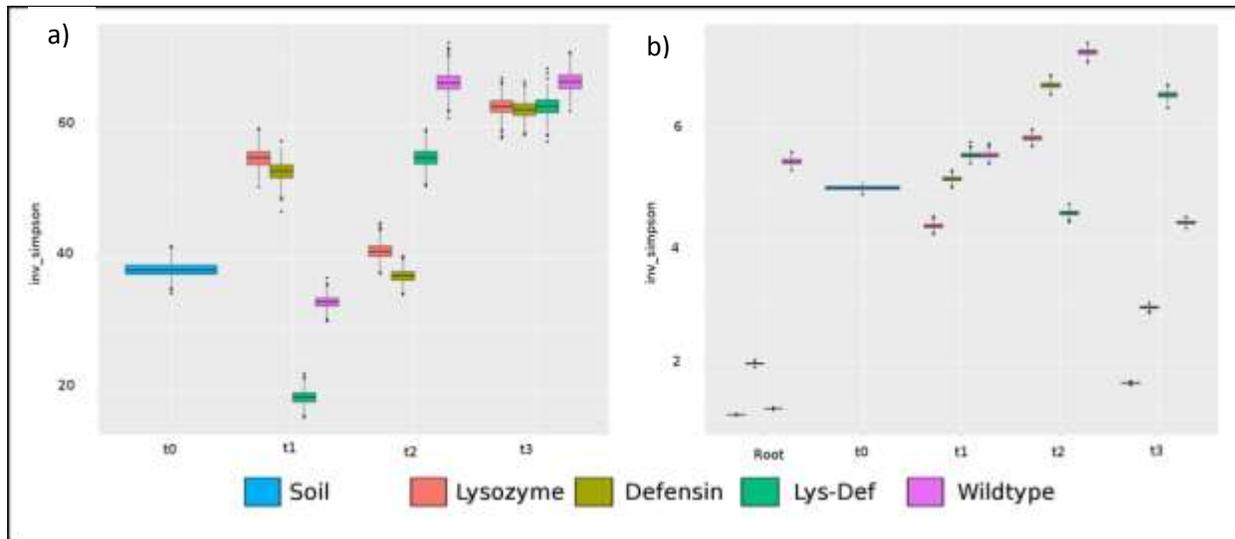


Figura 22. Comparación de diversidad a partir del índice Inversa de Simpson. Se muestra en a) datos de bacteria y en b) datos de hongos.

En comparación con los datos de S.obs, estos indicadores, Inv-Simpson, Shannon, entre otros; están influenciados por los componentes de uniformidad y riqueza por lo que aun cuando muestras como "Wildtype_t3" de bacterias poseen menor filotipos observados que otras muestras, su diversidad es mayor, el índice de Simpson es más sensible a la abundancia mientras Shannon a la riqueza (Hughes, *et al.*, 2004) por lo que las muestras de estos experimentos con riqueza similar tienen índice de Shannon muy similares en comparación con el índice de Simpson. Considerando que diversos estudios plantean que la capacidad de las plantas de moldear la microbiota en sus raíces se limita a modificar la abundancia de los organismos presentes, podría resultar más adecuado el índice de Simpson en la comparación de poblaciones. Se observa también el cambio rápido en las poblaciones de bacterias pasado el "tiempo 0" y como en el "tiempo 3", las diferencias son menores, en las comunidades de hongos el cambio parece ser más lento dado que en el "tiempo 1" la estructura es muy similar al "tiempo 0", por lo que de ser

cambios asociados a la selección de organismos en raíz podría esperarse que al igual que en bacterias, las muestras lleguen a tener una diversidad similar entre ellas.

Consistente con los taxos encontrados, Aislabie, *et al.*, (2013), menciona que de todos los filos entre los cuales están subdivididas las bacterias, 10 de ellos son los más abundantes en suelo, en mayor proporción se encuentran las *Proteobacterias*, *Acidobacterias* y *Actinobacterias*, mientras en menor proporción están las *Verrucomicrobia*, *Bacteroidetes*, *Firmicutes*, *Chloroflexi*, *Planctomycetes* y *Gemmatimonadetes* principalmente; los primero 5 filos enlistados en conjunto comprenden en promedio el 92% de las librerías en suelo (Janssen, 2006).

Estos organismos cumplen diferentes funciones en el ambiente, las *Proteobacterias* por ejemplo están implicadas en la fijación de nitrógeno, oxidación de metano, oxidación de nitrógeno y amonio, reducción de sulfato y descomposición de materia orgánica entre otros, las *Acidobacterias* reducen nitrato y nitrito.

Estos puntos denotan las condiciones de salud del suelo aun con la presencia de plantas OGM, la lisozima al atacar la pared celular es más efectiva con bacterias gram positivas mientras la defensina con bacterias gram negativas, Aislabie, *et al.*, (2013) menciona que los microorganismos con pared celular gram positiva tienden a ser más abundantes en suelo, entre ellos los representativos son las *Actinobacterias* y *Firmicutes* mientras por otro lado las *Proteobacterias* representan el grupo de bacterias gram negativas, ninguno de los cuales se vio alterado en comparación con el control.

Por otro lado la dominancia en raíz de *Fusarium sp.*, ejemplifica la ausencia de un efecto negativo, debido al tratamiento de los árboles, ya que se ha reportado la eficacia de las defensinas contra hongos como *Fusarium*, *Phytophthora*, *Phaesoripsis*, entre otros, así como muchos patógenos de humanos, plantas e insectos, además de la efectividad que tienen en el control de hongos en plantas que expresan constitutivamente este péptido (Swathi, *et al.*, 2008; Aerts, *et al.*, 2007). En diferentes estudios se describe *Fusarium sp.*, como un agente infeccioso causante de distintos padecimientos en plantas; sin embargo, naturalmente existen suelos que suprimen el desarrollo de la enfermedad causada por el hongo debido a características físicas del suelo y o factores bióticos (Larkin, *et al.*, 1996), en Chateaufrenard Francia y en Salinas Valley en California, la supresión es atribuible a *F. oxysporum* no patógeno y *Pseudomonas fluorescens* respectivamente (Amir & Alabouvette, 1992; Larkin, *et al.*, 1996), a su vez se ha comprobado que el empleo de estas cepas como agentes de biocontrol ya que reducen el desarrollo de enfermedades generadas por cepas de este hongo (Lemanceau & Alabouvette, 1991).

Para evaluar si la distribución está influenciada por efectos estocásticos, se puede aplicar la estrategia de Bootstrapping no paramétrico, donde réplicas de muestreos con remplazo permite determinar la consistencia de los resultados. La frecuencia con la cual una rama es encontrada en la misma configuración representa el puntaje de Bootstrap (Makarenkov, *et al.*, 2010).

A diferencia del diagrama generado a partir de la tabla completa de OTUs de Bacterias, el submuestreo de lecturas y construcción de árbol consenso mejora la asociación de las muestras considerando que el factor que muestra un mayor impacto es el tiempo; sin embargo, en hongos continúa teniendo un comportamiento aleatorio, que de acuerdo a los datos observados en el mapa de calor, son diferencias que parecen estar asociados a la abundancia de un pequeño grupo de OTUs específicos.

El análisis AMOVA, es un método para evaluar patrones y relaciones determinadas por análisis de agrupamiento que estima la diferencia entre poblaciones a partir de la varianza calculada entre y dentro de los grupos, es aplicada al estudio de la estructura genética de poblaciones y al estudio de variación alélica, supone que dos poblaciones evolucionan de la misma forma, sin embargo ciertas características pueden ser seleccionadas en función de diferentes factores. Aplicado al análisis de poblaciones, las muestras se asignan a diferentes grupos de estudio, en este caso “tiempo de muestreo” y “tratamiento”, a lo que el análisis sugiere que no hay una diferencia significativa entre grupos del diseño “tratamiento”; sin embargo, sí las hay en el diseño “tiempo”. Considerando el cambio en las características del suelo como producto de la introducción de los árboles, el cambio en las poblaciones estaría relacionado con la selección de microorganismos en la rizósfera, un proceso que es similar en todos los tratamientos.

En general este tipo de análisis de poblaciones por secuenciación de amplicones, están limitados por el método de extracción de DNA ya que pueden sesgar el resultado sobreestimando o subestimando ciertos grupos, ya que cada protocolo facilita la extracción de una parte de la población diversa (Delmont, *et al.*, 2011), por otro lado la abundancia de productos de PCR puede no reflejar la abundancia relativa de las cadenas molde de DNA debido a las diferencias de unión del primer y a la eficiencia de elongación (Hughes, *et al.*, 2004), en un estudio de Tanabe, *et al.*, (2015), evalúa la amplificación de 3 regiones del gen 18S, entre sus resultados explica que las regiones con mayor variabilidad, en este caso V1-V3, genera un menor tasa de amplificación en comparación con regiones de menor variabilidad, además que la presencia de secuencias erróneas puede resultar en la sobreestimación del número de OTUs (Chen, *et al.*, 2013).

Conclusiones

Se evaluó el efecto de la expresión de antimicrobianos en árboles de limón mexicano sobre las comunidades de microorganismos que habitan en el suelo agrícola. Considerando que aproximadamente la mitad de los datos de secuenciación son filtrados debido a la presencia de bases con baja calidad, podría considerarse aceptable disminuir los rangos de evaluación de manera que se conserve una mayor cantidad de información; sin embargo, debe considerarse que la mayoría de las asignaciones erróneas provienen de secuencias de baja calidad además que estas a su vez incrementan el número de OTUs sobreestimando el valor real.

Con base en los indicadores S_{obs} y S_{chao1} se determina que la cobertura de filotipos en las diferentes muestras es de aproximadamente del 90%, la cobertura en muestras de raíz fue más baja y se observa que sus curvas de acumulación de especies, tienen un crecimiento más lento indicativo de baja diversidad.

Ya que los cambios esperados por la interacción de las raíces con las comunidades del suelo, involucran un cambio en la abundancia de los microorganismos, es de mayor utilidad la información del índice de Simpson. Considerando la información de S_{obs} , e indicadores de diversidad, se observa que el comportamiento de los tratamientos es muy similar y no se presenta disminución en diversidad que pueda ser asociado a la presencia de los antimicrobianos.

Comparando la estructura de las poblaciones a partir de la distancia de Bray-Curtis, el agrupamiento de de las muestras no presenta una relación en función del tratamiento sino con la temporalidad de las muestras. En hongos el agrupamiento es muy aleatorio además que el nivel de disimilitud es mayor al presentado en bacterias, observando la estructura de hongos en el mapa de calor, se aprecia que las diferencias de los grupos pueden estar asociadas a la abundancia de ciertos grupos taxonómicos.

Los filotipos encontrados en este estudio presentan mucha similitud con las descripciones de poblaciones de microorganismos en estudios reportados, además que parece no haber una alteración en los grupos de bacterias gram positivas y gram negativas en ninguno de los tratamientos. Por otro lado la alta colonización de *Fusarium* sp. en las raíces de los árboles de limón descarta un efecto negativo de los tratamientos sobre el establecimiento de poblaciones.

El análisis de varianza AMOVA rechaza la hipótesis nula de que las agrupaciones de las muestras en base al tratamiento son diferentes entre ellas, en comparación con agrupaciones en función del tiempo de muestreo que sí presentan diferencias significativas entre grupos, diferencias que podrían estar asociadas a la selección de los microorganismos en las rizósfera de los árboles de limón.

Perspectivas

Incluir réplicas biológicas para mejorar la inferencia y comparación de indicadores.

Extender el tiempo de muestreo para el análisis de hongos de manera que pueda ser evaluado si aumenta la diferencia entre tratamientos.

Evaluar la presencia de los antimicrobianos en la raíz.

Bibliografía

- Acuerdo por el que se dan a conocer las medidas fitosanitarias que deberán aplicarse para el control del Huanglongbing (candidatus liberibacter spp.) y su vector. (2010). Diario Oficial de la Federación.
- Aerts, A. M., Thevissen, K., Bresseleers, S. M., Sels, J., Wouters, P., Cammue, B. P. A., & François, I. E. J. A. (2007). *Arabidopsis thaliana* plants expressing human beta-defensin-2 are more resistant to fungal attack: Functional homology between plant and human defensins. *Plant Cell Reports*, 26(8), 1391–1398.
- Aislabie, J., & Deslippe, J. R. (2013). Soil microbes and their contribution to soil services. In *Soil Microbial Diversity* (pp. 143–161). New Zealand.
- Amir, H., & Alabouvette, C. (1993). Involvement of soil abiotic factors in the mechanisms of soil suppressiveness to fusarium wilts. *Soil Biology and Biochemistry*, 25(2), 157–164.
- Arredondo-Bernal, H., Sánchez-González, J., & Mellín-Rosas, M. (2013). Taller Subregional de Control Biológico de *Diaphorina citri*, vector del HLB. Panamá: SENASICA.
- Aubert, B. (1992). Citrus greening disease, a serious limiting factor for citriculture in Asia and Africa. *Proc. Int. Soc. Citriculture*. Montpellier Cedex.
- Bais, H. P., Weir, T. L., Perry, L. G., Gilroy, S., & Vivanco, J. M. (2006). The role of root exudates in rhizosphere interactions with plants and other organisms. *Annual Review of Plant Biology*, 57, 233–266.
- Baker, G. C., Smith, J. J., & Cowan, D. a. (2003). Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods*, 55(3), 541–555.
- Bennett, A. B., Chi-Ham, C., Barrows, G., Sexton, S., & Zilberman, D. (2013). Agricultural Biotechnology: Economics, Environment, Ethics, and the Future. *Annual Review of Environment and Resources*, 38(1), 249–279.
- Carme Ruíz de Villa, M., & Sánchez-Pla, A. (2010). *Análisis de Datos de Microarrays*. Barcelona: Universitat Oberta de Catalunya.
- Chaparro, J. M., Badri, D. V., & Vivanco, J. M. (2014). Rhizosphere microbiome assemblage is affected by plant development. *The ISME Journal*, 8(4), 790–803.
- Chassy, B. (2002). The Role of Agricultural Biotechnology in World Food Aid. *Economic Perspectives*, 8(3), 20–22.
- Chassy, B. M. (2007). The history and future of GMOs in food and agriculture. *Cereal Foods World*, 52(4), 169–172.
- Chen, W., Zhang, C. K., Cheng, Y., Zhang, S., & Zhao, H. (2013). A Comparison of Methods for Clustering 16S rRNA Sequences into OTUs. *PLoS ONE*, 8(8).
- Clark, N., Mugabe, J., & Smith, J. (2014). GOVERNING AGRICULTURAL BIOTECHNOLOGY IN AFRICA. Building Public Confidence and Capacity for Policy-Making. African Centre for Technology Studies (ACTS)
- Cotta, S. R., Dias, A. C. F., Marriel, I. E., Gomes, E. A., Van Elsas, J. D., & Seldin, L. (2013). Temporal dynamics of microbial communities in the rhizosphere of two genetically modified (GM) maize hybrids in tropical agrosystems. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 103(3), 589–601.

- Delmont, T. O., Robe, P., Cecillon, S., Clark, I. M., Constancias, F., Simonet, P., ... Vogel, T. M. (2011). Accessing the soil metagenome for studies of microbial diversity. *Applied and Environmental Microbiology*, 77(4), 1315–1324.
- Düring, K. (1993). Can lysozymes mediate antibacterial resistance in plants? *Plant Molecular Biology*, 23(1), 209–14.
- Excoffier, L. (1995). AMOVA 1.55 (analysis of molecular variance). University of Geneva, Geneva, (1), 1–6.
- FAO. (2016). Climate is changing. Food and agriculture must too. Retrieved August 1, 2016, from <http://www.fao.org/world-food-day/2016/theme/en/>
- Ganz, T. (2004). Defensins: Antimicrobial peptides of vertebrates. *Comptes Rendus - Biologies*, 327(6), 539–549.
- Garcia-Armisen, T., Papalexandratou, Z., Hendryckx, H., Camu, N., Vrancken, G., De Vuyst, L., & Cornelis, P. (2010). Diversity of the total bacterial community associated with Ghanaian and Brazilian cocoa bean fermentation samples as revealed by a 16 S rRNA gene clone library. *Applied Microbiology and Biotechnology*, 87(6), 2281–2292.
- García-Lara, S., & Bergvinson, D. J. (2007). Programa integral para reducir pérdidas poscosecha en maíz. *Agricultura Técnica En México*, 33(2), 181–189.
- Gasman, F., Ocampo, H., Alonso, S., & García, S. (2009). La bioseguridad en México y los organismos genéticamente modificados: cómo enfrentar un nuevo desafío. *Estado de Conservación Y Tendencias de Cambio*. México: CONABIO.
- Gómez-Jaimes, R., López-Arroyo, J. I., Hernández-Fuentes, L. M., Urias-López, M. A., Velázquez-Moreno, J. J., & Borrayo-Zepeda, A. (2011). Evaluación de Bactericidas para el Control de Huanglongbing (*Candidatus Liberibacter asiaticus*) en Lima Persa. In 2° Simposio Nacional sobre Investigación para el Manejo del Psilido Asiático de los cítricos y el Huanglongbing en México (pp. 399–404). inifap.
- Gotelli, N., & Colwell, R. (2011). Estimating species richness. In *Biological Diversity. Frontiers in Measurement and Assessment* (pp. 39–54).
- Gundale, M. J., Kardol, P., Nilsson, M. C., Nilsson, U., Lucas, R. W., & Wardle, D. a. (2014). Interactions with soil biota shift from negative to positive when a tree species is moved outside its native range. *New Phytologist*, 202(2), 415–421.
- Gutiérrez-Galeano, D., Ruiz-Medrano, R., Xoconostle-Cázares, B., (2015). Estado Actual de los cultivos genéticamente modificados en México y su contexto internacional. Centro de Investigación y Estudios Avanzados del IPN. México
- Guttman, D. S., McHardy, A. C., & Schulze-Lefert, P. (2014). Microbial genome-enabled insights into plant–microorganism interactions. *Nature Reviews Genetics*, 15(12), 797–813.
- Haney, C. H., Samuel, B. S., Bush, J., & Ausubel, F. M. (2015). Associations with rhizosphere bacteria can confer an adaptive advantage to plants. *Nature Plants*, 1(June).
- Huang, J., Nandi, S., Wu, L., Yalda, D., Bartley, G., Rodriguez, R., ... Huang, N. (2002). Expression of natural antimicrobial human lysozyme in rice grains. *Molecular Breeding*, 10, 83–94.
- Huerga, M., & San-Juan, S. (2005). EL CONTROL DE LAS PLAGAS EN LA AGRICULTURA ARGENTINA Miguel Huerga y AGRICULTURA ARGENTINA. Estudio Sectorial Agrícola Rural. Buenos Aires.
- Hughes, J. B., Hellmann, J. J., Ricketts, T. H., & Bohannon, B. J. M. (2001). Counting the Uncountable : Statistical Approaches to Estimating Microbial Diversity MINIREVIEW Counting the Uncountable :

- Statistical Approaches to Estimating Microbial Diversity. *Applied and Environmental Microbiology*, 67(10), 4399–4406.
- Hughes, J., & Bohannan, B. (2004). Application of ecological diversity statistics in microbial ecology. In *Molecular Microbial Ecology Manual* (2nd ed., pp. 1321–1344). Holanda: Kluwer Academic Publishers.
- Icoz, I., & Stotzky, G. (2008). Cry3Bb1 protein from *Bacillus thuringiensis* in root exudates and biomass of transgenic corn does not persist in soil. *Transgenic Research*, 17(4), 609–620.
- Janssen, P. H. (2006). Identifying the Dominant Soil Bacterial Taxa in Libraries of 16S rRNA and 16S rRNA Genes MINIREVIEWS Identifying the Dominant Soil Bacterial Taxa in Libraries of 16S rRNA and 16S rRNA Genes. *Applied and Environmental Microbiology*, 72(3), 1719–1728.
- Kim, M., Lee, K.-H., Yoon, S.-W., Kim, B.-S., Chun, J., & Yi, H. (2013). Analytical tools and databases for metagenomics in the next-generation sequencing era. *Genomics & Informatics*, 11(3), 102–13.
- Krebs, C. J. (2014). Similarity Coefficients and Cluster Analysis. In Pearson (Ed.), *Ecological Methodology* (3rd ed.). Vancouver.
- Larkin, R., Hopkins, D., & Martin, F. (1996). Suppression of *Fusarium* wilt of watermelon by nonpathogenic *Fusarium oxysporum* and other microorganisms recovered from disease-suppressive soil. *Phytopathology*.
- Lee, S.-H., Kim, C.-G., & Kang, H. (2011). Temporal dynamics of bacterial and fungal communities in a genetically modified (GM) rice ecosystem. *Microbial Ecology*, 61(3), 646–659.
- Abbasian, F., Lockington, R., Mallavarapu, M., & Naidu, R. (2015). A pyrosequencing-based analysis of microbial diversity governed by ecological conditions in the Winogradsky column. *World Journal of Microbiology and Biotechnology*.
- Lemanceau, P., & Alabouvette, C. (1991). Biological control of fusarium diseases by fluorescent *Pseudomonas* and non-pathogenic *Fusarium*. *Crop Protection*, 10(4), 279–286.
- Makarenkov, V., Boc, A., Xie, J., Peres-neto, P., Lapointe, F., & Legendre, P. (2010). Weighted bootstrapping : a correction method for assessing the robustness of phylogenetic trees.
- Mongomery, D. (2004). *Diseño y Análisis de Experimentos* (2da ed.). México: Limusa Wiley.
- Nakajima, H., Muranaka, T., Ishige, F., Akutsu, K., & Oeda, K. (1997). Fungal and bacterial disease resistance in transgenic plants expressing human lysozyme. *Plant Cell Report*, 16(1994), 674–679.
- Nguyen, N.-P., Warnow, T., Pop, M., & White, B. (2002). Comparative integrated omics: identification of key functionalities in microbial community-wide metabolic networks. *Npj Biofilms and Microbiomes*, 57(6), 10–13.
- Nocker, A., Burr, M., & Camper, A. K. (2007). Genotypic microbial community profiling: A critical technical review. *Microbial Ecology*, 54(2), 276–289.
- NORMA Oficial Mexicana NOM-004-SEMARNAT-2002, Protección ambiental.- Lodos y biosólidos.- Especificaciones y límites máximos permisibles de contaminantes para su aprovechamiento y disposición final, (2003). Diario Oficial de la Federación. Retrieved from http://dof.gob.mx/nota_detalle.php?codigo=691939&fecha=15/08/2003
- Ozor, N. (2008). Challenges and impacts of agricultural biotechnology on developing societies. *African Journal of Biotechnology*, 7(4), 322–330.

- Ozor, N., & Igbokwe, E. M. (2007). Roles of agricultural biotechnology in ensuring adequate food security in developing societies. *African Journal of Biotechnology*, 6(14), 1597–1602. <http://doi.org/10.4314/ajb.v6i14.57704>
- Pérez, M. (2015, December 17). Afectados por plagas , 70 por ciento de cultivos de café en Chiapas y Oaxaca. *La Jornada*.
- Quail, M., Smith, M. E., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., Salzberg, S. (2012). A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics*, 13(1), 341.
- Rao, A. R. (2008). Analysis of molecular variance. *Bioinformatics and Statistical Genomics*, (1), 1–6.
- Salcedo, D., Mora, G., Ignacio, C., Cintora, C., Hinojosa, R., DePaolis, F., & Mora, S. (2011). Perspectives Assessment of the economic impact of Huanglongbing (HLB) disease on Mexico ' s citrus chain. *COMUNICA Agricultural Health*, (January-July), 40–47.
- Salipante, S. J., Kawashima, T., Rosenthal, C., Hoogestraat, D. R., Cummings, L. A., Sengupta, D. J., ... Hoffman, N. G. (2014). Performance comparison of Illumina and Ion Torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. *Applied and Environmental Microbiology*, 80(24), 7583–7591.
- Santivañez, T., Mora, G., Díaz, G., López, J. I., & Vernal, P. (2013). Marco Estratégico para la Gestión Regional del Huanglongbing en América Latina y el Caribe. Santiago de Chile: FAO.
- Schröder, J. M., & Harder, J. (1999). Human beta-defensin-2. *International Journal of Biochemistry and Cell Biology*, 31(6), 645–651.
- SENASICA, (2016). Situación fitosanitaria actual [Recurso electrónico]. <http://senasica.gob.mx/?id=4608>
- SINAVEF. Mapa Dinámico Fitosanitario [Recurso electrónico]. Consulta mapa HLB. Fecha: 02/08/2016. <http://www.sinavef.gob.mx/MDF/>
- Smit, E., Smit, E., Leeflang, P., Leeflang, P., Glandorf, B., Glandorf, B., ... Wernars, K. (1999). Analysis of Fungal Diversity in the Wheat Rhizosphere by Sequencing of Cloned PCR-Amplified Genes Encoding 18S rRNA and Temperature Gradient Gel Electrophoresis. *Applied and Environmental Microbiology*, 65(6), 2614–2621.
- Spain, A. M., Krumholz, L. R., & Elshahed, M. S. (2009). Abundance, composition, diversity and novelty of soil Proteobacteria. *The ISME Journal*, 3(8), 992–1000.
- Swathi Anuradha, T., Divya, K., Jami, S. K., & Kirti, P. B. (2008). Transgenic tobacco and peanut plants expressing a mustard defensin show resistance to fungal pathogens. *Plant Cell Reports*, 27(11), 1777–1786.
- Tanabe, A. S., Nagai, S., Hida, K., Yasuike, M., Fujiwara, A., Nakamura, Y., ... Katakura, S. (2015). Comparative study of the validity of three regions of the 18S-rRNA gene for massively parallel sequencing-based monitoring of the planktonic eukaryote community. *Molecular Ecology Resources*, 402–414.
- Van de Peer, Y., Chapelle, S., & De Wachter, R. (1996). A quantitative map of nucleotide substitution rates in bacterial rRNA. *Nucleic Acids Research*, 24(17), 3381–3391.
- Van de Peer, Y., Neefs, J.-M., De Rijk, P., & De Wachter, R. (1993). Reconstructing evolution from eukaryotic small-ribosomal-subunit RNA sequences: calibration of the molecular clock. *Journal of Molecular Evolution*, 37, 221–232.

- Whiteley, A. S., Jenkins, S., Waite, I., Kresoje, N., Payne, H., Mullan, B., ... O'Donnell, A. (2012). Microbial 16S rRNA Ion Tag and community metagenome sequencing using the Ion Torrent (PGM) Platform. *Journal of Microbiological Methods*, 91(1), 80–88.
- Zarraonandia, I., Owens, S. M., Weisenhorn, P., West, K., Hampton-marcell, J., Lax, S., Gilbert, A. (2015). The Soil Microbiome Influences Grapevine-Associated Microbiota, 6(2), 1–10.
- Zecua-Nájera, R. (2013). Obtención de variedades de limón mexicano expresando la lisozima humana para controlar la enfermedad de Huanglongbing (HLB). (Tesis de maestría). Departamento de Biotecnología y Bioingeniería, CINVESTAV. México.