



CENTRO DE INVESTIGACIÓN Y ESTUDIOS AVANZADOS DEL
INSTITUTO POLITÉCNICO NACIONAL
UNIDAD IRAPUATO

Título de la tesis

Ecología genómica y funcional del Microbioma de la raíz coraloide

Tesis que presenta

BIOL. JOSÉ ANTONIO CORONA GÓMEZ

Para Obtener el Grado de

Maestro en Ciencias

En la especialidad de

Biología de Plantas

Directora de Tesis

Dra. Angélica Cibrián Jaramillo

Irapuato, Guanajuato.

Agosto, 2016

El presente trabajo se desarrolló en el Laboratorio de Genómica Ecológica y Evolutiva perteneciente a la Unidad de Genómica Avanzada, del Centro de Investigaciones y Estudios Avanzados del IPN, Unidad Irapuato, bajo la dirección de la Dra. Angélica Cibrián Jaramillo. Financiado con el proyecto CONACYT de ciencia básica No. 169701 titulado: *“Population Genomics of Cycads and their Cyanobacteria: Unraveling the Evolutionary Basis of Adaptation to Novel and Changing Enviroments”* y la beca para estudios de posgrado CONACYT No. 338379.

Irapuato, Guanajuato. Agosto, 2016

AGRADECIMIENTOS

Al CONACYT por el apoyo económico otorgado durante la realización de este estudio (CVU/Bacario: 629615).

A la Dra. Angelica Cibrián Jaramillo, por el apoyo, las ideas y los consejos que me ha dado; por el conocimiento brindado y la confianza que ha puesto en mi durante mi formación académica tanto en Licenciatura como en Maestría.

Al comité académico, conformado por los doctores Francisco Barona Gómez y Víctor Olalde Portugal, porque a lo largo de este proyecto aportaron valiosas observaciones y consejos.

A Flor Zamudio, por su invaluable apoyo en laboratorio y consejos técnicos.

A mis compañeros del Laboratorio 11 y 14 por su compañerismo y camaradería, en especial a Pablo Cruz, Pablo Suarez y Daniel Díaz por colaborar en el proyecto de cícadas.

Al Dr. Luis Delaye y al Dr. Cei Abreu por sus valiosas enseñanzas en Bioinformática.

A mi familia por su apoyo incondicional e inculcarme el deseo de seguir adelante, de superación y perseverancia.

A mis amigos dentro y fuera de Cinvestav por siempre poner una sonrisa en mi rostro.

Al estudiante Rafa Rincón por su apoyo en la parte Bioinformática.

Un agradecimiento especial a mi pareja Mónica Rodríguez Uribe, por ser una amiga, confidente, consejera, correctora de ortografía; por el apoyo en los momentos difíciles y por ser la persona que más tiempo pasa conmigo.

Índices

Introducción.....	3
Antecedentes	7
El microbioma de las plantas.....	7
Historia natural de las Cícadas.....	8
Raíces coraloides de cícadas y su simbiosis con bacterias endófitas	10
Comunicación planta-cianobacteria.....	12
Establecimiento de la simbiosis bacteria-cicada.	13
Taxonomía de las bacterias de la raíz coraloide.	14
Hipótesis.....	17
Objetivo general	17
Objetivos particulares	17
Metodología.....	18
Colecta y procesado de material biológico	18
Colecta de muestras.....	18
Lavado de raíces coraloides.....	19
Aislado de bacterias	20
Crecimiento del Co-Cultivo	21
Inoculación	21
Medios	21
Condiciones y tiempo de crecimiento	21
Extracción de ADN	22
Extracción de ADN de cultivos de cianobacterias	22

Extracción de ADN Co-Cultivos	22
Secuenciación	23
Secuenciación 16S rRNA.....	23
Secuenciación metagenomas.....	24
Edición de secuencias	24
Edición y limpieza de datos Metagenómicos.....	25
Análisis Taxonómico.....	25
Filogenia 16S rRNA	27
Árbol de árboles	27
Edición de árboles filogenéticos.....	28
Análisis taxonómico de Metagenomas	28
Análisis ecológicos y de composición de comunidad	29
Análisis Funcional de Metagenomas	30
Resultados tesis.....	32
Colecciones Biológicas de raíces coraloides.....	32
Colectas bacterianas y co-cultivos	33
Cepario.....	34
Taxa identificados.....	34
Identificación taxonómica mediante el marcador 16S rRNA.....	34
Índices de diversidad entre los diferentes ambientes	45
Composición funcional en diferentes ambientes.....	48
Co-cultivos fotosintéticos a un año de ser sembrados	52
Comparación taxonómica de los co-cultivos con diferente tiempo de incubación	52
Comparación funcional de los co-cultivos con diferente tiempo de incubación..	53

Discusión.....	57
Revelando el microbioma de la raíz coraloide: más allá de las cianobacterias..	57
Diversidad entre medios de cultivo.....	58
Distribución de los géneros bacterianos de acuerdo al medio ambiente	59
Diversidad cultivos de 1 año.....	60
Funciones del microbioma de la raíz coraloide: fijación de nitrógeno	60
Microbioma de la raíz coraloide: una comunidad bacteriana asociada a las cianobacterias.....	63
Conclusion.....	66
Perspectivas.....	67
Bibliografía	68
Anexos	79
Programas para identificación taxonómica.....	84
Programas para edición de árboles filogenéticos.....	91
Arboles Filogenéticos de árbol de arboles	93
Programa para análisis de diversidad	107

RESUMEN

Las cícadas son repositorios de una larga historia evolutiva, que constituyen un enlace entre las plantas con semilla y la diversificación de las angiospermas, y los orígenes de las asociaciones del microbioma con plantas vasculares con semilla. En este estudio nos centramos en describir la diversidad taxonómica y funcional del microbioma raíz coraloide de *Dioon merolae*, en poblaciones naturales; se mostró que la raíz coraloide alberga una amplia diversidad de taxa bacterianos, incluyendo géneros nunca antes reportados en esta simbiosis. Se encontraron diferencias en la composición del microbioma entre los ambientes donde se colectó *D. merolae* (seco y húmedo), también entre métodos y en particular en cultivos con un tiempo de incubación prolongado. En contraste a nivel funcional se encontraron muchas similitudes entre ambientes, donde se encontró la presencia de las rutas metabólicas básicas relacionadas entre sí. Se encontró que hay roles específicos de las cianobacterias en los co-cultivos como la fijación de nitrógeno. Lo que nos lleva a suponer está es una simbiosis abierta a diferentes grupos de bacterias que crecen en consorcio con cianobacterias.

ABSTRACT

Cycads are repositories of a long evolutionary history, which constitute a link between seed plants and the diversification of angiosperms, and with them, the origins of the microbiome associations with vascular seed plants. In this study we focused on describing the taxonomic and functional diversity of the microbiome of coralloid root *Dioon merolae* in natural populations. We show that the root coralloide hosts a wide variety of bacterial taxa, including novel genres reported for the first time in this symbiosis. Differences in the composition of the microbiome between environments (dry and humid); and between culture methods, in particular in cultures with a prolonged incubation, were found. In contrast, there were functional similarities between environments, with the presence of basic shared metabolic pathways. We found there are specific functional roles of cyanobacteria in co-cultures, such as nitrogen fixation. This leads us to suppose it is a symbiosis open to different groups of bacteria, but always growing in consortium with cyanobacteria.

INTRODUCCIÓN

Los procesos ecológicos que llevan a las especies a adaptarse y a sobrevivir a los nuevos ambientes cambian a lo largo de la historia y son una de las grandes bases de la evolución (Losos and Ricklefs 2009). Prueba de esto es la gran diversidad de especies que tenemos en el planeta, cada una de ellas con adaptaciones únicas que les permiten sobrevivir y mantener su descendencia (Tilman 1999). En general las especies cuentan con un metabolismo central el cual mantiene todas sus funciones vitales y de este se desprende un segundo 'metabolismo adaptativo' (Haslam 1994), el cual no solo les permite sobrevivir, sino que también les da la habilidad para responder con adaptaciones a los cambios que el ambiente les impone en diferentes escalas temporal y espaciales (Schilling, Letscher and Palsson 2000). El desarrollo de nuevas rutas metabólicas o la modificación del 'metabolismo adaptativo' suelen ser eventos que no suceden con una simple mutación si no que requieren de muchas de éstas y un proceso de selección natural que lleva tiempo (Rieseberg and Willis 2007).

Agregando a la complejidad de estos procesos, los organismos no existen solos en la naturaleza, están obligados a interactuar con quienes comparten el mismo ambiente, lo que genera una gama de interacciones biológicas y ecológicas entre ellos y su hábitat. Desde el punto de vista evolutivo, la simbiosis se considera un proceso en el que las especies involucradas estrechan cada vez más su relación a lo largo del tiempo, donde una especie aprovecha las habilidades que no poseen de la otra especie y viceversa, y en la cual la selección natural favorece estas ventajas y los individuos involucrados en la relación proliferan (Futuyma 1983; Xu *et al.* 2014). La simbiosis es una interacción entre dos organismos que conviven y uno vive dentro del otro y/o en asociación; a su vez una asociación involucra la interacción de dos especies en las cuales una depende de la otra ya sea que se brinden un hábitat o ayuden en una parte del ciclo de vida (Margulis and Sagan 2008). La simbiosis representa una unidad integrada por dos entidades biológicas originalmente independientes, que gracias

a una serie de eventos están juntas y coevolucionan creando un mecanismo que origina nuevas especies y nichos ecológicos (Mereschkowsky 1926; Margulis and Sagan 2008). Dependiendo de la naturaleza de la interacción biológica entre especies, pueden tomar distintos caminos evolutivos (Poulin 1996). Por ejemplo, cuando dos especies compiten en una situación de depredación, por lo general se da una carrera armamentista entre las especies y la que desarrolla las mejores estrategias de defensa o caza es la que prospera a costa de la otra (Vermeij 1994). Pero lo que sucede cuando la relación es benéfica como en una asociación simbiótica, las dos especies llevan a cabo un proceso de co-adaptación que requiere de cambios fisiológicos y metabólicos (Rai 1990).

Actualmente utilizamos el término simbiosis para describir la interacción de dos especies donde hay beneficios para uno o los dos organismos, cuando ambas especies se benefician mutuamente en una simbiosis se le considera a esta del tipo mutualista, la existencia de beneficios mutuos entre dos especies no es fácil de definir en todos los casos, por lo que en muchas interacciones complejas simplemente se utiliza el termino general de simbiosis para describirlo (Kneip *et al.* 2007a). Las asociaciones simbióticas presentan distintos grados de integración entre las especies, desde una simple asociación por comportamiento, hasta la *simbiogénesis* donde dos especies terminan fusionando sus genomas dando origen a un organismo único, como el caso de las mitocondrias y los cloroplastos en eucariotas (Khakhina, et al., 1992). Las asociaciones simbióticas son uno de los casos más exitosos en el mundo natural, ya que es más fácil que una especie se relacione con otro aprovechando las adaptaciones que tienen en conjunto, a que una especie evolucione por si sola y adquiera las habilidades de una segunda especie (Margulis, 1991). En estos casos es probable que existan estructuras funcionales altamente especializadas para la interacción simbiótica (Rai *et al.* 2002).

Las asociaciones simbióticas en plantas tienen particularidades ya descritas. Todas las plantas en su interior se encuentran habitadas por diversas comunidades de microorganismos compuestas por bacterias, arqueas, protistas y hongos (Lindow and Brandl 2003). Estos microorganismos con un estilo de vida

endófito juegan un papel importante en el desarrollo, crecimiento, adaptación y diversificación de las plantas (Hardoim *et al.* 2015). De manera general, la composición de los endófitos de las plantas depende del genotipo de la planta y de las condiciones ecológicas donde crece el organismo (Boon *et al.* 2014). La relación más común entre los endófitos y su hospedero es el comensalismo, con funciones aún desconocidas, menos común son las relaciones mutualistas en plantas (Hallmann *et al.* 1997).

En este trabajo me concentro en la simbiosis entre cícadas (Zamiaceae) y bacterias endófitas de la raíz. Las cícadas son gimnospermas (Cycadales) distribuidas en una gran diversidad de ambientes distintos, entre ellos algunos de los más hostiles para la vida vegetal como los desiertos e islas volcánicas (Osborne *et al.* 2012b). Poseen raíces especializadas llamadas coraloides en las que alojan bacterias, fijadoras de nitrógeno para la planta, y ésta a su vez esta brinda hospedaje y una fuente de carbono a las cianobacterias (Elmerich and Newton 2007). Se presume que es más eficiente que la planta se adapte a una simbiosis bacteriana, a que desarrollen todo el mecanismo de la fijación de nitrógeno (Lindblad 1990), mismo que ha sido optimizado por cianobacterias previo a la existencia de las plantas (Berman-Frank, Lundgren and Falkowski 2003a). Esta es una idea interesante desde un contexto evolutivo, ya que es probable que la huella de esta relación y adaptación evolutiva, haya quedado registrada tanto en el genoma de la planta como en de la cianobacteria, pudiendo ser observable desde un contexto genético que ayude a entender este tipo de interacciones biológicas. Asimismo, es probable que la simbiosis entre cícadas y bacterias, haya permitido que las cícadas siendo el grupo de plantas con semilla más antiguo sigan adaptándose continuamente a nuevos ambientes e incluso diversificándose hasta hace relativamente poco (Nagalingum *et al.* 2011).

Pese a una serie de antecedentes taxonómicos y fisiológicos (descritos a detalle en la siguiente sección), no sabemos las bases biológicas y evolutivas de la relación cícada-cianobionte, desconocemos la composición de la comunidad de endófitos desde un punto de vista taxonómico, genómico o metabólico, y sabemos aún menos de cómo cambia esa comunidad en un contexto ecológico

(i.e. diversos ambientes). En este sistema además, surge la posibilidad de que algunas de las habilidades metabólicas de las cianobacterias y el microbioma asociado a la raíz coraloide hayan pasado a las cícadas en un evento de intercambio lateral (horizontal gene transfer, HGT), o si se haya dado un evento de disminución de los genomas donde los alelos redundantes entre las dos especies se eliminan estrechando aún más la relación al hacerla más dependiente (Roughgarden 1983). Con el objetivo de entender las bases genómico-ecológicas de la interacción entre cícadas, cianobacterias y sus endófitos bacterianos, o microbioma, seleccionamos a *Dioon merolae*, una de las 54 especies de cícadas reportadas en México. *Dioon merolae* tiene poblaciones con distinta distribución geográfica y crece en ambientes distintos. Un sistema biológico de esa naturaleza permite evaluar la naturaleza del microbioma de raíz en el contexto de la simbiosis y en un marco evolutivo, el impacto de las condiciones ecológicas que influyen en la asociación entre simbioses.

ANTECEDENTES

El microbioma de las plantas

Los organismos unicelulares aparecieron hace unos 3500 millones de años y se encuentran en todas partes, pero solo en raras ocasiones están solos (Staley and Konopka 1985), Es claro que muchos microorganismos dependen de otros organismos para crecer y reproducirse (Hug *et al.* 2012). La definición de microbioma más simple es: el grupo de microorganismos que habitan un lugar en un tiempo específico y con factores abióticos definidos. También se puede interpretar como todos los organismos que comparten un cuerpo (Lederberg and McCray 2001). Siendo más específicos, se considera que el microbioma es el conjunto de microorganismos y sus genomas en un ambiente en particular, sin involucrar interacciones ecológicas (Boon *et al.* 2014).

El concepto del microbioma asociado a plantas se define como la colección de organismos y sus genomas que viven asociados al hospedero. Esto ha dado lugar a nuevas ideas sobre la evolución de las plantas donde las fuerzas de selección no actúan solo sobre el genoma de la planta, sino que también sobre la microcomunidad asociada a estas (Hallmann *et al.* 2011). Rasgos importantes de adaptación pueden estar asociados al microbioma. El uso de las tecnologías de secuenciación masiva como el análisis de metagenomas, permiten el estudio de los microbiomas en su totalidad y ayudan a tener profundidad a nivel funcional. Las ciencias *-ómicas* son una herramienta invaluable para explorar, identificar, y caracterizar la contribución genética y de elementos metabólicos involucrados en la interacción entre la planta hospedera y sus endófitos. La secuenciación de metagenomas de microbiomas vegetales han revelado funciones importantes relacionados con la supervivencia de los endófitos bacterianos dentro de la planta (Sessitsch *et al.* 2012) y el estudio de metabolomas a demostrados efectos beneficios para la planta por metabolitos producidos por bacterias (Scherling *et al.* 2009).

Establecer un nombre a los de interacciones entre los miembros del microbioma y los hospederos han cambiado y se han redefinido con el tiempo. Los términos "simbiosis asociativa" y "fijación de nitrógeno asociativa" se han utilizado para describir la interacción entre *Azospirillum* y otras bacterias de la rizosfera como las cianobacterias y sus plantas hospederas. Ninguno de ellos es totalmente satisfactorio como un término genérico. El proceso y el grado de interacción entre la bacteria y la planta pueden diferir entre las especies, ya que hay diferentes estructuras diferenciadas en las raíces inducidas por las bacterias, y el alcance de la colonización no siempre está bien definido, y el beneficio de la asociación a menudo ha sido cuestionada (Hardoim *et al.* 2015). Por ejemplo, las bacterias fijadoras de nitrógeno del suelo que se pueden encontrar en estrecha asociación con las raíces, pueden ser denominadas "bacterias fijadoras de nitrógeno asociativas" (Elmerich 2007). Para ser más específicos se acuñó el término "fijadores de nitrógeno endófitos" usado para describir un tipo de asociación específica entre plantas y bacterias fijadoras de nitrógeno, donde las bacterias viven dentro de la planta pero no crean síntomas patogénicos (Costa and Lindblad 2002a). Es evidente que las cianobacterias hasta ahora descritas en simbiosis con plantas, sean todas; "fijadores de nitrógeno endófitos" (Rai *et al.* 2002).

Historia natural de las Cícadas

Este trabajo se centró en la descripción del microbioma de la raíz coraloide de las cícadas. Las cícadas (Cycadales) son de las gimnospermas más basales después del *Ginkgo biloba* que aún persisten en la actualidad, con 300 millones de años de haber aparecido en el planeta (Hilerio *et al.* 2008). Hasta hace 100 millones de años atrás las cícadas tenían una distribución cosmopolita, dominando todos los continentes, hoy en día las pocas especies que han sobrevivido se distribuyen por todos los continentes excepto Europa y la Antártida, generalmente se las encuentra en pequeñas poblaciones en los trópicos y subtropicos de ambos hemisferios (Costa and Lindblad 2002a). Sin embargo, pese a la distribución reducida, estudios recientes apuntan a que las cícadas actuales no son mucho mayores de 12 millones de años (Nagalingum *et al.* 2011). Se caracterizan por ser

relativamente no ramificadas, con tallos gruesos y blandos de madera escasa, grandes hojas compuestas, con conos terminales o subterminales. Plantas dioicas, con apariencia de palmas. De las cuales se pueden encontrar en el registro fósil hasta el triásico (Vovides and Iglesias 1994). Poseen características casi exclusivas dentro de las gimnospermas, como la producción de anterozoides (gametos masculinos), característica que comparten únicamente con el género *Ginkgo* (Norstog and Nicholls 1997). Se estima que poseen genomas muy grandes de entre 23 y 63 Gpb (Dolezel *et al.* 2003; Burleigh *et al.* 2012). Adicionalmente, llevan a cabo una gran cantidad de interacciones con múltiples especies (Norstog and Nicholls 1997; Schneider *et al.* 2002), particularmente, son interesantes las interacciones que realizan con hongos para formar micorrizas y con cianobacterias para la fijación de nitrógeno en raíces coraloides especializadas (Fisher and Vovides 2004; Bergman, Rai and Rasmussen 2007; Adams and Duggan 2012), descritas en detalle en las siguientes secciones.

Se han registrado alrededor de 331 especies de cícadas en el mundo, clasificadas dentro de 3 familias (Cycadaceae, Stangeriaceae y Zamiaceae) y 10 géneros: *Bowenia* (2), *Ceratozamia* (27), *Cycas* (107), *Dioon* (14), *Encephalartos* (65), *Lepidozamia* (2), *Macrozamia* (41), *Microcycas* (1), *Stangeria* (1) y *Zamia* (71) (Osborne *et al.* 2012a). México forma el escenario perfecto para responder preguntas básicas sobre ecología evolutiva en cícadas y su microbioma, ya que existen aproximadamente 160 especies de cícadas en el mundo distribuidas en tres familias: Cycadaceae, Stangeriaceae y *Bowenia*, de las que 54 especies se encuentran reportadas para el país (Lindblad, 2009), clasificadas en tres géneros; *Ceratozamia*, *Zamia* y *Dioon*. De éstas, un 75% de las especies son endémicas de México, lo que hace al país un importante centro de diversidad de cícadas neotropicales (Osborne, *et al.*, 1999; Vovides and Iglesias, 1994).

Dentro de los géneros de cícadas presentes en México resalta particularmente el género *Dioon*, ya que 13 de las 14 especies que alberga son endémicas a algunas de las provincias del territorio Mexicano, solamente 1 se encuentra en Honduras (Osborne *et al.* 2012a). Este género de plantas ha sido comúnmente dividido en 2 grupos mediante estudios morfológicos y filogenéticos

(Norstog and Nicholls 1997; Bogler and Francisco-Ortega 2004; González, Vovides and Bárcenas 2008). De acuerdo a sus características morfológicas, un grupo se caracteriza por poseer especies con frondas y conos generalmente más grandes y troncos bien desarrollados (eje. *D. spinulosum*, *D. mejiae* y *D. rzedowskii*). El segundo grupo posee especies menos robustas, con troncos y conos generalmente más pequeños y frondas considerablemente más cortas (eje. *D. caputoi*, *D. holmgrenii*, *D. purpusii*, *D. califanoi*, *D. tomasellii*) (Norstog and Nicholls 1997). Los análisis filogenéticos indican que existen dos clados importantes. El primero está compuesto por: *D. angustifolium*, *D. mejiae*, *D. spinulosum*, *D. rzedowskii* y *D. edule*. El segundo está formado por: *D. argenteum*, *D. purpusi*, *D. merolae*, *D. caputoi*, *D. holmgrenii*, *D. califanoi*, *D. tomaselli* y *D. sonorensis* (González, Vovides and Bárcenas 2008).

En este trabajo nos enfocamos en *Dioon merolae* (Figura 1). Esta especie tiene como características frondas aplanadas con foliolos lineales-lanceolados que surgen del raquis en un ángulo agudo, imbricados y se insertan oblicuamente en él, excepto en la porción basal de la fronda (De Luca, Sabato and Torres 1981). De acuerdo con “The IUCN Red List of Threatened Species 2013” el estatus en el que se *D. merolae* es “vulnerable” (VU). Este hecho puede ser debido al aumento del pastoreo y de las tierras agrícolas, así como la extracción ilegal de plántulas y semillas que en general han afectado enormemente a las poblaciones de cícadas (Norstog and Nicholls 1997; Vovides *et al.* 2003). *D. merolae* se distribuye tanto en la provincia de la Sierra Madre del Sur, como parcialmente en la provincia de la costa del Pacífico Mexicano (Morrone 2006; Cabrera-Toledo *et al.* 2010). *D. merolae* presenta una distribución amplia y densidades de aproximadamente 0.01 ind/m², cuenta con poblaciones separadas hasta 440 km de distancia y algunas poblaciones tienen más 1000 individuos (Cabrera-Toledo *et al.* 2010; Lázaro Zermeño *et al.* 2011).

Raíces coraloides de cícadas y su simbiosis con bacterias endófitas

La simbiosis bacteria-cícadas tiene dos características únicas: Primero, son la única simbiosis conocida entre gimnospermas y procariotas diazótrofos. En

segundo lugar, son la única simbiosis en la que las bacterias invaden los tejidos de las raíces especializadas de una planta (Bergman, Rai and Rasmussen 2007; Cuddy *et al.* 2012). En el registro fósil no se ha encontrado la zona simbiótica con las cianobacterias, pero todas las cícadas forman esta simbiosis, lo que sugiere que surgió como un solo evento evolutivo hace 250 Ma. (Döbereiner 1992; Costa and Lindblad 2002a). Asimismo, se presume que es una simbiosis mutualista donde ambas especies se benefician, esta simbiosis emplea un modo biotrófico de nutrición en el cual el intercambio de nutrientes se hace entre células vivas.

Se pueden identificar dentro de las cícadas tres tipos de raíces distintas: la raíz primaria que es similar a la raíz principal de muchos tipos de plantas, raíces laterales y un tercer tipo de raíces altamente especializadas, engrosadas y con geotropismo negativo, con forma similar a la de un coral, las cuales son colonizadas por bacterias, y se les conoce como raíces coraloides (Lindblad *et al.* 1985; Costa and Lindblad 2002a). Las raíces coraloides han sido encontradas en todos los géneros, y en cada una de las especies reportadas (Lindblad 1990). La forma estructural de la raíz coraloides se forma hasta que se da la simbiosis con las bacterias. Previo a la colonización de las cianobacterias se forman las primeras papillas o callos, de los cuales se generarán las raíces coraloides y después de la colonización las raíces toman su forma típica de coral (Ahern and Staff 1994a). La formación de las raíces coraloides inicia en zonas adventicias y es a partir del hipocotilo inmediatamente abajo de los peciolos cotiledonarios del cual emergen. Seguido a este desarrollo tenemos la fase de maduración, colonización de las cianobacterias, la formación de las ramificaciones coraloides, senescencia y la regeneración (Ahern and Staff 1994a). Las raíces coraloides no presentan la misma morfología en todas las cícadas, incluso hay diferencias entre plantas de la misma especie, ya que entran en juego factores como la edad, presencia ausencia de cianobacterias, humedad, y el modo de origen, en las formas de la raíz.

La mayor parte de la evidencia previa a este trabajo describe una asociación entre cícadas y cianobacterias exclusivamente (ver Tabla 1). Las cianobacterias están confinadas a un lugar específico de las raíces coraloides de

las cícadas, en un corte transversal de una raíz coraloide infectada se ve a simple vista un halo verde azulado, formado por las cianobacterias (Lindblad *et al.* 1985; Costa and Lindblad 2002a). Las cianobacterias tienen una gran amplitud ecológica, y tienen relaciones de simbiosis con un gran rango de organismos, incluyendo musgos, helechos, gimnospermas, angiospermas y hongos (Costa and Lindblad 2002b; Rai *et al.* 2002; Schenk, Weingart and Ullrich 2008; Adams and Duggan 2012; Cuddy, Neilan and Gehring 2012). En todas las situaciones en que las cianobacterias están asociadas a tejidos de plantas se encuentran ligadas a la fijación de nitrógeno (Elmerich and Newton 2007). Las distintas especies de cianobacterias han desarrollado estrategias adaptativas sorprendentes, desde su capacidad de ser autótrofos o heterótrofos, dependiendo de la fuente de carbono. Dada la evidencia previa en la literatura, el papel fundamental de las cianobacterias en las raíces coraloideas de cícadas, parece ser la de fijación de nitrógeno.

Comunicación planta-cianobacteria

Similar a la relación simbiótica entre *Rhizobium* y leguminosas (van Rhijn and Vanderleyden 1995), las señales químicas de la planta son esenciales en la formación de simbiosis con cianobacterias. Hay evidencia experimental de tres señales químicas en el establecimiento de la simbiosis de las cianobacterias y las plantas (Campbell, Wong and Meeks 2003), cuando la planta se encuentra en estrés de nitrógeno excreta *HIF*, con lo cual aumenta la producción de hormogonios móviles en las poblaciones de cianobacterias que se encuentren en el ambiente (Campbell and Meeks 1989). El segundo paso es para incrementar la probabilidad de infección, la planta produce sustancias químicas atrayentes para dirigir el movimiento de los hormogonios hacia el punto de infección y se lleve a cabo la colonización (Knight and Adams 1996). El tercer paso después de la colonización es la producción de un factor de represión del hormogonio (*HRF*, *Hormogonium repressing factor*) dentro de la planta, *HRF* es dominante sobre *HIF* y se cree que desplaza la formación de hormogonios, permitiendo el crecimiento de los filamentos y de este modo se diferencien las células vegetales en

heterocistos y expresen la nitrogenasa (Cohen and Meeks 1997).

Hasta el momento no se conoce la identidad química de los factores *HRF* y *HIF*, también se desconocen las señales exactas que manda la planta para el alto número de células vegetativas se diferencien de heterocistos y cuáles son sus blancos dentro de las cianobacterias. Se sabe muy poco de los genes que se ven activados durante el proceso de la simbiosis, como se traducen estas señales dentro de las cianobacterias, como es el efecto en la regulación transcripcional y cómo es que evolucionaron estas adaptaciones tan especializadas (Campbell, Wong and Meeks 2003), también es difícil entender como es el intercambio de nutrientes entre la planta y la bacteria.

Establecimiento de la simbiosis bacteria-cicada.

El mecanismo de infección de las cianobacterias sigue siendo incierto, se cree por una parte que la infección se podría dar por las lenticelas de la raíz (Costa and Lindblad 2002a), mientras otros apoyan la idea que la colonización se da por las grietas de las capas exteriores de la raíz (Nathanielsz and Staff 1975). El mecanismo de infección en raíces de cícadas sigue siendo escasa, tanto como el escenario en el que se da la infección y el modo preciso de entrada resulta complicado de evaluar (Ahern and Staff 1994a), una de las razones es la relativamente lenta germinación de las grandes semillas de las cícadas, lo que dificulta los experimentos de infección a gran escala bajo condiciones de laboratorio. Es evidente que estos aspectos deben ser estudiados con mayor detalle, el desarrollo de las raíces coraloides esta genéticamente determinado por la planta huésped, estas se forman en ausencia de cianobacterias, pero si hay presencia de cianobacterias competentes para la simbiosis y se lleva a cabo la infección, se llevan a cabo modificaciones de las raíces en la zona simbiótica (Ahern and Staff 1994a).

Las modificaciones de las raíces coraloides producidas por la infección de las cianobacterias, sugiere que la cianobacteria es capaz de interferir con la maquinaria de desarrollo de la planta, para iniciar los comportamientos necesarios

para la simbiosis. Compuestos fenólicos están presentes en el material mucilaginoso que llena la zona donde se encuentran las cianobacterias, que también puede ser inducido por la simbiosis. Estos compuestos fenólicos pueden funcionar como un método para excluir a otros organismos (Obukowicz, Schaller and Kennedy 1981) incluso podría excluir a otras cianobacterias.

Los cianobiontes exhiben todos los mismos patrones de ciclo de vida y la diferenciación celular como los ilustrados en la Figura 2, en la cual la fase móvil (hormogonio) se utiliza para llevar a cabo la infección de las raíces coraloideas. Pero una modificación más evidente en los cianobiontes es el aumento en la frecuencia de heterocistos, este aumento de la fijación de nitrógeno podría darse por distintos motivos: el aumento de carbono proporcionado por la planta huésped o los factores del ambiente interno de la raíz coraloide como la ausencia de oxígeno, nitrógeno, ausencia de luz, etc. o algún compuesto producido por la planta, esto es algo que aún no se define. Sin embargo, parece que hay una correlación entre el grado de heterotrofia y la frecuencia de heterocistos en simbiosis (Döbereiner 1992).

Taxonomía de las bacterias de la raíz coraloide.

En cuanto a la taxonomía de las cianobacterias asociadas a la raíz coraloide de cícadas, Reinke se dio cuenta de la asociación de cianobacterias con un tipo específico de raíces de cícadas ya en 1872 y se caracterizó el cianobionte como una "*Anabaena verdadera*". Las cianobacterias que forman asociaciones con raíces coraloideas de cícadas han sido desde entonces taxonómicamente atribuido indistintamente a *Anabaena*, *Nostoc punctiforme*, *N. cycadeae*, o simplemente *Nostoc sp.* (Livingstone and Jaworski 1980; Costa and Lindblad 2002a; Rasmussen and Nilsson 2002) pero, en la mayoría de los casos se han identificado los cianobiontes como *N. punctiforme* (Tabla 1). De hecho, la cianobacteria aislada de *Macrozamia riedli* fue seleccionado como el tipo de cepa del género *Nostoc* bajo la designación *Nostoc PCC 73102* (Rippka *et al.* 1979). Por otra parte, el primer genoma de una cianobacteria simbióticamente

competente que fue secuenciado era la de *Nostoc ATCC 29133*, una cepa equivalente a *Nostoc PCC 73102* (Meeks *et al.* 2002). Esto, sin duda condujo a una mejor comprensión de la definición del concepto "competencia simbiótica " de cianobacterias. Por qué si bien la simbiosis no es específica de una sola especie de cianobacteria, se ve que en la mayoría de los estudios predomina *Nostoc* como el género más común (Tabla 1).

En análisis previos donde se ha utilizado caracterización por técnicas genéticas se ha demostrado que numerosas cepas de cianobacterias genéticamente distintas están en simbiosis con las cícadadas (Costa, Paulsrud and Lindblad 1999; Costa and Lindblad 2002a; Zheng *et al.* 2002), en algunos estudios se ha encontrado a *Calothrix* y *Tolypothrix* como simbionte (Grobbelaar *et al.* 1987; Gehringer *et al.* 2010). Esta variación puede deberse al hecho de que las cícadadas son infectadas de novó después de la germinación de las semillas, con lo que la flora de cianobacterias existente en el ambiente relacionada con el entorno y las condiciones puede infectar la raíz (Bergman, Rai and Rasmussen 2007).

Con estos antecedentes, surge la siguiente pregunta: **¿Cuál es el valor evolutivo del microbioma de la raíz coraloide en cícadadas? ¿Es posible que las cícadadas se adapten a los diferentes ambientes y al paso del tiempo, mediante un núcleo o un grupo de diversas bacterias simbióticas?**, quizá adquiriendo algunos de sus metabolitos y cualidades adaptativas en el proceso de la simbiosis. Esta pregunta es muy general y se necesita analizar y caracterizar a profundidad el microbioma y la planta como hospedero. En este trabajo me concentro en la primera parte de la caracterización del microbioma de la planta como se describe en las siguientes secciones.

Tabla 1 Especies de bacterias identificadas en simbiosis con cícadas en la literatura.

Especie de cícada	Bacterias encontradas	Referencia
<i>Zamia integrifolia</i> , <i>Dioon</i> sp.	<i>Oscillatoria</i> , <i>Calothrix</i> , <i>Nostoc</i> , <i>Anabaena</i>	(Zimmerman and Rosen 1992)
<i>Cycas circinalis</i> , <i>C. rumphii</i> , <i>Encephalartos lebomboensis</i> , <i>E. villosus</i> , <i>Zamia pumila</i>	<i>Nostoc</i>	(Costa, Paulsrud and Lindblad 1999)
<i>Cycas circinalis</i> , <i>Ceratozamia mexicana</i> , <i>Encephalartos villosus</i>	<i>Rizoplano</i> (<i>Bacillus</i> , <i>Streptomyces</i> , <i>Arthrobacter</i> , <i>Myxobacterales</i> , <i>Rhodococcus</i> , <i>Flavobacterium</i> , <i>Cytophaga</i>)	(Lobakova, Orazova and Dobrovol'skaya 2003)
<i>Macrozamia</i> , <i>Lepidozamia</i> , <i>Bowenia</i> , <i>Cycas</i> , <i>Encephalartos</i> y <i>Zamia</i>	<i>Nostoc</i>	(Costa, Romero and Lindblad 2004)
<i>Cycas circinalis</i> , <i>C. Rumphii</i> , <i>C.</i> <i>Revoluta</i>	<i>Calothrix</i> y <i>Nostoc</i>	(Thajuddin <i>et al.</i> 2010)
<i>Macrozamia</i> spp.	<i>Calothrix</i> y <i>Nostoc</i>	(Gehring <i>et al.</i> 2010)
<i>Macrozamia</i> , <i>Lepidozamia</i> , <i>Bowenia</i> and <i>Cycas</i>	<i>Leptolyngya</i> , <i>Calothrix</i> y <i>Nostoc</i>	(Cuddy, Neilan and Gehring 2012)
<i>Cycas revoluta</i>	<i>Nostoc</i>	(Yamada <i>et al.</i> 2012)
<i>Dioon caputoi</i> , <i>Dioon merolae</i>	<i>Nostoc</i> , <i>Tolypothrix</i> , <i>Cianobacteria del grupo 5</i>	Corona-Gómez <i>et al.</i> , unpublished

HIPÓTESIS

Existe un núcleo funcional y taxonómico bacteriano en el interior de la raíz coraloide que se mantendrá constante, aún en distintos ambientes.

OBJETIVO GENERAL

Caracterizar la diversidad taxonómica y funcional de la microbiota endófito de la raíz coraloide de individuos de la especie *D. merolae* provenientes de dos ambientes contrastantes.

Objetivos particulares

- Procesar muestras de raíz coraloide de *D. merolae*: colectado, transporte al laboratorio, lavado, macerado y almacenamiento en glicerol.
- Sembrado de co-cultivos y cultivos bacterianos a partir de los macerados de la raíz coraloide.
- Secuenciación de amplicon 16s rRNA de aislados, y secuenciación masiva del ADN obtenido de los co-cultivos.
- Identificar taxonómicamente la comunidad bacteriana de la raíz coraloide, determinando los OTUs que componen la comunidad y representarlos filogenéticamente usando (16s rRNA, metagenomas).
 - Determinar la composición ecológica de la comunidad bacteriana de la raíz coraloide.
- Minar bioinformáticamente los genomas y metagenomas en búsqueda de rutas metabólicas relacionadas con funciones simbióticas.
 - Secuenciar los genomas de los principales géneros bacterianos
 - Minar las funciones relacionadas con el funcionamiento de la comunidad
 - Obtención de los genes involucrados en el funcionamiento del consorcio cianobacteria-bacterias, mediante minería genómica.
 - Descifrar el origen filogenético de los genes funcionales de la comunidad, así como su posible historia evolutiva
- Condensar la información de los objetivos anteriores en un núcleo funcional y otro taxonómico de la comunidad.

METODOLOGÍA

Colecta y procesado de material biológico

Colecta de muestras

Las poblaciones de cícdas silvestres utilizadas para el muestreo fueron seleccionadas de acuerdo a las coordenadas establecidas en trabajo previo de Lázaro-Zermeño (2011) el cual indica cómo encontrar las poblaciones de *D. merolae* en México (Figura 1). La colecta de *D. merolae* se llevó a cabo entre los meses de abril y mayo de 2015 en las localidades de Jiquipilas (JP; Lat 16.655, Long -92.39972222) y Raymundo Flores (RF; Lat 16.02611, Long -93.5991666), ambas pertenecientes al estado de Chiapas. Se tomó registro de las coordenadas de cada una de las poblaciones muestreadas y les fue asignada una clave y un número de control a cada planta.



Figura 1, Puntos geográficos donde se encuentra poblaciones de *D. merolae*. Tomado de Lazaro-Zermeño, 2011 y actualizado por Suarez Moo, 2016. Los recuadros de color azul muestran los sitios muestreados en este estudio.

Para tomar las muestras de raíz coraloide: i) se revisó si los ejemplares de cícdas tenían raíces superficiales, de no ser encontradas, ii) se excavó con pala

de mano previamente desinfectada (alcohol al 70%), y se removió el suelo de la base de la planta (30 cm aproximadamente alrededor de la planta) siguiendo la raíz primaria hasta que se encontraran las raíces coraloides; iii) una vez localizadas las raíces fueron cortadas desde su origen en la raíz lateral; iv) las muestras de raíz se limpiaron con agua destilada para eliminar el exceso de tierra y se colocaron en tubos Falcón de 15mL que contenían 5mL de agua destilada estéril para ser transportadas al laboratorio (en condiciones de oscuridad y a temperatura ambiente).

Lavado de raíces coraloides

En el laboratorio, las raíces coraloides fueron lavadas para eliminar microorganismos rizosféricos y dejar solamente a los microorganismos endófitos de la raíz, para esto se tomaron cada una de las raíces y se introdujeron en tubos Falcón de 50 ml con 10 ml de cada una de las siguientes soluciones siendo agitadas lenta y constantemente, como se describe a continuación: 3 minutos en solución de H₂O₂, 7 minutos en etanol al 70%, 30 segundos en agua MQ, 4 minutos en NaClO al 6%, y finalmente 3 lavados con agua MQ durante un minuto (Figura 2). El agua del último lavado se utilizó como inóculo de los medios usados como control negativo, si las placas del control negativo presentaban crecimiento bacteriano el material biológico era descartado. Todo este proceso se realizó en campana de flujo laminar para evitar contaminación biológica procedente del laboratorio.

El inóculo fue preparado macerando las raíces con agua MQ. Posteriormente, se tomaron 100µl de macerado de raíz y se sembraron en 6 medios de enriquecimiento que se describen en el siguiente apartado. El cultivo de bacteria en placas fue nombrado *t0*, mientras que los co-cultivos (el macerado de raíz inoculado directamente en medio líquido), fueron denominados *t1*. lo cual se partió el cultivo de bacterias en placas, al cual nombramos como *t0*. Para los co-cultivos se tomó macerado de raíz y se inoculo directo en medio líquido y se dejó crecer, a este método lo denominamos *t1*.

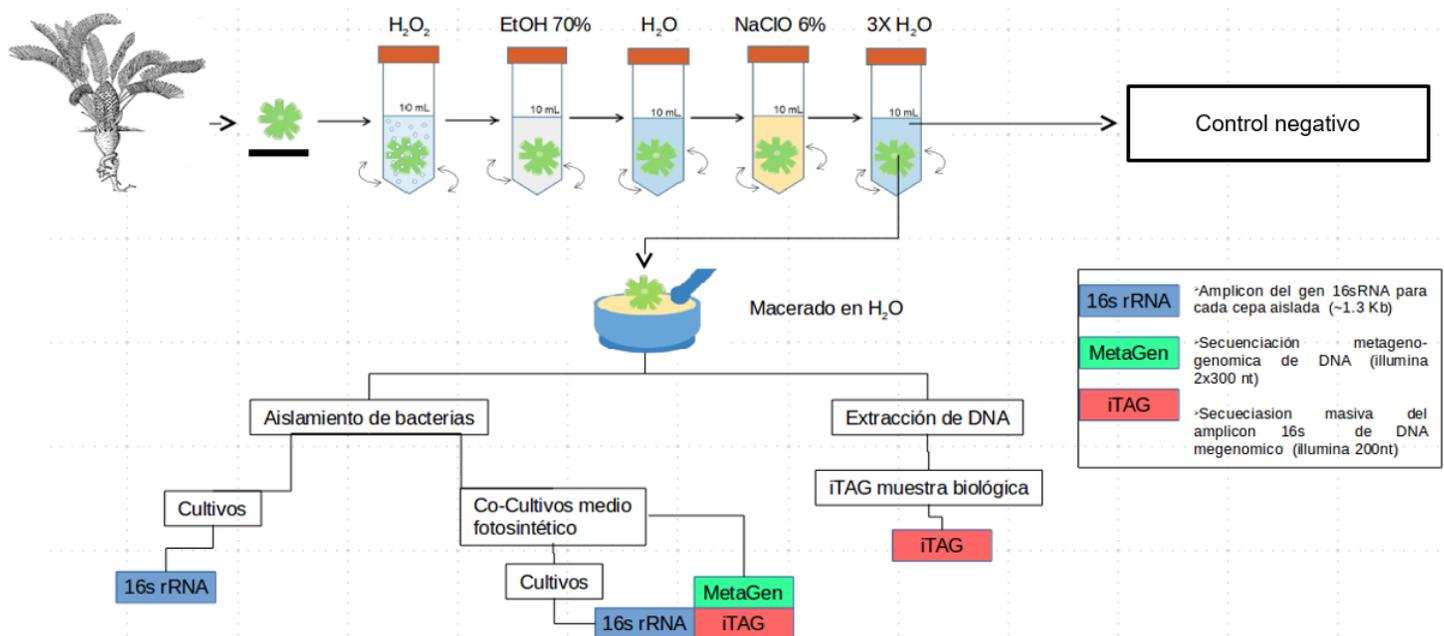


Figura 2. Diagrama de lavado y procesamiento de las raíces coraloides en laboratorio, se muestra el flujo de trabajo utilizados desde el lavado de la raíz hasta la obtención de los cultivos y co-cultivos así como el tipo de análisis molecular que se le hizo a cada uno.

Aislado de bacterias

El aislamiento de las bacterias endófitas de la raíz coraloides se llevó a cabo en distintos medios de cultivo para tratar de recuperar la mayor diversidad posible. Teniendo en cuenta trabajos previos con endófitos de otras plantas y lo que esperábamos obtener del interior de la raíz fueron seleccionados los medios i) Caulobroth (Medio 1), este medio rico fue escogido por el descubrimiento de caulobacter creciendo al lado de las cianobacterias simbiotes, se modificó a partir de "Caulobacter medium" de (Atlas, 2004); ii) Cycabroth (Medio 2), medio rico en carbohidratos, basado en el medio de *Caulobacter* pero adicionado con Manitol (Atlas, 2004); iii) Rhizobium Media (Medio 3), medio rico basado en el crecimiento de bacterias de suelo y endófitos (Subba & ATCC); iv) Nutrient Broth (Medio 4), este medio se utiliza para aislar gran variedad de bacterias (Difco nutrient broth); v) BG-11 (Medio 5), medio salino diseñado para el aislamiento de organismos fotosintéticos como cianobacterias y microalgas (Ripca et al 1979); vi) ISP Medium 4 (Medio 6), usados para el aislamiento de actinomicetos.

Se hizo un enriquecimiento de las bacterias de la raíz coraloides que consistió en colocar 100µl de macerado de raíz, (raíz previamente lavada), en 5 mL de medio

cultivo y se incubó a 25° C durante 48 y 96 horas respectivamente. Estos enriquecimientos fueron plaqueados en medio sólido usando diluciones seriadas para obtener colonias de bacterias separadas.

Crecimiento del Co-Cultivo

Inoculación

Con los co-cultivos se pretende hacer crecer a toda la comunidad bacteriana junta, o cultivar una subcomunidad a partir de la comunidad natural. Se parte de un inóculo de macerado de raíz coraloide previamente pasada por el protocolo de limpieza y asegurando que el control negativo estuviera limpio de bacterias de rizósfera. Para los experimentos se partió de macerado de raíz coraloide fresco, después se inocularon matraces con medio fotosintético y otros con medio eutrófico en una relación 1/1000 inóculo volumen de medio. La composición del medio fotosintético y eutrófico se describen en el siguiente apartado.

Medios

Todos los medios de los co-cultivos fueron escogidos, pensando en aumentar la biomasa bacteriana, para facilitar la extracción de ADN y los posteriores análisis genómicos. También se enriquecen las interacciones entre la comunidad bacteriana, dado que las bacterias crecen juntas en un ambiente común. Por lo que los medios escogidos para los co-cultivos fueron dos: I) (Cycabroth) Un medio rico, con el cual se espera seleccionar la mayor parte de la comunidad bacteriana oligotrófica (Atlas, 2004), II) (BG-11) Medio sin una fuente de carbono orgánica, el cual se usó para enriquecer a la comunidad microbiana que estuviera relacionada con las cianobacterias (Rippka *et al.* 1979).

Condiciones y tiempo de crecimiento

Los co-cultivos sin fuente de carbono se dejaron crecer 30 días en matraces con 150 mililitros de medio (BG-11), agitación constante y 16 horas luz 8 oscuridad,

para promover la fotosíntesis por parte de las cianobacterias. Para los co-cultivos en Cycabroth se dejaron crecer a 25 °C durante 60 horas con agitación constante en oscuridad.

Los co-cultivos fotosintéticos se dejaron en reposo y en condiciones de 16 horas luz 8 oscuridad, a una temperatura controlada de 25°C, sin agitación durante un año, lo cual generó un cultivo estable que se mantuvo vivo sin la adición de ningún nutriente o fuente de carbono al medio durante todo este periodo de incubación.

Extracción de ADN

Extracción de ADN de cultivos de cianobacterias

Del cultivo de cianobacterias crecido en medio BG-11, se transfirieron 30mL a tubos falcón de 50mL y se centrifugaron a 4000 rpm por 10 minutos para formar pallets. Los pallets de cianobacterias fueron resuspendidos con 250µL de agua destilada estéril y se procedió a la extracción de ADN utilizando el kit “ZR Fungal/Bacterial DNA MiniPrep” de ZYMO Research y siguiendo sus especificaciones. Posterior a la extracción el ADN se cargó en un gel de agarosa al 1% mas RedGel y se corrió en electroforesis por 45 min 100v y se observó que hubiera bandas no degradadas de más de 10Kb de tamaño; también el ADN se cuantificó por espectrofotometría a 260 nm (Nanodrop 2000, Thermo Scientific). Por último, las muestras fueron almacenadas a -20°C.

Extracción de ADN Co-Cultivos

Una vez crecidos los co-cultivos en diferentes medios, se tomaron 10 ml de cada uno y se transfirieron a tubos falcón de 50 ml, después fueron centrifugados a 7000 rpm durante 10 minutos con el fin de separar el contenido celular del medio, el sobrenadante fue decantado con cuidado para no suspender el pallet celular. Posteriormente, para resuspender el pallet celular, se agregó 1 ml de buffer de lisis (Triss 50 mM pH 8.0, EDTA 50 mM y SDS 3%). Para ayudar a lisar las células bacterianas, este contenido se pasó a un tubo eppendorf de 2 ml, se agregaron perlas de lisis 425-600 µm (SIGMA) y se procedió a dar 5 minutos de vórtex a

máxima potencia. Luego se incubó a 65 °C durante 3 horas. El resultado de la lisis fue centrifugado y el sobrenadante fue recuperado en un tubo limpio, al cual se le agregó un volumen de fenol cloroformo isoamílico (25:24:1). Para eliminar la mayor cantidad de proteínas y contenido celular la mezcla fue agitada suavemente durante 5 minutos. En seguida, esta mezcla de fenol fue centrifugada a 13000 rpm por 5 min. La fase acuosa fue colectada.

Para precipitar el ADN se utilizaron 250 µl de acetato de sodio 3 M, y 1 ml de etanol al 100% y se incubó a 4 °C por 30 minutos. Esta mezcla fue centrifugada a 14000 rpm por 10 minutos y a 4 °C. El sobrenadante fue decantado y el pellet resultante fue lavado cuidadosamente con etanol al 70% y se volvió a centrifugar a 14000 rpm por 5 minutos; el lavado se repite una vez más y el pellet fue secado con miVac (Accesolab). Por último, el ADN fue resuspendido en 100 µl H₂O milliQ y almacenado a -20 °C para su posterior análisis.

Como método alternativo de extracción de ADN, se utilizó el kit de extracción “FastDNA™ Spin Kit For Soil” (MP BIOMEDICALS), por su buena efectividad en extracción de ADN de suelos y muestras difíciles de lisar.

Secuenciación

Secuenciación 16S rRNA

Para la identificación molecular de los aislados bacterianos se utilizó una reacción de PCR para amplificar la región de ARN 16s ribosomal. Se utilizaron oligonucleótidos específicos para 16s de bacterias (Lane 1991; McDonald *et al.* 1999; Weinbauer *et al.* 2002), la reacción de PCR fue hecha con 50% Go-Taq Green Master Mix (Promega), 26% agua milliQ, 8% DMSO, 4% oligo F27 (5' AGA GTT TGA TCM TGG CTC AG 3') y 4% oligo R1492 (5' TAC GGY TAC CTT GTT ACG ACT T 3'), y 8% de DNA como templado, y una temperatura de alineamiento de 56 °C. Se amplificó un fragmento de aproximadamente 1300pb. Los productos de PCR fueron corridos en gel de agarosa al 1% con marcador 1Kb plus (Invitrogene) para asegurar que la PCR haya amplificado correctamente. El

producto de PCR se purificó con el kit de purificación ilustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) y el producto resultante se cuantificó en Nanodrop. Los fragmentos amplificados fueron secuenciados por ambos lados mediante el método Sanger. Este servicio fue proporcionado por el departamento de Servicios Genómicos de LANGEBIO.

Secuenciación metagenomas

El ADN metagenómico extraído de los co-cultivos fue cuantificado en nanodrop y posteriormente fue enviado a secuenciación masiva: MiSeq 2x300 y HiSeq del sistema Illumina, en Servicios Genómicos UGA (Tabla 2), donde se obtuvieron entre ~1.3 y 9.8 millones de lecturas por cada metagenoma.

Tabla 2 Análisis de calidad (FastQC) de los metagenomas obtenidos de los co-cultivos y secuenciados por Illumina

Metagenoma	Codificado	Lecturas Totales	Secuencias filtradas	Longitud de secuencia	%GC	Formato
RF1	Sanger / Illumina 1.9	2994716	0	35-301	60	MiSeq 2x300
RF3	Sanger / Illumina 1.9	1485660	0	35-301	63	MiSeq 2x300
JP2	Sanger / Illumina 1.9	5568140	0	35-301	49	MiSeq 2x300
JP6	Sanger / Illumina 1.9	1351348	0	35-301	65	MiSeq 2x300
RFPOOL	Sanger / Illumina 1.9	3309530	0	35-301	44	MiSeq 2x300
JPPOOL	Sanger / Illumina 1.9	2931034	0	35-301	39	MiSeq 2x300
RF3-oneyear	Sanger / Illumina 1.9	9875428	0	101	58	HiSeq 2x100
JP6-oneyear	Sanger / Illumina 1.9	9574972	0	101	64	HiSeq 2x100

Edición de secuencias

Los electroferogramas resultantes de la secuenciación Sanger se editaron en el programa Geneious (Kearse *et al.* 2012) mediante el cual se eliminaron las colas de secuenciación y se ensamblaron las secuencias directas y reversas para obtener un concatenado del cual se extrajo la secuencia consenso. Las secuencias que tenían inconsistencias en la lectura del electroferograma fueron Blasteadas (Buhler *et al.* 2007) contra la base de datos 16s de NCBI y se tomaron los mejores 5 hits, después se alinearon y se curaron las partes del

electroferograma que no fueran consistentes con la secuencia consenso del alineamiento.

Edición y limpieza de datos Metagenómicos

Los resultados de secuenciación masiva fueron analizados por FastQC v0.10.1 , de donde se obtuvieron el número de lectura, la calidad y el contenido de GC de cada metagenoma (Tabla 2), los gráficos resultantes se depositaron en la sección de anexos. Después de esto, los metagenomas se cargaron a MG-RAST (Wilke *et al.* 2013) donde se hizo anotado. También se cargaron a PATRIC donde se encuentran almacenado.

Análisis Taxonómico.

Una vez curados los datos de secuenciación se procedió a hacer un Blastn con un corte de *e-value* de 0.1 (Buhler *et al.* 2007) con las secuencias en la base de datos SILVA (Pruesse *et al.* 2007; Quast *et al.* 2013), de la cual se extrajeron los 10 mejores hits de cada secuencia, y se añadieron a un archivo junto a nuestras secuencias, posteriormente a este archivo fasta se le hizo un alineamiento con MUSCLE V 3.8.31 (Edgar 2004).

Mediante un programa de python v2.7.6, llamado IdentificadorRF.py (codigo fuente anexos1), el cual fue desarrollado para este proyecto, al Blastn se le identificó en qué posición se encontraban las secuencias, El programa después de identificar el sentido de las secuencias ordena todas de 5' a 3', para no generar errores en los alineamientos posteriores (Figura 3).

Para la identificación taxonómica se hizo la reconstrucción filogenética del alineamiento de secuencias, utilizando el programa MrBayes (V3.2.2 x64) (Ronquist *et al.* 2012), con una distribución gamma, a un millón de generaciones. Con el árbol filogenético resultante se le dió el nombre a las secuencias (dependiendo con cuál secuencia de la base de datos Silva se agruparon en el árbol). Para asignar el nombre (género), se tomó en cuenta el 97% de identidad a nivel de nucleótidos (Figura 3).

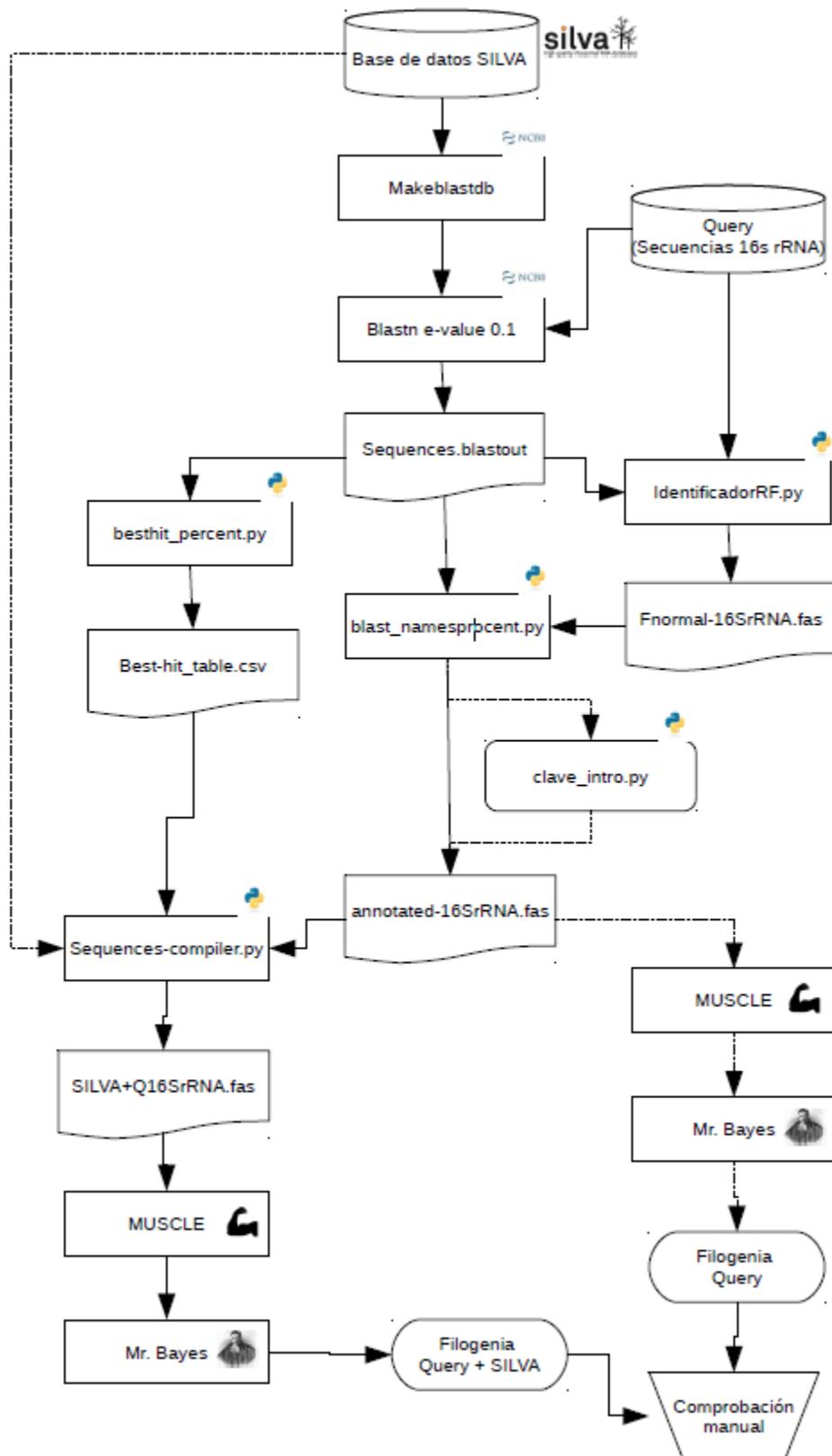


Figura 3 Método de identificación taxonómica para secuencias de 16S rRNA, en la cual se muestran los programas y los códigos de programación utilizados, el código se puede consultar en el anexo de anexos

Filogenia 16S rRNA

Una vez identificadas las secuencias taxonómicamente marcadas con su nombre de taxa y código de cepario, se hizo su alineamiento con MUSCLE V 3.8.31 (Edgar 2004) editando las colas de las secuencias para dejar un tamaño uniforme de 1100pb en todo el alineamiento y con este alineamiento se procedió a hacer la reconstrucción filogenética con el programa MrBayes (V3.2.2 x64) (Ronquist *et al.* 2012) con los parámetros antes mencionados.

Los programas que se usaron para la identificación taxonómica de las secuencias 16S rRNA, fueron agrupados en un solo fichero para hacer un programa de análisis automático, el cual es alimentado por la base de datos SILVA (Quast *et al.* 2013) y las secuencias Query obtenidas a partir de la secuenciación Sanger y curadas previamente (Figura 3). Toda esta línea de programas puede ser ejecutada de forma automática en una terminal Unix mediante un archivo o programa *shell* maestro (anexos). Los resultados principales que se obtienen de la ejecución del programa maestro son: Una tabla donde se colocan los mejores resultados de blast, porcentaje de identidad, tamaño del alineamiento, e-value y otros parámetros que ayudan a inferir la taxonomía en caso de duda. También se obtiene un árbol de las secuencias Query con su nombre y anotación, y otra filogenia con las secuencias más los mejores resultados de blast.

Al final se realizó una comprobación manual de los resultados, es decir una comparación de los árboles resultantes para ver la distribución y separación de los clados taxonómicos. Con la ayuda de la tabla de resultados de blast se pudo determinar con buena precisión el género de cada secuencia.

Árbol de árboles

Con la ayuda del paquete ETE (Environmental for Tree Exploration) (Huerta-Cepas, Dopazo and Gabaldón 2010) se generó un programa en python v2.7.6, que coloca un árbol base y en las ramas finales del mismo se colocan filogenias individuales

de cada género. Para generar el árbol base, se extrajo una secuencia consenso de cada uno de los géneros, mediante el alineamiento de las secuencias que se tienen por cada género. Todas las secuencias consenso se alinearon en MUSCLE (Edgar 2004) y la filogenia fue generada en MrBayes (Ronquist *et al.* 2012) con los parámetros anteriormente descritos. El árbol resultante se guardó en formato Newik. Para las filogenias de especies, se utilizó el mismo procedimiento de alineación y ensamblado de filogenia. Teniendo todas las filogenias fueron ensambladas por el programa que denominamos multi-tree.py (anexos).

Edición de árboles filogenéticos

La edición de los árboles filogenéticos se hizo con python v2.7.6 y la ayuda del paquete ETE (Huerta-Cepas, Dopazo and Gabaldón 2010), al cual mediante comandos se le fueron adicionando los metadatos más importantes, para ayudar a una mejor visualización de los datos. Los metadatos escogidos para aparecer en el árbol fueron: El tiempo donde se logró aislar la cepa, el ambiente de dónde provenía el material biológico, el medio donde fue crecida la cepa y por último, la clave del organismo de donde se tomó el inculo de la raíz coraloide. Los nombres de los géneros fueron corroborados por el árbol de árboles (Figura 4) y les fue asignado un color para diferenciarlos. Todos estos pasos quedaron codificados en el programa ETE.py (Anexos).

Análisis taxonómico de Metagenomas

Luego de descargar las secuencias de los metagenomas en la computadora de trabajo (AMD FX-8320) se procedió a encontrar qué organismos conformaban el metagenoma. Esto se hizo mediante marcadores filogenéticos universales de copia única para bacterias, siguiendo el método propuesto por Sunagawa *et al.* 2013 (Sunagawa *et al.* 2013), el cual está basado en generar unidades taxonómicas operacionales a partir de metagenomas (mOTUS). El método extrae de los metagenomas, 40 genes marcadores que han sido definidos previamente para identificar especies de procariotas (Mende *et al.* 2013). Para este estudio se

utilizaron 10 marcadores que daban los mejores resultados (Tabla 3) Posteriormente, se definieron los perfiles taxonómicos basados en el modelo de Markov que utiliza los datos disponibles de metagenomas y genomas referenciados para inferir la taxonomía. Este análisis fue usado hasta nivel de genero con un 93% de identidad de las secuencias para que los datos sean comparables con los producidos por las cepas cultivables.

Tabla 3. Marcadores de filogenéticos de copia única utilizados para análisis mOTUs

Clave	Marcador
COG0012	<i>Predicted GTPase, probable translation factor</i>
COG0016	<i>Phenylalanyl-tRNA synthetase alpha subunit</i>
COG0018	<i>Arginyl-tRNA synthetase</i>
COG0172	<i>Seryl-tRNA synthetase</i>
COG0215	<i>Cysteiny-tRNA synthetase</i>
COG0495	<i>Leucyl-tRNA synthetase</i>
COG0525	<i>Valyl-tRNA synthetase</i>
COG0533	<i>Metal-dependent proteases with possible chaperone activity</i>
COG0541	<i>Signal recognition particle GTPase</i>
COG0552	<i>Signal recognition particle GTPase</i>

Análisis ecológicos y de composición de comunidad

Los análisis taxonómicos se hicieron en el programa estadístico R V3.0.2 (Team 2014) con ayuda del paquete “Vegan” V2.3-5 (Oksanen *et al.* 2007). Los análisis realizados fueron los siguientes:

1. Curva de acumulación de especies

Se utilizó el método de rarefacción basado en el número de proporcional de secuencias para cada metagenoma (Definir bien el valor), donde la riqueza de mOTUs es sub-muestreada aleatoriamente del total de la comunidad y el tamaño de la muestra debe ser más pequeño que el tamaño total de la comunidad. La función de la rarefacción se basó en Hurlbert's (1971).

2. Índices de diversidad

Se escogió el índice de Shanon-Weaver H' (1962) porque es una manera de representar la diversidad específica. El índice refleja la heterogeneidad de una comunidad sobre la base de dos factores: el número de especies presentes y su abundancia relativa. Conceptualmente es una medida del grado de incertidumbre asociada a la selección aleatoria de un individuo en la comunidad (Pla 2006).

Para representar mejor la diversidad se escogió un segundo método, el índice de Simpson D' (1949) o índice de dominancia, Donde se toma el número de especies encontradas en el muestreo y su abundancia relativa y con ellos se calcula la probabilidad de tomar dos individuos al azar y que estos pertenezcan a la misma especie.

Análisis Funcional de Metagenomas

Para describir los microbiomas y su estructura es necesario saber que especies o taxa la componen. Normalmente se han usado métodos basados en el cultivo de bacterias y la amplificación del gen ribosomal 16S (Ward, Weller and Bateson 1990). Para estudiar interacciones y funciones es difícil interpretarlo solo con un análisis de 16S rRNA, se necesita explorar los genes funcionales que componen al microbioma y pruebas de laboratorio (Boon *et al.* 2014). Se utilizó MG-Rast V 3.3.6 (Wilke *et al.* 2013) como método de anotación y ensamblado para los metagenomas, MG-Rast comienza con un control de calidad (las secuencias con mala calidad son eliminadas). Posteriormente hace una predicción de proteínas, que son agrupadas por función y luego son anotadas en base a similitud a las secuencias que se encuentran almacenadas en su base de datos. para hacer esto se vale de herramientas bioinformáticas diseñadas para el análisis de metagenomas secuenciados por *shootgun metagenomics*.

Posteriormente los resultados de la anotación funcional de MG-Rast, fueron descargados en forma de tabla con la nomenclatura de función usada por Pfam (Sonnhammer, Eddy and Durbin; Bateman *et al.* 2002). Se cargaron los datos a el programa STAMP (Parks and Beiko 2010a), el cual es un programa que mediante una interfaz gráfica se pueden explorar datos metagenómicos desde el punto de

vista taxonómico y funcional. Con ayuda de algoritmos estadísticos, en él se hizo el gráfico de dispersión donde se compararon ambos ambientes, los metagenomas de 1 año contra sus originales y por último los metagenomas donde se mide la actividad de fijación de nitrógeno.

Posterior a esto se hizo un análisis de componentes principales a partir de los datos de metagenómica, donde se pusieron a prueba los resultados de anotación taxonómica como los de anotación funcional, con el objetivo de encontrar patrones entre los metagenomas que nos ayudaran a contestar la pregunta de investigación.

Para visualizar las rutas metabólicas que se compartían entre los metagenomas se utilizó el visualizador de rutas metabólicas de KeggMapper (Kanehisa *et al.* 2011), donde a partir de los datos de anotación de los metagenomas genera un mapa de las posibles rutas metabólicas, en este programa cargamos los metagenomas de los co-cultivos fotosintéticos de un mes, y comparamos entre los dos ambientes que rutas metabólicas se sobreponían. Como valores se usó un corte de *e-value* $1e-5$ y una identidad mínima del 60%.

RESULTADOS TESIS

Colecciones Biológicas de raíces coraloides

En las colectas que se efectuaron en el estado de Chiapas se lograron coleccionar seis plantas adultas de *D. merolae* con raíces coraloides desarrolladas, las cuales fueron procesadas como se indica en la metodología. Las muestras se originaron en dos ambientes distintos: JP identificado como matorral xerófilo con 560 metros sobre el nivel del mar con una precipitación promedio 320 mm y temperatura media entre los 18° C, mientras RF se identificó como selva baja perennifolia con 900 metros sobre nivel del mar precipitación promedio de 2500mm y temperatura media de 25° C (Lázaro Zermeño *et al.* 2011).

Se obtuvo un número limitado de muestras de raíz coraloide debido a la complejidad inherente a encontrar las raíces en organismos de vida silvestre. Ya que a pesar que las comunidades cuentan con gran número de ejemplares de *D. merolae*, solo se logró revisar unas 40 plantas por sitio, de las cuales la mayoría fueron de difícil acceso dado que crecen entre rocas y pendientes muy pronunciadas, y las raíces coraloides no se encuentran muy expuestas ya que las raíces principales pueden ser muy profundas. En 12 de las plantas se logró identificar las raíces coraloides la mayoría de estas eran ejemplares juveniles, algunas de las raíces fueron encontradas secas por lo que se descartaron por lo que solo se obtuvieron seis muestras frescas, tres de cada ambiente respectivamente.

Las seis muestras obtenidas en campo, posterior a ser lavadas como se indica en la sección de métodos, se les hizo un control de contaminación microbiana, donde nos aseguramos que las raíces no generaran unidades formadoras de colonia en las cajas control las cuales fueron sembradas con el agua del último lavado, lo que nos aseguró que la limpieza de las raíces coraloides fue exitosa y se lograron eliminar los organismos de la rizosfera. De este modo todos los organismos que se logren identificar en los resultados, son de la endosfera de la raíz coraloide. Se hizo mucho énfasis en este control, por el hecho de que queremos ver

exclusivamente endófitos que se encuentran en la raíz coraloide.

Colectas bacterianas y co-cultivos

Se montaron seis co-cultivos fotosintéticos a partir de cada una de las muestras de raíces coraloideas, posterior a un mes de crecimiento (métodos) se detectó una falta de crecimiento en los co-cultivos JP1 y RF9, por lo que se descartaron para el análisis de Metagenómica por la falta de cianobacterias en el co-cultivo lo cual se corroboró mediante microscopía, aun así, se procedió a aislar las bacterias que se pudieran encontrar en estos. No se tiene una respuesta clara de por qué en estos co-cultivos no se desarrollaron las cianobacterias, podrían ser múltiples factores entre ellos problemas de adaptación al medio, o que el lavado pudo dañar a las bacterias del interior.

De los cuatro co-cultivos fotosintéticos que dieron resultado positivo a la presencia de cianobacterias por el color verde azul del medio y microscopía (JP2, JP6, RF1 y RF3). Se obtuvo a partir de ellos glicerol con el objetivo de su preservación para posteriores análisis. También se obtuvieron los inóculos que sirvieron para el aislamiento de las cepas pertenecientes al tiempo t1, y ADN con el cual a partir de secuenciación masiva se obtuvieron sus metagenomas. Dos de los co-cultivos (JP6 y RF3) se dejaron desarrollar a través de 1 año entero en condiciones de laboratorio como se describe en los métodos y posterior a esto se extrajo ADN y se hizo la secuenciación masiva y se obtuvieron sus metagenomas.

Independientemente a los co-cultivos fotosintéticos se hicieron dos co-cultivos en medio oligotrófico (Cycabroth), los cuales partieron de pools de las muestras: JPPool (JP2, JP6), RFPool (RF1 y RF3). Estos co-cultivos al no depender de la luz y el CO₂ como fuente de energía crecieron más rápido, solo se necesitó 60 horas de incubación (métodos), posterior a esto se generaron glicerol de los co-cultivos para su posterior análisis, También se obtuvieron los metagenomas a partir de extracciones de ADN.

Cepario

Se lograron aislar bacterias tanto de los co-cultivos como de los macerados de raíz t1 y t0 respectivamente, con el uso de medios distintos (métodos). El total de cepas aisladas fue de 470 las cuales fueron registradas con una clave y guardadas en glicerol a -80 °C (Anexos Cepario general XLS) no se esperaba un número tan amplio de bacterias dado que lo que conocíamos por la literatura (Lobakova, Dol'nikova and Korzhenevskaya; Rai *et al.* 2002; Gehringer *et al.* 2010). No se pudo analizar a todas las bacterias que se lograron encontrar por cuestiones del tamaño del análisis y no sabíamos que diversidad podríamos encontrar, por lo que aproximadamente solo la mitad del cepario conseguido fue identificado con el marcador 16s rRNA mediante PCR. El resto se guardo para futuros estudios y de ser necesario experimentar con ellas.

Taxa identificados

Identificación taxonómica mediante el marcador 16S rRNA

El uso del marcador 16S rRNA ha sido muy extendido para la identificación taxonómica de bacterias, el gen 16S rRNA es un marcador filogenético que tiene una tasa de cambio muy baja dado que es vital para la síntesis de proteínas, estas características lo hacen un buen blanco de amplificación por PCR ya que esta conservado en todas las bacterias. Pero hay que tomar en cuenta que este marcador por su propia naturaleza de baja tasa de cambio, vuelve imposible separar algunos grupos bacterianos, en este estudio encontramos el problema en diferenciar géneros bacterianos como, por ejemplo, *Rhizobium* de *Agrobacterium* y otros casos.

Para mejorar el análisis de identificación desarrollamos nuestro propio método bioinformático de análisis del gen 16S rRNA (métodos) que combina las dos herramientas más ampliamente utilizadas para identificación taxonómica, y compensa algunas de las desventajas de cada una. La primera herramienta es BLAST (Altschul *et al.* 1990), un algoritmo de búsqueda muy poderoso que nos permite buscar en grandes bases de datos a una buena velocidad, su problema es

que solo se basa en la homología de las secuencias y no toma en cuenta inferencias filogenéticas. La segunda herramienta es de naturaleza filogenética donde usamos inferencia bayesiana para hacer arboles filogenéticos (Ronquist *et al.* 2012) de los resultados de BLAST, lo que permite encontrar entre los homólogos de la secuencia problema cuales son las más cercanas filogenéticamente. Por ultimo juntamos estos dos análisis en un árbol que se divide por géneros bacterianos y por cada género hace una filogenia, lo que hace que dentro de cada género sean más notorias las distancias filogenéticas. Cada uno de los géneros es ordenado en su rama correspondiente en el árbol central y así se genera un árbol de árboles (Figura 4) el cual se puede consultar visualmente y asegurar que la anotación taxonómica que estamos dando es la correcta.

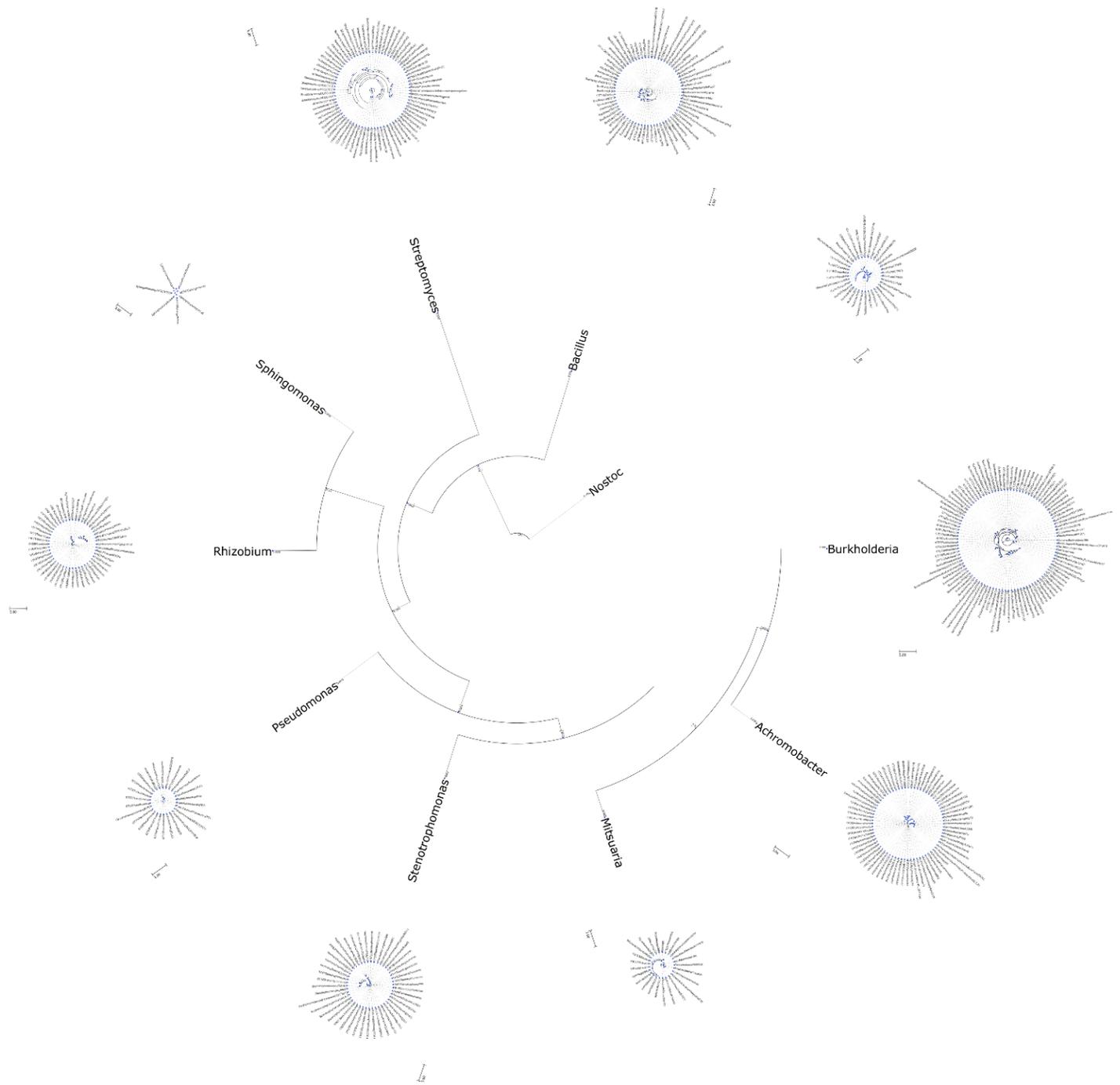


Figura 4. Árbol de árboles, el árbol central se hizo concatenando las secuencias de cada género y obteniendo un consenso. los árboles que se encuentran en cada una de las hojas se hicieron combinando la secuencia problema con los mejores resultados de BLAST de la base de datos SILVA. Por la gran resolución de la figura se agrega cada filogenia por separado en anexos.

El árbol de árboles presenta los 10 géneros de bacterias más abundantes que encontramos en los aislados de la raíz coraloide. También muestra las secuencias más cercanas a cada género extraídas de la base de datos SILVA (Pruesse *et al.* 2007; Quast *et al.* 2013) por homología. Se destaca *Nostoc* como el más basal de los géneros bacterianos, los géneros que mejor se dividen taxonómicamente son *Burkholderia* y *Streptomyces* mientras que en otros casos como el de *Rhizobium* y *Pseudomonas* se forman peines que no permiten una buena separación entre especies. Este árbol permitió corroborar que la anotación taxonómica utilizada fuera correcta.

Taxa identificados en *t0*

De los aislados de macerados de muestras de raíces coraloide fueron identificados molecularmente hasta género 121 cepas. Entre los géneros identificados se destacan por su gran abundancia los géneros: *Stenotrophomonas*, *Rhizobium* y *Achromobacter*, los cuales conforma el 76.8% de los OTUS identificados para *t0*, mientras que el 23.2% restante está conformado por: *Dyadobacter*, *Luteibacter*, *Microbacterium*, *Mitsuaria*, *Mucilaginibacter*, *Ochrobactrum*, *Shinella*, *Sphingobium*, *Variovorax* y *Xanthomonas* (figura árbol general). Se reconocieron un total de 14 géneros, cabe destacar que no están incluidas las cianobacterias ya que el medio fotosintético usado para aislarlas fue inoculado con los co-cultivos (*t1*).

Taxa identificados en *t1*

Las cepas aisladas a partir de los inóculos tomados de los co-cultivos fotosintéticos fueron denominadas *t1* por la abreviación de tiempo1, donde el co-cultivo pasó un tiempo de enriquecimiento en un medio fotosintético sin fuente de carbono orgánico esto forzó a la comunidad a depender de organismos fotosintéticos como las cianobacterias, de la cual se sabe que juega un papel

clave en la simbiosis que se presenta en la raíz coraloide. Para aumentar la resolución, el crecimiento y forzar las interacciones los co-cultivos se dejaron crecer más de 30 días.

El uso de los co-cultivos permitió la obtención de mayor diversidad de bacterias, aumento de interacciones y de dependencia de las bacterias y su habilidad para crecer en comunidad ya sea aprovechándose de los nutrientes que ofrece cada una o formando parte de cadenas metabólicas lo cual les permite crecer.

A partir de los co-cultivos se lograron identificar 120 cepas hasta el nivel de género, de los cuales los más abundantes fueron *Burkholderia*, *Pseudomonas*, *Streptomyces* y *Bacillus* que representan el 79.1% de los OTUS identificados para *t1* y el porcentaje restante está compuesto por los géneros: *Achromobacter*, *Enterobacter*, ***Nostoc***, *Pantoea*, *Peanibacillus*, *Ralstonia*, *Sphingomonas*, *Stenotrophomonas* y ***Tolypothrix***. Entre los grupos menos abundantes se resalta la presencia de las cianobacterias *Nostoc* y *Tolypothrix*, esto puede deberse a la complejidad de aislar a las cianobacterias del resto de las bacterias ya que son de lento crecimiento comparadas con las demás bacterias; además, tienen medios muy exclusivos, en este proyecto o sólo se usó BG-11 Agar para aislarlas por medio de estriado del resto de las bacterias (Stanier *et al.* 1979).

Se observó que sólo las cianobacterias y *Streptomyces* fueron exclusivas a un medio, mientras que todas las demás bacterias se distribuyen sin un patrón específico con respecto a los otros cinco medios usados. El uso de diversos medios fue una prueba ya que no se sabía qué esperar de las bacterias de la raíz coraloide. Para futuros trabajos se recomienda el uso de medio agar nutritivo como medio oligotrófico y generalista. Mientras que el uso de medios especializados es recomendado para aumentar la biodiversidad (medios fotosintéticos y diazotróficos, que seguramente capturará a partes de la comunidad con las habilidades de crecer dichos medios).

Taxa identificados por metagenómica

Para identificar la composición taxonómica de los metagenomas se utilizaron 10 marcadores de copia única (Tabla 3) de los 40 identificados en el trabajo de Sungawa (Mende *et al.* 2013; Sunagawa *et al.* 2013), ya que son excelentes marcadores taxonómicos y se han extraído de los genomas anotados en la base de datos NCBI. El total de genomas a la fecha de estos marcadores es de 1820 por lo que todas las similitudes están ligadas a este número de genomas. Para este trabajo se decidió tomar como parámetro el 93% de identidad, para identificar hasta nivel de género (generalmente este porcentaje de similitud está muy ligado al nivel taxonómico de los géneros bacterianos) (Ward, Weller and Bateson 1990; Lane 1991).

A partir de los datos de metagenómicos extraídos de los co-cultivos, se lograron identificar 77 géneros bacterianos, pertenecientes a 34 familias de 18 órdenes bacterianos (Figura 6). En los co-cultivos fotosintéticos encontramos que las taxa más comunes fueron: *Pseudoxanthomonas*, *Stenotrophomonas*, *Rhizobium*, *Methylobacterium*, *Agrobacterium*, *Bradyrhizobium*, *Starkeya*, *Rhodopseudomonas*, *Mesorhizobium*, *Microbacterium*, *Burkholderia*, *Nostoc* y *Sphingomonas*. Mientras que para el medio oligotrófico los géneros más numerosos fueron *Bacillus*, *Pseudomonas*, *Burkholderia* y *Enterobacter* (Figura 6).

Taxa identificados en metagenomas de 1 año

En los metagenomas que se obtuvieron después de un año de incubación de los co-cultivos JP6 y RF3 se lograron identificar 12 géneros nuevos entre los cuales destaca la presencia de *Anabaena* y *Streptomyces*, mientras que 24 géneros se volvieron a recuperar. Cabe destacar que los medios se mantuvieron en reposo y condiciones constantes durante el año del experimento y no fueron abiertos ni manipulados durante este tiempo. También se debe mencionar que el método de secuenciación dió diferentes cantidades de lecturas por la variabilidad intrínseca que hay en la secuenciación (Tabla 2).

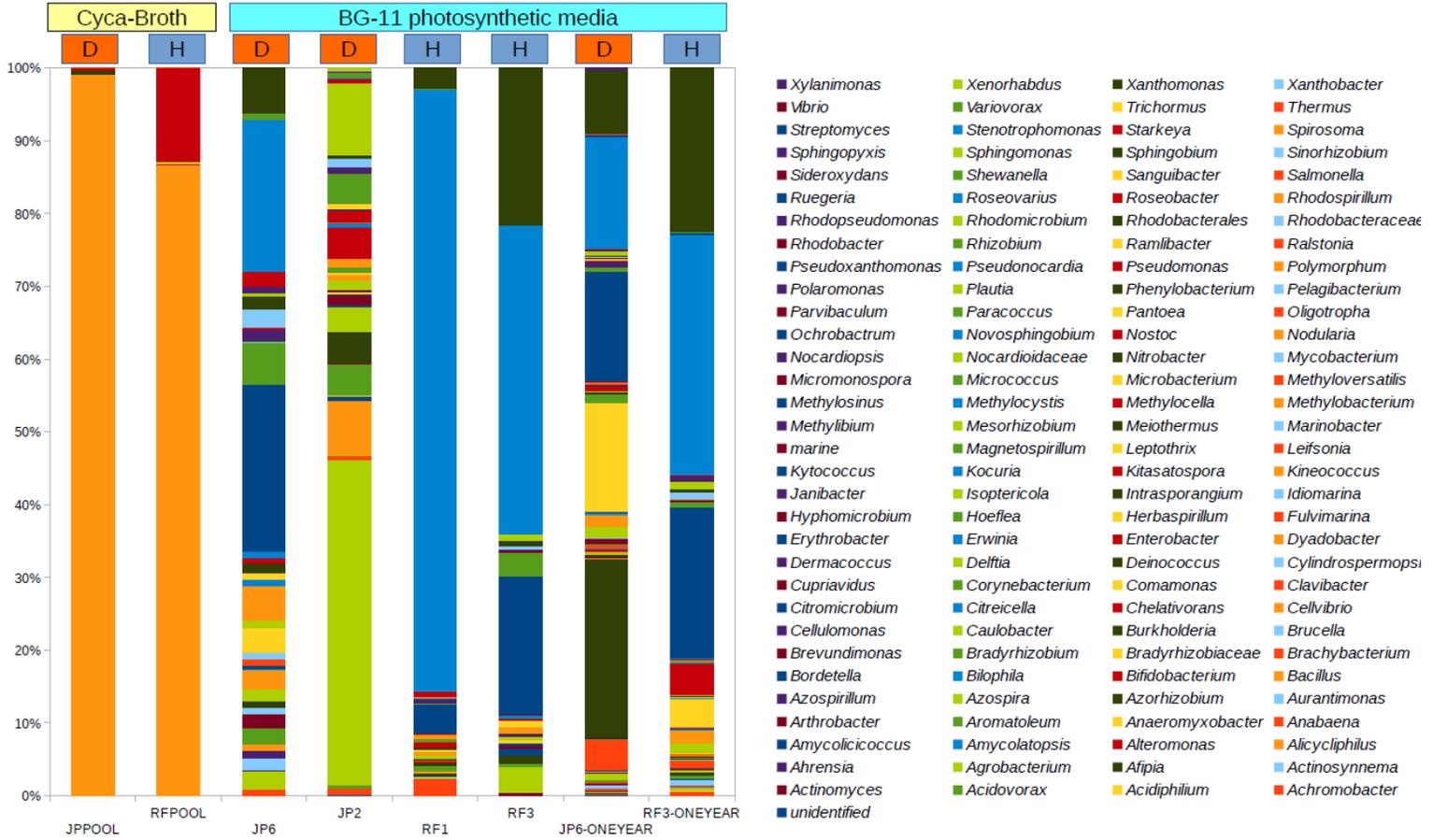


Figura 6. Diagrama de barras de la proporción de cada taxón en cada uno de los 8 metagenomas, echo con la anotación de MOTUs usando 10 marcadores universales.

Comparación Metagenómica y 16s rRNA: diferencias entre métodos

Hay diferencias entre el método de aislado de la bacteria y el uso de co-cultivos. A nivel de cepas aisladas podemos ver que entre t_0 y t_1 hay una gran diferencia en cuanto a los géneros que se pueden recuperar, (Figura 12), también se presentan diferencias con respecto a los índices de diversidad. Normalmente las especies que se pueden encontrar en t_0 son aquellas que se adaptan bien a ambientes ricos en nutrientes como el medio agar nutritivo y tienen una tasa de crecimiento alta como *Stenotrophomonas*, *Rhizobium* y *Achromobacter*. En los co-cultivos oligotróficos también se destacaron como géneros dominantes *Pseudomonas* y

Bacillus debido a que su tasa de crecimiento fue mayor y por ende pudieron aprovechar la alta disponibilidad de nutrientes del medio. Mientras que en el co-cultivo fotosintético no podía haber crecimiento de esta manera, ya que los nutrientes se encuentran registrados a cantidades de sales mínimas y sin una fuente de carbono orgánico, por lo que vemos que hay un enriquecimiento de más grupos que crecen lentamente, porque tienen más tiempo para desarrollarse y tienen que interactuar con más organismos en la comunidad para sobrevivir.

Si se consideran todos los aislamientos de bacterias como un muestreo biológico de la raíz coraloide, se puede realizar una curva de acumulación de especies, encontrando que estos aislados se encuentran lejos del número total de especies, (figura acumulación de especies 16s). Lo anterior indica que si se siguen aislando bacterias es muy probable que se sigan encontrando géneros nuevos. Esto se confirma con los resultados taxonómicos de los metagenomas. No hay que subestimar la técnica de aislamiento, ya que no solo proporciona la composición taxonómica (conocer parte de la diversidad), también nos permite tener las bacterias que la componen para poder experimentar con ellas.

Diversidad general de la raíz coraloide de *D. merolae*

Entre los distintos métodos encontramos 128 géneros bacterianos, resultado bastante contrastante con lo previamente reportado (Tabla 1). Se destaca el orden de las Proteobacteria como las más diversas y abundantes en la raíz coraloide. Sin embargo, también se encuentra la presencia de Firmicutes, Cianobacterias, Bacteroides, Proteobacterias y Actinobacterias lo que demuestra que la raíz coraloide tiene una dinámica comunidad de endófitos bacterianos. Este estudio fue realizado con una sola especie de las 331 especies de cícadas que hay en el mundo, por lo que aún falta investigar más para terminar de conocerla composición de la raíz coraloide y quiénes son sus integrantes principales.

Rarefacción

Los resultados de las curvas de rarefacción muestran de manera probabilística, la cantidad de géneros nuevos que se descubren conforme al número de cepas (en el caso de 16S rRNA) u homologías con marcadores taxonómicos (en el caso de mOTUs). Para el caso del análisis de 16s tenemos un total de 242 OTUs identificados, distribuidos en 26 géneros. Al observar la curva general se encuentra una pendiente muy pronunciada (Figura 7), por lo que se infiere que de haber seguido aislando e identificando bacterias se hubiera seguido encontrando géneros nuevos. Las curvas por muestreo o por ejemplar (Figura 7) muestran la cantidad de géneros identificados contra el número de individuos que se lograron analizar, siendo el más prominente JP6 con 15 géneros identificados. Para JP6, RF3, RF1 y JP2 se analizaron en promedio 50 aislados por cada uno, mientras que JP1 y RF9 tuvieron un menor número de aislados dado que el co-cultivo no fue viable.

En el caso de los co-cultivos que analizamos con mOTUs, incluyendo los 8 metagenomas tenemos un total de 128 géneros identificados, a partir de más de 37 millones de lecturas. En este caso por la naturaleza del análisis vemos un mayor número de géneros. Se destaca que las curvas correspondientes a JP2 y JP6 son las más prominentes y son muy similares. También se observa similitud entre RF1 y RF3 (Figura 7). La identificación de géneros más bajos se encuentra en los metagenomas provenientes de los co-cultivos oligotróficos (RFPOOL y JPPOOL). En los co-cultivos de un año vemos que en comparación a sus contrapartes de 1 mes hubo más especies encontradas como lo vemos en la curva de rarefacción (Figura 7).

Composición taxonómica de ambientes

Entre los aislados de t_0 y t_1 se observan taxa únicos para cada ambiente *Stenotrophomonas*, *Rhizobium Stenotrophomonas*, *Rhizobium* y *Mucilaginibacter*. Además, se observan otros géneros que son compartidos *Bacillus*, *Achromobacter*, *Burkholderia*, *Pseudomonas*, *Mitsuaria* y las cianobacterias (*Nostoc* y *Tolypothrix*). Las diferencias entre ambientes pueden referirse a una selección por ambiente, en la cual la cícada recluta las bacterias dependiendo del ambiente en que se encuentre (diagramas de Ven ambiente). En el orden de las cianobacterias se encontró en un ambiente *Tolypothrix* y en otro *Nostoc*, por lo tanto, se infiere que podría ser una selección por función más que específica a taxonomía. En investigaciones previas otros autores ya han reportado que la simbiosis de las cianobacterias con las cícadas es abierta (Gehring *et al.* 2010), y que las raíces coraloides pueden reclutar diferentes grupos de bacterias (Zheng *et al.* 2002; Costa, Romero and Lindblad 2004). El aporte de esta investigación sobre el conocimiento de las raíces simbióticas de cícadas es que se cuenta con

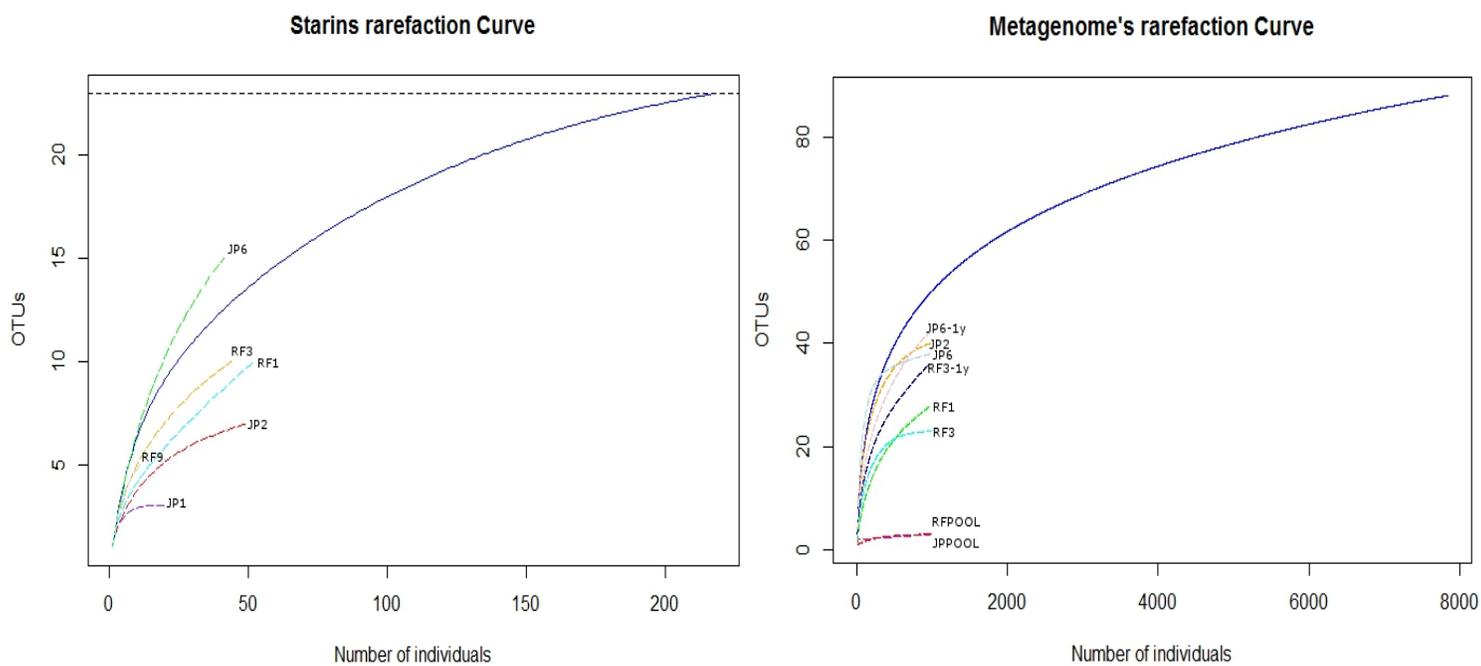


Figura 7. Curvas de rarefacción de cada una de las raíces, a la izquierda están las curvas generadas por las bacterias identificadas por el gen 16S y a la derecha las identificadas dentro de los metagenomas con los marcadores del método mOTUs.

el reporte completo del microbioma de la raíz coraloide para *D. merolae*.

También se encontraron diferencias con respecto a los individuos muestreados, lo que revela que el sistema es dinámico, es decir, que carece de un perfil taxonómico estático pues encontramos más homologías entre los organismos del ambiente húmedo que con los del ambiente seco (Figura 8). Las raíces coraloides demostraron tener un microbioma muy variable con respecto a otros órganos simbióticos como los nodos de las leguminosas (Rosenblueth and Martínez-Romero 2006; Hallmann *et al.* 2011; Hardoim *et al.* 2015).

En los perfiles taxonómicos de los co-cultivos se encuentra una diferencia marcada entre ambientes que apoya lo encontrado a nivel de aislados, donde se observa una mayor diversidad en el ambiente seco, y se tienen especies que se comparten entre los ambientes (Tabla 4). Sin embargo, parece que algunas son exclusivas de cada ambiente. Estos resultados ofrecen la primera pauta de la diversidad real que podríamos esperar en las raíces coraloides de diversas cícadas.

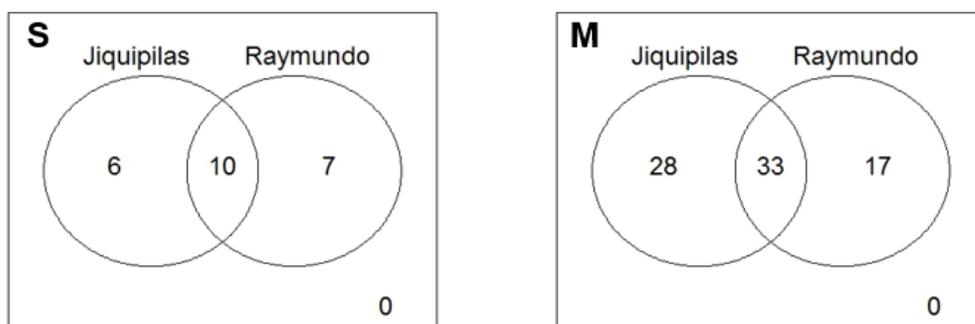


Figura 8. Diagrama de Venn donde se comparan el número de géneros entre ambientes, Jiquipilas = Seco, Raymundo = Húmedo. a la izquierda están los datos generados por las bacterias identificadas por el gen 16S (S) y a la derecha los datos generados por los metagenomas con los marcadores del método mOTUs.

Índices de diversidad entre los diferentes ambientes

Los índices de diversidad son una manera de representar de manera numérica y proporcional la diversidad de un ambiente. Utilizando los géneros identificados por 16S rRNA y mOTUs pudimos calcular los índices de diversidad para la raíz

coraloide de *D. merolae*. En ambos análisis se nota que el ambiente seco (JP) es más diverso que el ambiente húmedo, lo que muestra que hay diferencia entre la composición del microbiana en distintos ambientes. La comparación entre métodos demuestra grandes diferencias por método (Tabla 4).

Tabla 4 Índices de diversidad de cada una de las muestras de aislados y co-cultivos

Muestra	Shannon	Simpson	Gama	Método
JP1	0.95027054	2.27272727	3	16s rRNA
JP2	1.2559371	2.38095238	7	
JP6	2.29827783	6.75100402	15	
RF1	1.51177868	3.1588785	10	
RF3	1.78397463	4.34080717	10	
RF9	1.46813994	3.90322581	5	
RFPOOL	0.39301096	1.29372497	3	mOTUs
JPPOOL	0.04597413	1.01414709	3	
RF1m	0.84286465	1.40915124	28	
RF3m	1.71955768	3.68832229	23	
JP2m	2.25594583	4.30289847	40	
JP6m	2.77053588	8.63158225	38	

Composición funcional del microbioma de la raíz coraloide

Funciones importantes

En los metagenomas pertenecientes a los co-cultivos fotosintéticos (JP2, JP6, RF3 y RF1) se identificó un total de más de 30 mil genes. Los grupos principales de genes fueron del metabolismo de proteínas, aminoácidos y derivados, metabolismo de carbohidratos y subsistemas basados en cluster. Los genes que se encuentran en menor proporción con respecto del total son los relacionados con la función de la fotosíntesis, metabolismo secundario, metabolismo de potasio y metabolismo de nitrógeno (Figura 9).

Se destaca la baja proporción de los genes relacionados con la fotosíntesis a pesar de que en todos los metagenomas provenientes de los co-cultivo fotosintéticos se detectó la presencia de cianobacterias (Figura 6) y se sabe que las cianobacterias forman la base de la cadena alimenticia de estos co-cultivos ya

que carecen de una fuente de carbono orgánico. Esto se pudo comprobar claramente en los metagenomas oligotróficos (JPPOOL y RFPOOL), donde dada la naturaleza del medio no se tienen señales de la presencia de cianobacterias o de funciones relacionadas con la fijación de nitrógeno. Lo anterior indica que las cianobacterias solo son una pequeña parte de la complejidad funcional de los metagenomas.

Los genes involucrados con el metabolismo del nitrógeno y de la fotosíntesis se encuentran en menor proporción; los genes de fijación de nitrógeno que se lograron anotar están relacionados filogenéticamente con los de las cianobacterias. En los metagenomas donde no se encuentra la presencia de las cianobacterias (JPPOOL y RFPOOL) no se haya la presencia de genes relacionados con la fijación de nitrógeno, esto da un argumento válido para afirmar que las cianobacterias son la base de la simbiosis en la fijación de nitrógeno.

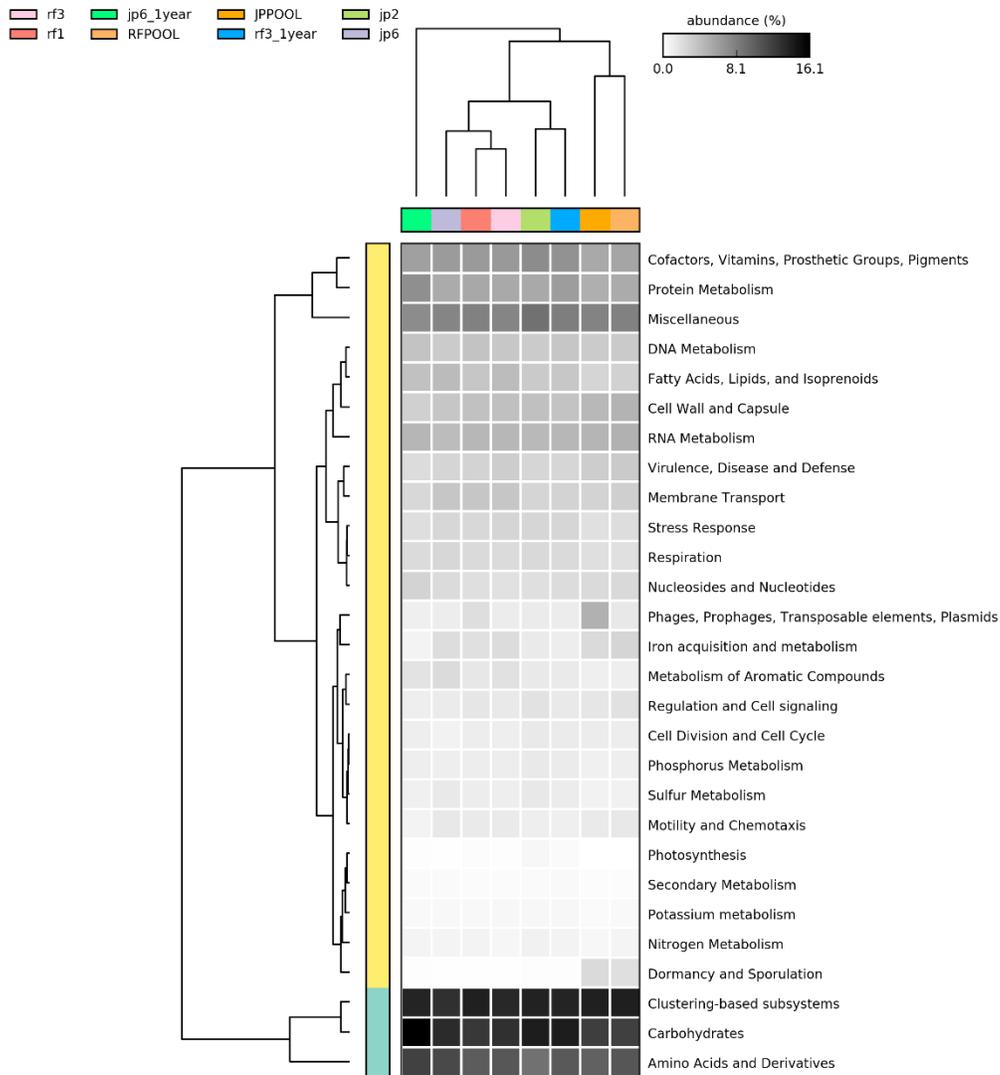


Figura 9 Mapa de calor hecho con la anotación funcional de los metagenomas. tomado el porcentaje de cada categoría funcional dentro del metagenoma.

Composición funcional en diferentes ambientes

Se compararon los metagenomas pertenecientes a los co-cultivos fotosintéticos (JP2, JP6, RF3 y RF1) por ambiente, ya que estos demostraron tener una gran diversidad (Figura 6) y en su metodología sólo difieren en el ambiente al que pertenecen. Se puede apreciar que mediante distancias euclidianas de similitud se agrupan los metagenomas (Figura 9) y haciendo un análisis de dispersión de puntos (Figura 10) podemos ver que las funciones identificadas se mantienen muy cerca de la diagonal en la mayor parte de los casos, con ligeras desviaciones a

cada uno de los dos ambientes.

En las funciones que más diferencias podemos encontrar entre ambientes es en el metabolismo de carbohidratos y aminoácidos con sus derivados, esto demuestra que las rutas metabólicas principales se mantienen, como es el caso de la fijación de nitrógeno y la fotosíntesis (Figura 11), que si bien se mantiene en una proporción muy baja comparada a los demás genomas se encuentra conservado

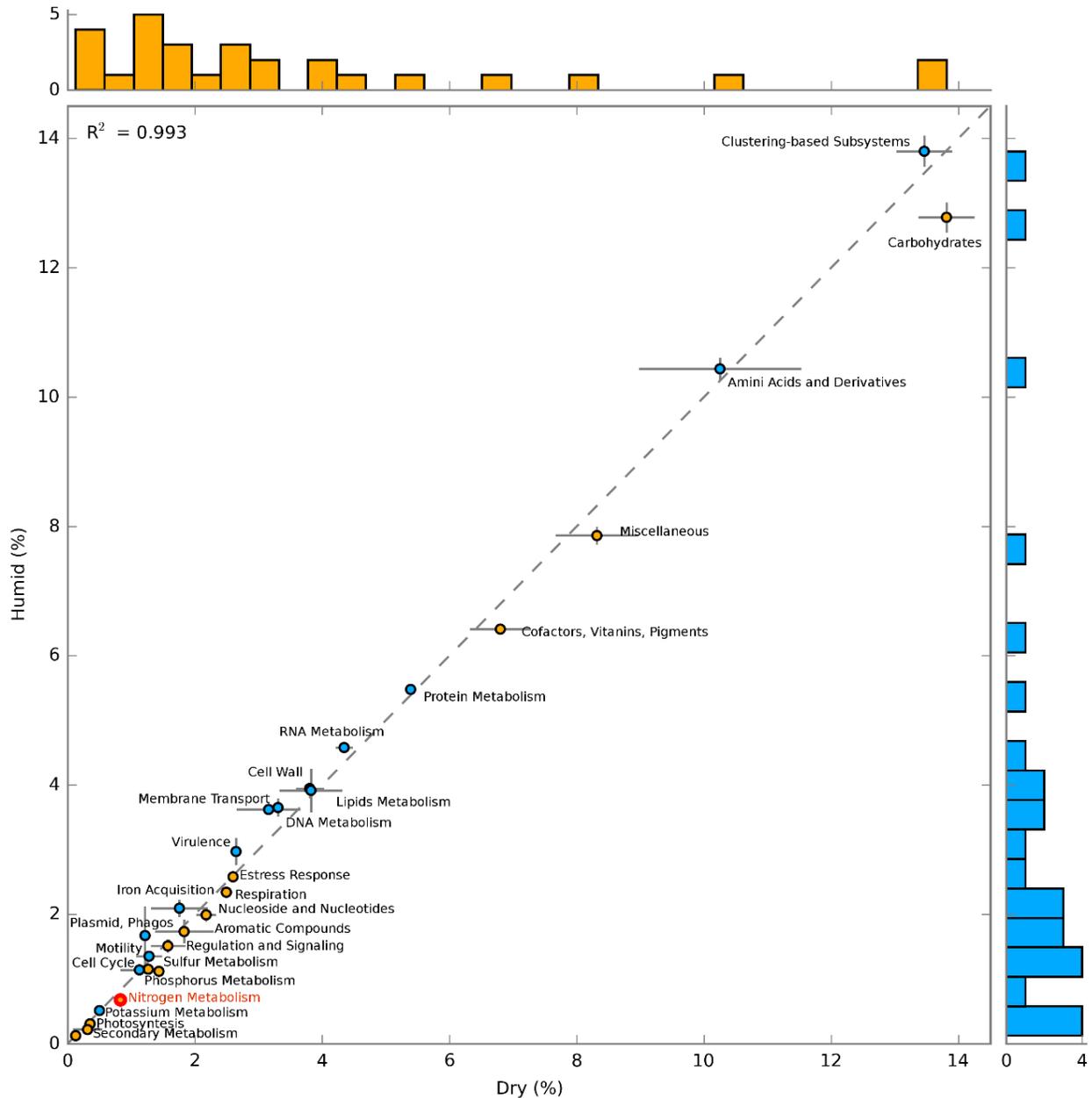


Figura 10. Grafica de dispersión, donde se compara la proporción de funciones anotadas en los metagenomas entre ambientes. Ambiente húmedo (Humid), Ambiente seco (Dry).

en todos. Se han analizado estas dos rutas metabólicas importantes y hemos encontrado que por homología de secuencias se puede deducir que son rutas pertenecientes al grupo de las nostocales, lo que apoya el resultado de que la cianobacteria es una de las bases para generar el nicho de la raíz coraloide.

También se encontró que hay diferencias en menor proporción en las funciones de los metagenomas de los distintos ambientes. El origen de estas diferencias puede tener distintas razones: i) El nivel de profundidad de los metagenomas afecta la cantidad de rutas metabólicas que se pueden identificar, ii) Las diferencias entre la composición de las especies de los metagenomas da plasticidad a tener distintas rutas metabólicas dentro del microbioma. Ambos puntos tienen gran validez, uno es un problema técnico y el otro una conjetura basada en la biología de la microcomunidad. Si hacemos de lado los problemas técnicos de la secuenciación y basados en la diversidad bacteriana observada se puede apoyar la idea de que diferentes taxa en el microbioma pueden aportar variabilidad a las funciones del microbioma.

Basados en la gran similitud que hay a nivel de función se puede apreciar que se mantiene un gran grupo de funciones metabólicas. Para probar que estas funciones forman un núcleo que es la base de la microcomunidad de la raíz coraloides se decidió buscar más que homologías funcionales, un análisis que busque a nivel de homología en secuencias para buscar genes ortólogos entre los metagenomas, para descartar que todas las funciones vienen de un mismo organismo.

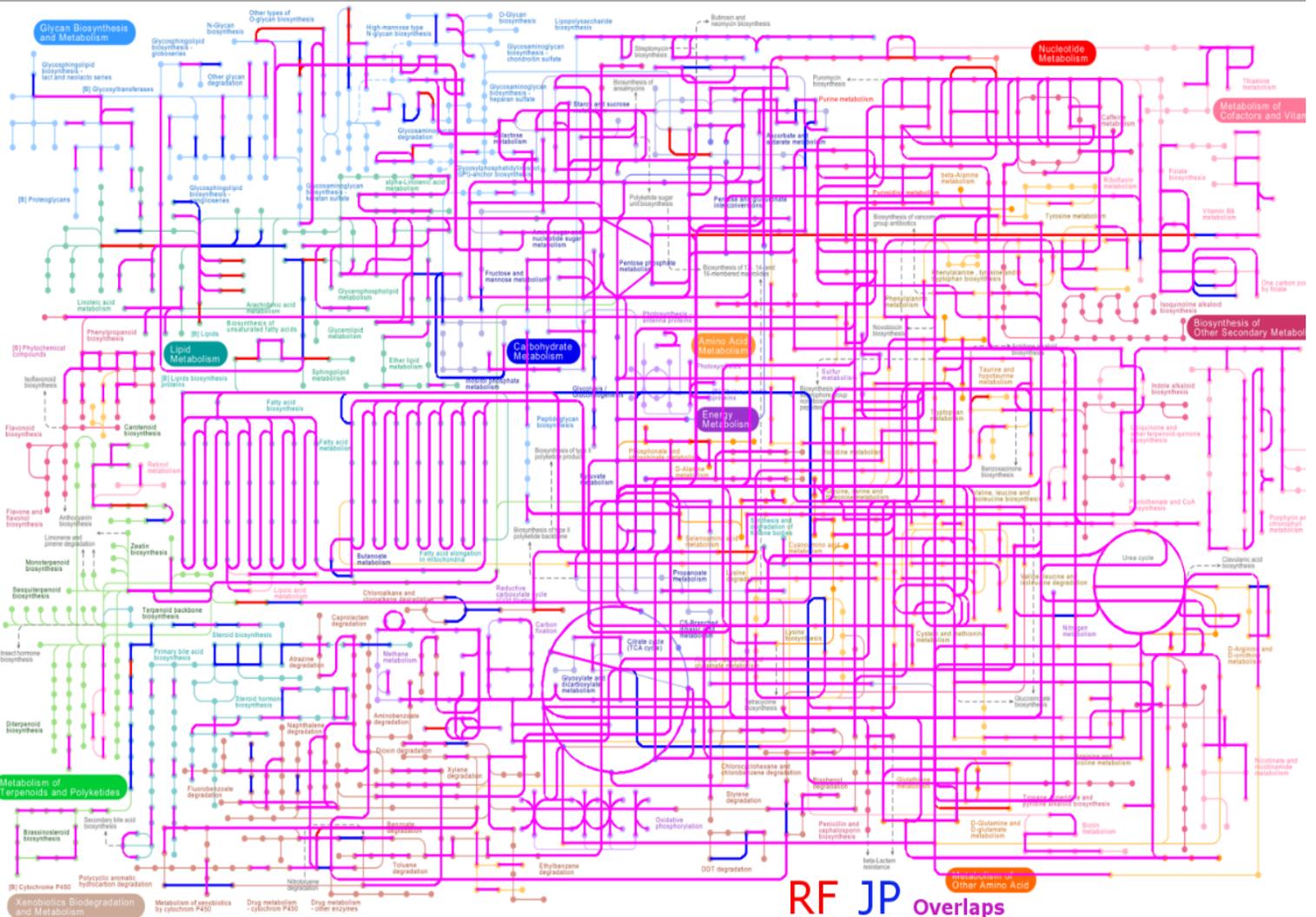


Figura 11. Mapa de rutas metabólicas Keeg, donde se comparan las principales rutas metabólicas de los metagenomas del ambiente húmedo con color rojo (RF), el ambiente seco en calor azul (JP) y las rutas metabólicas sobrelapadas en los dos ambientes en color morado.

Co-cultivos fotosintéticos a un año de ser sembrados

Comparación taxonómica de los co-cultivos con diferente tiempo de incubación

Debido a que los co-cultivos fotosintéticos se mantuvieron creciendo en reposo se forman biopelículas hechas de mucílago y una gran cantidad de biomasa respecto al líquido del matraz donde se encontraban. Por el fenómeno presentado surgió la siguiente pregunta: ¿es posible que estos cultivos conserven la comunidad bacteriana identificada con 11 meses de anterioridad? Para responder a esta pregunta se repitió la extracción y la secuenciación en dos de estos medios, encogiéndose uno de cada ambiente (JP6 y RF3) los cuales renombramos como RF3y1 y JP6y1.

Se encontró que RF3y1 y JP6y1 aumentaron la diversidad de géneros encontrados a través de un año llegando a un promedio de 69 géneros mientras que sus contrapartes tienen un promedio de 32 géneros identificados, lo que muestra que la cantidad de especies que podemos detectar se duplicó con el paso del tiempo, esto está ligado a que las especies de lento crecimiento fueron desarrollándose durante el año que llevo el experimento y al momento de la secuenciación se pudieron detectar. También observamos que, aunque la diversidad aumentó, muchos de los géneros se conservaron (en promedio 23 géneros se mantienen) (Figura 12).

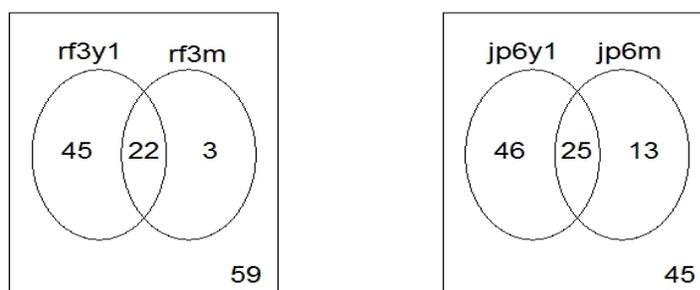


Figura 12. Diagrama de Venn en el cual se comparan los géneros encontrados en los co-cultivos de un mes comparados con el paso de un año.

Entre los géneros que se conservan entre los co-cultivos de un mes y un año destacan: *Stenotrophomonas*, *Pseudoxanthomonas*, *Xanthomonas*, *Microbacterium*, *Rhizobium*, *Methylobacterium*, *Agrobacterium*, *Caulobacter*, *Burkholderia* y *Nostoc*, estos géneros son los que están jugando un rol como base de la comunidad, mientras los géneros que se perdieron fueron: *Bacillus*, *Brucella*, *Methylibium*, *Roseobacter* y *Parvibaculum*. Los géneros perdidos en su mayoría son Proteobacterias, estas bacterias son de crecimiento rápido y necesitan medios ricos en nutrientes. Los géneros nuevos que se pudieron identificar en un año fueron: *Clavibacter*, *Streptomyces*, *Leifsonia*, *Rhodospirillum*, *Pseudomonas*, *Anabaena*, *Micrococcus* y *Mycobacterium*, estos géneros en su mayoría pertenecen al grupo de las Actinobacterias, estos organismos suelen tener tasas de duplicación muy lenta en la naturaleza y es por eso que en los cultivos de un mes no las detectamos, también resalta la presencia del género *Anabaena*, esta cianobacteria no la habíamos detectado antes y fue una de las primeras reportadas en la literatura como simbiote de cícadadas.

Comparación funcional de los co-cultivos con diferente tiempo de incubación

A nivel funcional se ven cambios en especial en JP6y1 ya que en un análisis de similitud se aleja de todos los demás metagenomas de medio fotosintético mientras que RF3y1 es más parecido funcionalmente a los co-cultivos de un mes de incubación (Figura 9). Comparando las funciones de los metagenomas de un año JP6y1 y RF3y1 con sus contrapartes de un mes, vemos que el metabolismo de carbohidratos se enriquece dentro de un año mientras que el metabolismo de adquisición de hierro y transportes de membrana se pierde un poco.

Como podemos ver en la figura (Figura 13) la mayor parte de las funciones se mantiene a través del tiempo del cultivo. Se esperaba un aumento en las funciones relacionadas con la obtención de nutrientes, como el metabolismo del nitrógeno y del potasio, pero estos se mantienen en la diagonal con cambios mínimos, lo que apunta a que las bacterias que tienen estas funciones son pocas,

pero se mantienen ya que son vitales para el funcionamiento de toda la comunidad.

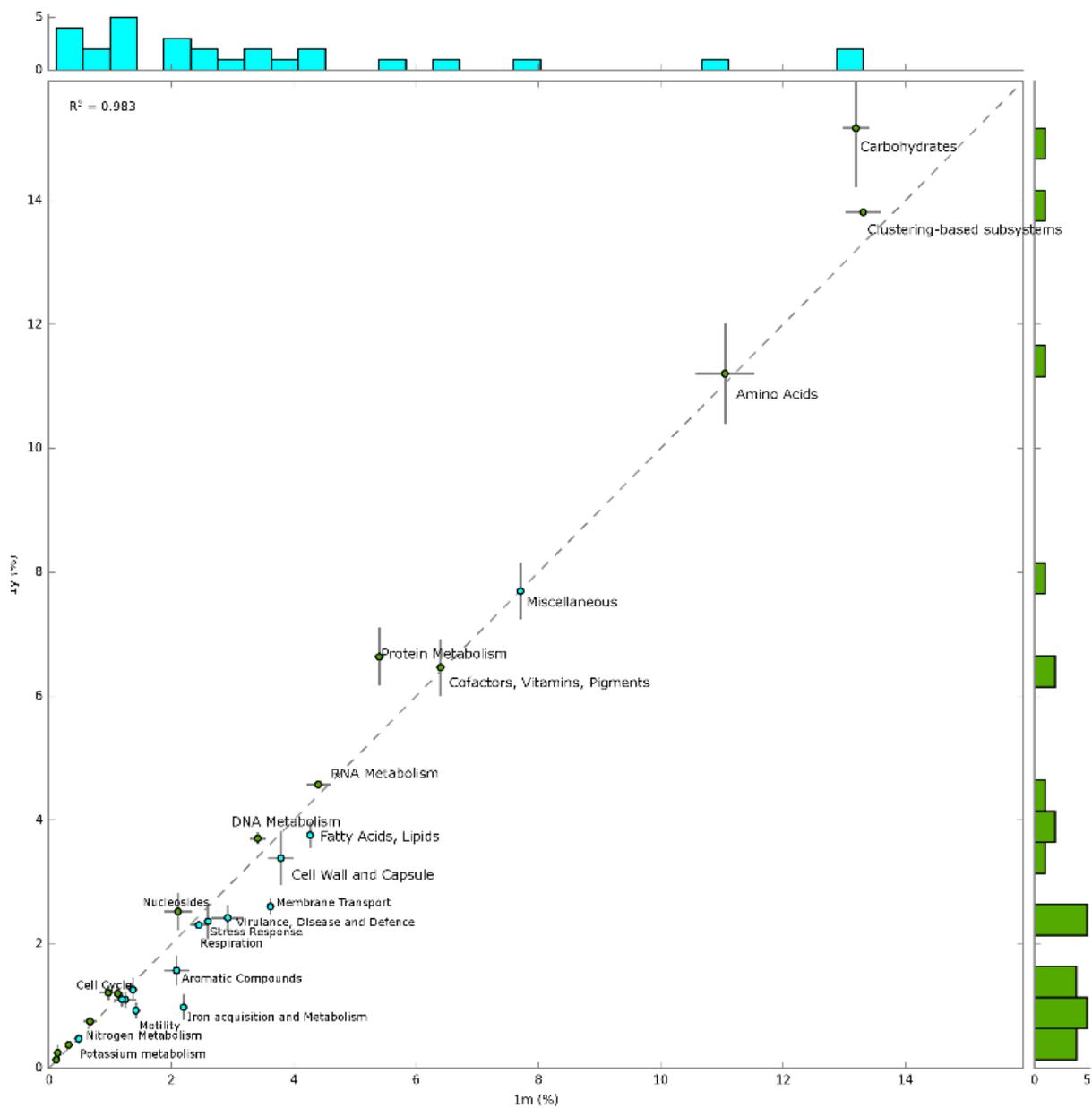


Figura 13. Grafica de dispersión de las funciones de los metagenomas de 1 año comparado contra la abundancia que estos tenían en el primer mes de incubación.

Es notorio que entre ambientes la mayor parte de las rutas metabólicas están compartidas (Figura 11). Pero a nivel taxonómico existen muchas diferencias ya que solo se comparte menos de la mitad de los OTUs entre ambientes. Con una comparación entre ambos ambientes a nivel taxonómico y a nivel funcional vemos las grandes diferencias que hay a nivel taxonómico (Figura 14), a nivel taxonómico solo pocos grupos se encuentran cerca de la diagonal mientras que en las funciones es evidente el caso opuesto.

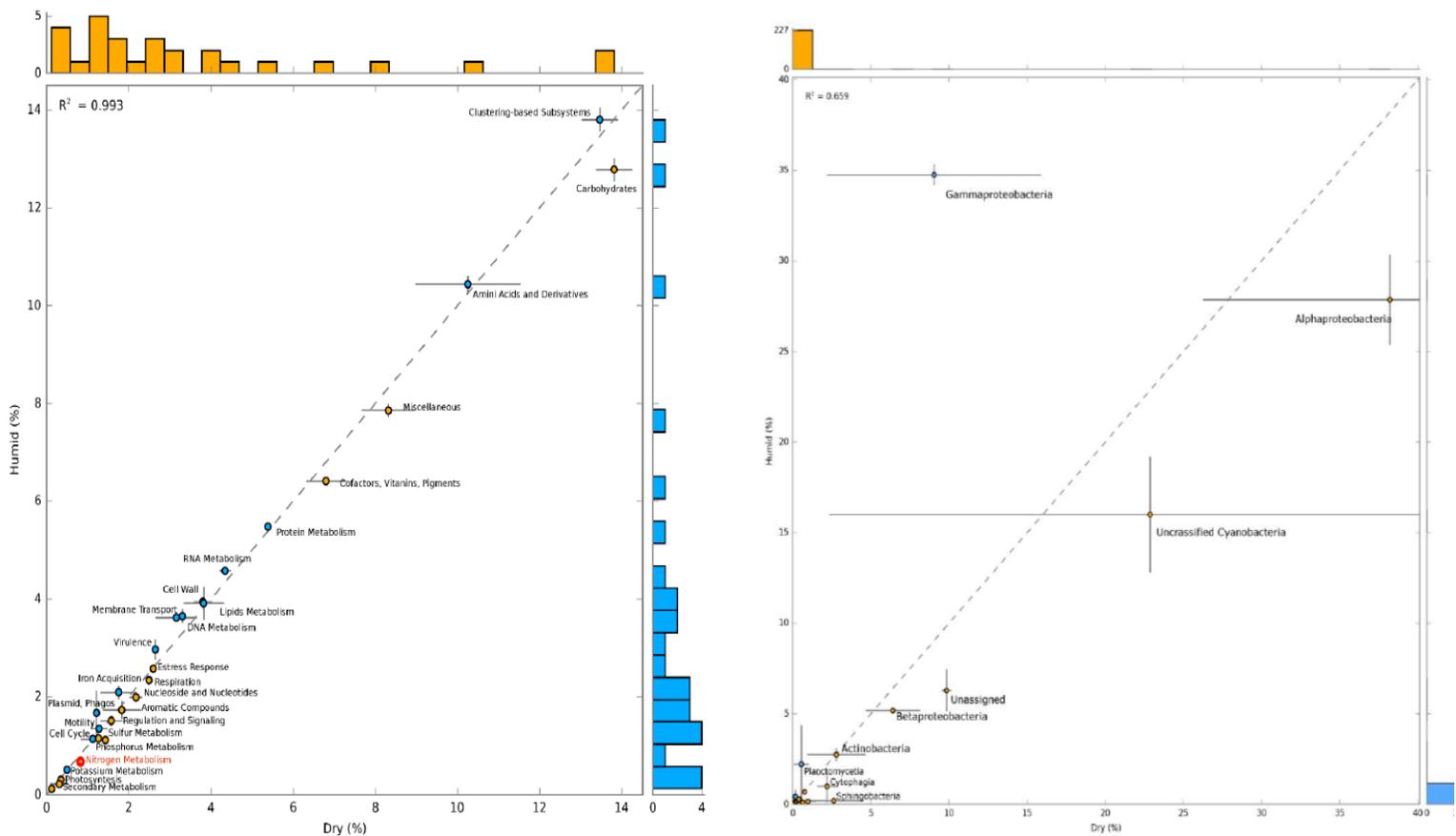


Figura 14. Graficas de dispersión, donde se compara la proporción de funciones anotadas en los metagenomas entre ambientes (izquierda) y a nivel taxonómico. Ambiente húmedo (Humid), Ambiente seco (Dry).

Para saber que factor afectaba más la composición taxonómica de los metagenomas se utilizó un análisis de componentes principales lo que demostró que hay una tendencia taxonómica y funcional al medio que se usa y el tiempo de incubación de este (Figura 15). Vemos que a nivel taxonómico hay una diferencia

muy marcada entre el medio eutrófico y los medios fotosintéticos por parte del componente principal 1 (PC1) que explica el 73% de la variación, a nivel funcional también tenemos esta separación, pero un poco menos marcada. Los medios fotosintéticos sin importar el tiempo de un mes o un año se mezclan a nivel taxonómico, pero a nivel funcional se pueden diferenciar un poco más, lo que nos indica que si hubo un cambio en el enriquecimiento de ciertas funciones al paso del tiempo.

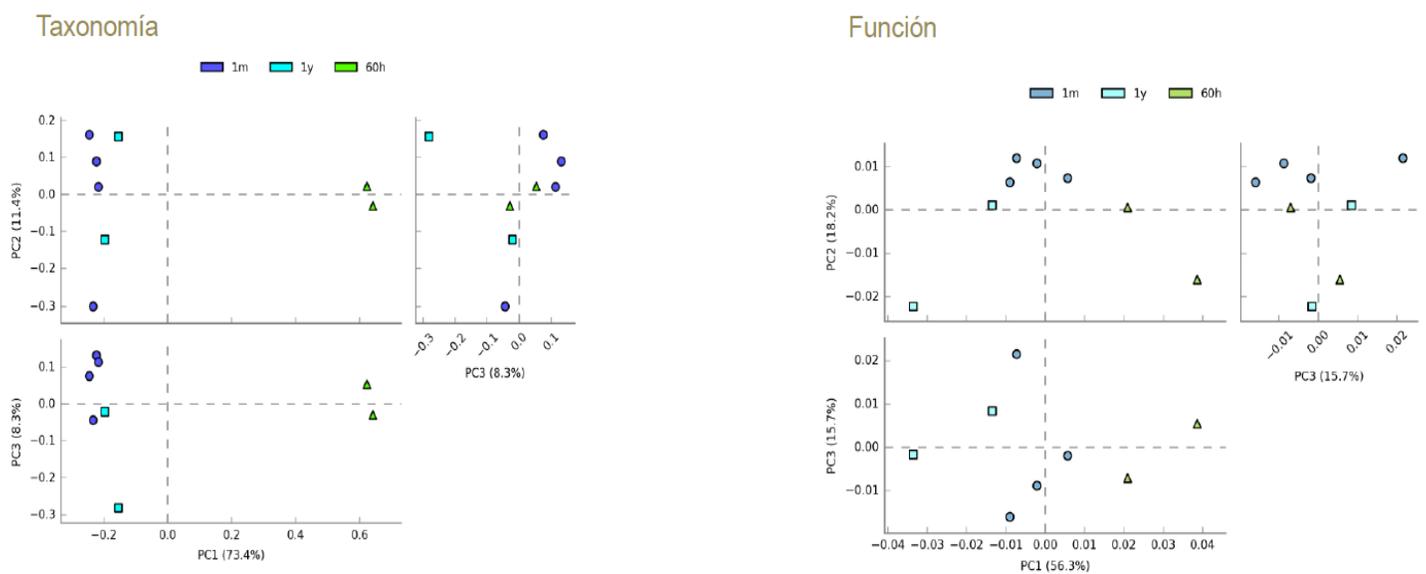


Figura 15 Análisis de componentes principales a nivel taxonómico y funcional de los metagenomas. Metagenomas pertenecientes a co-cultivos eutróficos (60 h), metagenomas fotosintéticos con un mes de incubación (1m) y metagenomas fotosintéticos con un año de incubación (1y)

DISCUSIÓN

Revelando el microbioma de la raíz coraloide: más allá de las cianobacterias

La primera evidencia de la presencia de un microbioma dentro de la raíz la encontramos en el trabajo de Lobakova y colaboradores en el 2003, donde mediante microscopia electrónica estudia la estructura de comunidades de cianobacterias durante la degradación de las raíces de las cícadas: *Cycas circinalis*, *Ceratozamia mexicana* y *Encephalartos hindelbrantii*. En este trabajo se encontró una gran cantidad de bacterias heterotróficas creciendo con las cianobacterias, pero no fueron identificadas a profundidad, ni se tomaron las condiciones experimentales necesarias para saber si el origen de estas bacterias es de la rizósfera o de la endosfera; tampoco se trató de incrementar la resolución de las interacciones bacterianas como se ha postulado (Cibrian-Jaramillo and Barona-Gómez 2015), ya que este no era su objetivo. Por otro lado, desde el trabajo de Zymmerman y Rosen en 1992, todo el esfuerzo de análisis más profundo y amplio en otros géneros, ha sido dirigido a aislar e identificar exclusivamente cianobacterias. En contraste con estos antecedentes, en este proyecto de tesis, se encontró un microbioma complejo taxonómicamente, donde se identificaron 128 géneros mediante metagenómica y 26 géneros por 16s rRNA que constituyen la mayor parte del microbioma de la raíz coraloide. Es concluyente que la cianobacteria no es el único simbiote de la raíz coraloide.

Aunque este trabajo no se basó en solo la identificación de cianobacterias se lograron identificar dos cianobiontes. Siendo de los pocos identificados para la especie *D. merolae*, las cianobacterias identificadas pertenecen al grupo de las Nostocales, y a los géneros *Nostoc* y *Tolypothrix*. Representan una nueva adición al estudio de cianobiontes y de los primeros para cicadas Mexicanas. Sorprendentemente, las Nostocales son la minoría de bacterias en cuanto a diversidad filogenética y en abundancia, lo que sugiere que los trabajos previos han sido sesgados por el tipo de mediciones (marcadores o técnicas específicas para cianobacterias) y por el tipo de muestreo (jardines botánicos).

La integración de colectas directas en campo, múltiples maneras de detección, y análisis filogenéticos, permitieron revelar la complejidad del microbioma de la raíz coraloide. Dando como resultado la identidad molecular de los endófitos diferentes a las cianobacterias de la raíz coraloide de *D. merolae*, comparando su 16S rRNA con el de una base de datos formada por exclusivamente secuencias de 16S rRNA de organismos reportados como endófitos, tomada del trabajo de Hardoim y colaboradores (2015). Vemos que el total de nuestras secuencias tienen más de un 97% de homología con secuencias depositadas en esta base de datos, por lo que no encontramos ninguna bacteria que ya no estuviera reportada como endófito en otra especie vegetal (Hardoim *et al.* 2015). Adicional a los aislados se obtuvieron los perfiles taxonómicos de los co-cultivos lo que reveló aún más diversidad y funciones de la microcomunidad endófito.

Diversidad entre medios de cultivo

La diferencia entre las bacterias que se pueden recuperar entre *t0* y *t1*, es muy amplia. De los géneros identificados en *t0* y *t1* sólo se solapan en el 20,8% de los taxa. Este resultado era predecible ya que la estrategia usada para *t0* (aislado directo de la muestra) y *t1* (co-cultivos) son complementarios entre sí y su objetivo es recuperar aspectos distintos de la composición taxonómica de la microbiota. Nuestros resultados también muestran que no hubo enriquecimiento para un grupo filogenético particular por parte de los seis medios oligotróficos utilizados para el aislamiento. Muchos de los géneros fueron aislados en varios medios de cultivo, sin predilección específica por alguno, siendo el más diverso en géneros bacterianos el medio 4 (Nutrient Broth) (Figura 5).

El patrón más notorio entre *t0* y *t1* fue que la mayoría de los géneros con sólo unas pocas especies se recuperaron en *t1*, y géneros con muchas especies se recuperaron en ambos, *t0* y *t1*. En *t0* se encuentran cuatro géneros abundantes en OTUs (*Streptomyces*, *Bacillus*, *Pseudomonas* y *Burkholderia*); y cuatro géneros pobres en OTUs (*Paenibacillus*, *Enterobacter*, *Pantoea*, *Ralstonia*). En co-cultivos (*t1*), se encuentran tres géneros abundantes en OTUs (*Rhizobium*,

Stenotrophomonas y *Achromobacter*), y 13 géneros pobres en OTUs fueron recuperados (Figura 5).

Estos patrones entre las diferencias en $t0$ y $t1$, pueden ser explicadas por varios escenarios: i) taxa raros (baja abundancia) sólo pueden ser recuperados en co-cultivos por el enriquecimiento en biomasa; ii) algunos géneros son de crecimiento rápido independientemente del medio de cultivo, y serán dominante su número en uno u otro método, simplemente por casualidad, iii) algunos grupos son más específicos a un medio en particular, y crecen mejor en BG11 ($t1$) que en los cultivos individuales ($t0$), o viceversa; y, iv) grupos en BG11 ($t1$) se recuperan como resultado de las interacciones funcionales entre los grupos de la sub-comunidad del co-cultivo.

Distribución de los géneros bacterianos de acuerdo al medio ambiente

Hay poca similitud en la diversidad taxonómica entre los ambientes a pesar de su alta diversidad, es notoria la diferencia taxonómica; al nivel de 16s rRNA y mOTUs, vemos que solo se comparten 43.4% y 42.3% géneros respectivamente (Figura 8). La distribución divergente de los grupos bacterianos en los ambientes contrastantes podría explicarse por un sesgo en el muestreo ambiental, es decir, en un ambiente diverso es más fácil reclutar a mas taxa bacterianos. Alternativamente, podría explicarse por la adaptación local del propio microbioma a diferentes ambientes, en la que hay filtrado de taxa por la planta sobre la base de las capacidades funcionales de las bacterias (Delaux *et al.* 2014; Bell *et al.* 2015; Chiu and Gilbert 2015).

De los organismos que encontramos en ambos ambientes destacan los géneros *Burkholderia*, *Bacillus*, *Achromobacter*, *Pseudomonas* y *Mitsuaria* que se aislaron en ambos entornos, adicional a estos grupos se identificaron en los metagenomas *Bradyrhizobium*, *Agrobacterium*, *Mesorhizobium*, *Sphingobium*, *Rhodopseudomonas*, *Sinorhizobium*, *Sphingomonas*, *Microbacterium* y

Caulobacter también en ambos ambientes. La presencia de grupos parcialmente en las 4 muestras ambientales, nos da un indicio de algunos endófitos que pudieran estar siempre presente dentro de la raíz coraloide aparte de la cianobacteria, algo inexplorado hasta este trabajo. *Bacillus*, *Achromobacter*, *Burkholderia*, *Agrobacterium*, *Sphingobium* y *Pseudomonas* tienen antecedentes de ser encontrados como endófitos comunes de otras plantas (Rosenblueth and Martínez-Romero 2006). *Mitsuaria* es un buen ejemplo de los géneros que encontramos en la raíz coraloide la cual es una de las bacterias asociadas a plantas que contribuyen con la función de la resistencia a los patógenos (Baysal *et al.* 2008; Rong *et al.* 2012)

Diversidad cultivos de 1 año

Entre los metagenomas de un año se vio un incremento en promedio del 100% de la diversidad de géneros con sus contrapartes de un mes (Figura 12). En ambos medios el cambio fue muy similar, donde las Proteobacterias decaen en abundancia y hay un aumento en bacterias gram positivas como las Streptomyces. Este fenómeno puede estar ligado a tres factores: i) El cambio en la disponibilidad de nutrientes hace que los taxa adaptados a estas condiciones se enriquezcan mientras que los que no decaigan, ii) El tiempo de incubación tan prolongado permite que los grupos de crecimiento muy lento tengan el tiempo suficiente para desarrollarse, y; iii) el metabolismo de las cianobacterias y sus exudados benefician o perjudica a ciertos grupos de la comunidad dependiendo del ambiente. No se puede saber cuál de estos factores es el determinante por la resolución de nuestro experimento, pero genera preguntas nuevas acerca del rol de las cianobacterias como la base de una microcomunidad.

Funciones del microbioma de la raíz coraloide: fijación de nitrógeno

De manera general las funciones metabólicas entre medios, estrategias de cultivo (1 año vs 1 mes) y ambientes, están conservadas en el microbioma de la raíz coraloide. Si hay diferencias entre las funciones, pero son cuestiones muy puntuales (Figura 11), en las cuales falta profundizar para poder concluir que apoyan a la adaptación local (Parks and Beiko 2010b). Incluso profundizando en

las funciones es algo difícil de probar experimentalmente hoy en día. Se sabe que muchos de los endófitos no tienen un efecto aparente en la planta hospedera y simplemente se desarrollan en el interior de los tejidos sin función aparente (Elmerich 2007), a estos se les ha llamado endófitos comensales, Otros endófitos confieren efectos benéficos para su planta hospedera como: defensa contra patógenos, adquisición de nutrientes como nitrógeno y hierro (Meeks *et al.* 1985; Reis *et al.* 2004; Kneip *et al.* 2007b; Sessitsch *et al.* 2012; Mendes *et al.* 2014). Las funciones del microbioma suelen ser complejas dado que funcionan en dependencia del genotipo del hospedero y las condiciones ambientales (Sessitsch *et al.* 2012; Boon *et al.* 2014; Hardoim *et al.* 2015), y es común que los endófitos pasen de comensales a patógenos o viceversa dependiendo del estrés que esté sufriendo el hospedero.

Las diferencias que hay a nivel taxonómico y a nivel funcional entre el microbioma de la raíz coraloide nos indican una tendencia a la selección del microbioma por función y no por taxonomía. Esto se puede traducir en que la planta recluta los organismos del ambiente por su función y la raíz funciona como un nicho específico, por lo que las bacterias que se encuentran dentro de él necesitan capacidades metabólicas específicas para sobrevivir en la endosfera de la raíz coraloide (Ahern and Staff 1994b). También podría ser importante e interesante la complementación de rutas metabólicas que hay entre las diferentes comunidades bacterianas, pero es algo que no se abordó en este proyecto, pero se especula y discute en la siguiente sección (microbioma de la raíz).

Dados los antecedentes en la literatura, y el hecho de que la mayoría de las bacterias encontradas en este trabajo son capaces de fijar nitrógeno, concentraré la discusión en esta función. El nitrógeno es un elemento muy abundante en la atmósfera, pero la habilidad de fijar nitrógeno atmosférico y convertirlo en compuestos aprovechables por la vía del complejo de enzimático de la nitrogenasa está restringida a algunos grupos de bacterias en los cuales se encuentran algunos géneros de cianobacterias. Estas bacterias se denominan diazotrofos, ya que obtienen todo su nitrógeno mediante la fijación de nitrógeno molecular (Berman-Frank, Lundgren and Falkowski 2003b). Por lo que los

organismos eucariotas como las cícadas y los demás organismos sólo son capaces de obtener nitrógeno fijado a través de sus interacciones simbióticas con bacterias fijadoras de nitrógeno (Kneip *et al.* 2007a). Las asociaciones de bacterias fijadoras de nitrógeno y especies hospederas ocurren en casi todos los ambientes y medios, acuático y terrestre, y presentan una gran diversidad de relaciones y estrategias (Kneip *et al.* 2007a).

En el microbioma de la raíz coraloide se identificaron una mayoría de grupos diversos grupos de bacterias identificadas como diazótrofes (Tabla 1 y 2 anexos bacterias fijadoras), algo que era de esperarse en una raíz simbiótica (Kistner and Parniske 2002). Sin embargo, cuando se minaron los metagenomas solo se encontraron los genes de la nitrogenasa codificados para cianobacterias (Berman-Frank, Lundgren and Falkowski 2003a). Para todos los diazótrofes la nitrogenasa tiene un complejo enzimático altamente conservado en distintos grupos de los reinos Bacteria y Archaea, los análisis filogenéticos sugieren que la información que codifica para este complejo tiene un ancestro en común (Berman-Frank, Lundgren and Falkowski 2003b). Pero a pesar de la diversidad de especies diazótrofes en los metagenomas solo se encontró la nitrogenasa ligada a las cianobacterias. Asimismo, y de manera notoria, una gran cantidad de genes del metabolismo del nitrógeno usados para la desnitrificación y producción de compuestos nitrogenados (~600) están codificados en los metagenomas de los demás géneros de la raíz coraloide. En conjunto, esto nos indica que la fijación de nitrógeno en esta microcomunidad podría ser una función específica de las cianobacterias, y los otros taxa ayudan con la fijación de nitrógeno de manera indirecta. Ya se ha visto que hay bacterias que se adhieren a la superficie a los heterocistos (Diferenciación celular de las cianobacterias para la fijación de nitrógeno) y promueven intercambio gaseoso y de metabolitos que ayudan a la cianobacteria con la fijación (Bershova, Kopteva and Tantsyurenko 1968).

Microbioma de la raíz coraloide: una comunidad bacteriana asociada a las cianobacterias.

En ecosistemas naturales se ha visto que grupos como las cianobacterias crecen junto a grupos definidos de bacterias que se asocian a la zona fibrilar y mucilago producidas por estas (Rasmussen and Nilsson 2002). Estas bacterias que se unen a las cianobacterias crecen en un microambiente que es producido por el metabolismo y los exudados de las cianobacterias. Este microambiente es un nicho complejo donde se crean gradientes químicos e intercambios gaseosos, producidos por el metabolismos como la fotosíntesis, la fijación de nitrógeno y la respiración de las bacterias epifitas que rodean a la cianobacteria (Meeks *et al.* 1985; Herrero *et al.* 2001; Berman-Frank, Lundgren and Falkowski 2003b; Jungblut *et al.* 2005). En numerosos grupos de cianobacterias como *Anabaena* y *Nostoc* se ha visto que este microambiente depende de la superficie donde se encuentren, por ejemplo, cuando están en simbiosis con plantas, y son estimuladas para la fijación de nitrógeno cambian su microambiente drásticamente. Es por eso que en el co-cultivo encontramos una gran diversidad que cambia con el tiempo (Figura 12), ya que está diversidad de bacterias se puede asociar a la cianobacteria. Creemos por nuestros resultados que el mismo fenómeno de asociación pasa en la raíz coraloide donde la cianobacteria entra a la raíz coraloide con sus bacterias asociadas y posteriormente por e microambiente de la raíz coraloide se selecciona una sub-comunidad que se relaciona con el metabolismo de nitrógeno y la vida endófito.

Pese a que las cianobacterias se encuentran en bajas densidades y tienen poca diversidad taxonómica, es claro que existe un microbioma cuya función, posiblemente fijar nitrógeno, es más importante que la composición taxonómica. En este marco, podemos especular en el papel de las cianobacterias en la raíz coraloide, y del microbioma asociado a ellas, en particular con algunos grupos taxonómicos específicos como *Pseudomonas* y *Caulobacter*, para los cuales existe evidencia previa de sus interacciones en comunidades o asociaciones con

otras bacterias (Bershova, Kopteva and Tantsyurenko 1968). En otros trabajos posteriores también se identificaron bacterias (*Pseudomonas* y *Caulobacter*) virtualmente siempre presentes en el mucilago de las cianobacterias (Kuentzel 1969; Caldwell and Caldwell 1978). Se identificó que dependiendo del ambiente (acuático, terrestres) pueden variar el tipo de bacterias asociadas a las cianobacterias.

Comparando con nuestros resultados, las cianobacterias de la raíz coraloide están asociadas en su mayoría a Proteobacterias, recreando un ambiente similar a ambientes acuáticos y de simbiosis con plantas. Con el paso del tiempo en los co-cultivos de un año donde la cianobacteria transforma el medio de cultivo, vemos que las Proteobacterias decaen en abundancia y hay un aumento en bacterias gram positivas como se ve en ambientes terrestres o rocosos (Albertano and Urzi 1999). En particular, en los co-cultivos de 1 año no hay acceso de nutrientes ni gases, pero las bacterias asociadas a las cianobacterias se encargan del reciclado de los nutrientes lo cual aprovechan para su desarrollo y crecimiento, regresando CO₂ al medio con la descomposición de los compuestos orgánicos y N₂ mediante procesos de desnitrificación, lo que crea un microcosmos estable al cambio del tiempo y el sistema se vuelve dinámico a nivel taxonómico dependiendo de los requerimientos metabólicos de la cianobacteria.

Esto nos da la primicia que el microbioma de la raíz coraloide está siendo influenciado por el microambiente creado por la cianobacteria. Como resultado, tenemos una comunidad relativamente estable de bacterias que sobreviven al medio transformado por la cianobacteria. Esto podría extrapolarse al ambiente que ocurre dentro de la raíz coraloide, una vez establecida la comunidad cianobacteria. Esta asociación alrededor de las cianobacterias, resultaría en un núcleo funcional, que no cambia drásticamente entre ambientes, y que no depende de la taxonomía de los miembros de la comunidad, sino de su capacidad de interactuar con las cianobacterias.

Por ello, proponemos que las cianobacterias, son la base del 'nicho

ecológico' de la raíz coraloide, y este nicho al ser reclutado por las cícadas puede generar ventajas al tomar no solo el metabolismo de la cianobacteria, si no todo el repertorio metabólico de la microcomunidad que las rodea, haciendo más eficientes los procesos fisiológicos como la fijación de nitrógeno, que, además, se ha visto que puede ser promovida en cianobacterias por grupos bacterianos que se unen a los heterocistos.

CONCLUSION

En las raíces coraloideas el núcleo funcional es notoriamente más estable y constante que el núcleo taxonómico. En los resultados solo el 42% de los géneros identificados se comparten entre los diferentes ambientes, mientras que la mayor parte de las rutas metabólicas se mantienen. La raíz forma un nicho específico en su endosfera para el desarrollo de las cianobacterias, el cual alberga a la cianobacteria y sus bacterias asociadas. Por lo tanto, creemos que se seleccionan bacterias que puedan crecer y desarrollarse dentro de la microcomunidad, lo cual está más ligado a sus funciones metabólicas que a su taxonomía. Este núcleo funcional está relacionado fuertemente con el microambiente que generan las cianobacterias, cambiando la distribución taxonómica de la microcomunidad, dependiendo de las condiciones en las que se encuentre la cianobacteria en su medio. Esto fue claro en los co-cultivos de un año. Entonces, las cianobacterias y sus bacterias asociadas se pueden conceptualizar como un holobionte (endófitos) dentro de otro (raíz coraloide).

PERSPECTIVAS

- Continuar con el trabajo de minería genómica para obtener las funciones y rutas metabólicas que sean de importancia para la microcomunidad de la raíz coraloide; Como, por ejemplo, las rutas metabólicas involucradas en el metabolismo del nitrógeno, las cuales pueden ser comprobadas con métodos experimentales (reducción de acetileno), y ver su influencia en la selección de la microcomunidad que crece junto a las cianobacterias de la raíz coraloide.
- Evaluar los efectos de compartimentación y la combinación de los genes en los metagenomas a nivel filogenético mediante la técnica llamada RADICAL (*Random Addition Concatenation Analysis*) (Narechania *et al.* 2012). El resultado que se espera es una biblioteca de árboles que representan una gran variedad de particiones de genes al azar de diferentes tamaños. En esta biblioteca se pueden identificar topologías únicas, valorar la topología general, y medir el apoyo a diferentes árboles. Esto podría dar una idea sobre los grupos de genes que hallan evolucionado para adaptarse a la simbiosis dentro de la raíz coraloide.
- Estudiar la microcomunidad asociada a las cianobacterias que colonizan plantas distintas a las cícadas, como el caso del helecho acuático *Azolla* con *Anabaena* (Van Hove and Lejeune 2002) y la angiosperma *Gunnera* con la cianobacteria *Nostoc* (Rasmussen, Bergman and Johansson 1994), con el fin de comparar diferencias y similitudes entre microbiomas.
- Realizar experimentos de cultivo de cícadas *in-vitro* para medir su adecuación con y sin microbioma. También es necesario un estudio del microbioma de la raíz coraloide en otras especies de *Dioon* para entender las diferencias que hay entre especies a nivel de composición bacteriana en la raíz coraloide.

BIBLIOGRAFÍA

- Adams DG, Duggan PS. Signalling in cyanobacteria--plant symbioses. *Signaling and Communication in Plant Symbiosis*. Springer, 2012, 93–121.
- Ahern CP, Staff IA. Symbiosis in cycads: The origin and development of coralloid roots in *Macrozamia communis* (Cycadaceae). *Am J Bot* 1994a:1559–70.
- Ahern CP, Staff IA. Symbiosis in cycads: The origin and development of coralloid roots in *Macrozamia communis* (Cycadaceae). *Am J Bot* 1994b;**81**:1559–70.
- Albertano, Urzi. Structural Interactions among Epilithic Cyanobacteria and Heterotrophic Microorganisms in Roman Hypogea. *Microb Ecol* 1999;**38**:201–14.
- Altschul SF, Gish W, Miller W *et al*. Basic local alignment search tool. *J Mol Biol* 1990;**215**:403–10.
- Bateman A, Birney E, Cerruti L *et al*. The Pfam Protein Families Database. *Nucleic Acids Res* 2002;**30**:276–80.
- Baysal F, Benitez M-S, Kleinhenz MD *et al*. Field management effects on damping-off and early season vigor of crops in a transitional organic cropping system. *Phytopathology* 2008;**98**:562–70.
- Bell CW, Asao S, Calderon F *et al*. Plant nitrogen uptake drives rhizosphere bacterial community assembly during plant growth. *Soil Biol Biochem* 2015;**85**:170–82.
- Bergman B, Rai AN, Rasmussen U. Cyanobacterial associations. *Associative and Endophytic Nitrogen-Fixing Bacteria and Cyanobacterial Associations*. Springer, 2007, 257–301.
- Berman-Frank I, Lundgren P, Falkowski P. Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. *Res Microbiol* 2003a;**154**:157–64.
- Berman-Frank I, Lundgren P, Falkowski P. Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. *Res Microbiol* 2003b;**154**:157–64.

- Bershova OI, Kopteva ZP, Tantsyurenko E V. The interrelations between the blue-green algae-the causative agents of “water bloom”--and bacteria. *Tsvetnie Vody* 1968:159–71.
- Bogler DJ, Francisco-Ortega J. Molecular systematic studies in cycads: Evidence from trnL Intron and ITS2 rDNA sequences. *Bot Rev* 2004;**70**:260–73.
- Boon E, Meehan CJ, Whidden C *et al.* Interactions in the microbiome: Communities of organisms and communities of genes. *FEMS Microbiol Rev* 2014;**38**:90–118.
- Buhler JD, Lancaster JM, Jacob AC *et al.* Mercury BLASTN: Faster DNA sequence comparison using a streaming hardware architecture. *Proc Reconfigurable Syst Summer Inst* 2007.
- Burleigh JG, Barbazuk WB, Davis JM *et al.* Exploring diversification and genome size evolution in extant gymnosperms through phylogenetic synthesis. *J Bot* 2012;**2012**.
- Cabrera-Toledo D, González-Astorga J, Nicolalde-Morejón F *et al.* Allozyme diversity levels in two congeneric *Dioon* spp.(Zamiaceae, Cycadales) with contrasting rarities. *Plant Syst Evol* 2010;**290**:115–25.
- Caldwell DE, Caldwell SJ. A Zoogloea sp. associated with blooms of *Anabaena flos-aquae*. *Can J Microbiol* 1978;**24**:922–31.
- Campbell EL, Meeks JC. Characteristics of hormogonia formation by symbiotic *Nostoc* spp. in response to the presence of *Anthoceros punctatus* or its extracellular products. *Appl Environ Microbiol* 1989;**55**:125–31.
- Campbell EL, Wong FCY, Meeks JC. DNA binding properties of the HrmR protein of *Nostoc punctiforme* responsible for transcriptional regulation of genes involved in the differentiation of hormogonia. *Mol Microbiol* 2003;**47**:573–82.
- Chiu L, Gilbert SF. The Birth of the Holobiont: Multi-species Birthing Through Mutual Scaffolding and Niche Construction. *Biosemiotics* 2015, DOI: 10.1007/s12304-015-9232-5.

- Cibrian-Jaramillo A, Barona-Gómez F. Increasing metagenomic resolution of microbiome interactions through functional phylogenomics and nearly-natural populations. *Front Microbiol* 2015;1–10.
- Cohen MF, Meeks JC. A hormogonium regulating locus, *hrmUA*, of the cyanobacterium *Nostoc punctiforme* strain ATCC 29133 and its response to an extract of a symbiotic plant partner *Anthoceros punctatus*. *Mol Plant Microbe Interact* 1997;**10**:280–9.
- Costa J-L, Lindblad P. Cyanobacteria in symbiosis with cycads. *Cyanobacteria in Symbiosis*. Springer, 2002a, 195–205.
- Costa J-L, Lindblad P. Cyanobacteria in Symbiosis with Cycads. *Cyanobacteria in Symbiosis*. Dordrecht: Kluwer Academic Publishers, 2002b, 195–205.
- Costa J-L, Paulsrud P, Lindblad P. Cyanobiont diversity within coralloid roots of selected cycad species. *FEMS Microbiol Ecol* 1999;**28**:85–91.
- Costa JL, Romero EM, Lindblad P. Sequence based data supports a single *Nostoc* strain in individual coralloid roots of cycads. *FEMS Microbiol Ecol* 2004;**49**:481–7.
- Cuddy WS, Neilan B a., Gehringer MM. Comparative analysis of cyanobacteria in the rhizosphere and as endosymbionts of cycads in drought-affected soils. *FEMS Microbiol Ecol* 2012;**80**:204–15.
- Cuddy WS, Neilan BA, Gehringer MM *et al.* Comparative analysis of cyanobacteria in the rhizosphere and as endosymbionts of cycads in drought-affected soils. *FEMS Microbiol Ecol* 2012;**80**:204–15.
- Delaux PM, Varala K, Edger PP *et al.* Comparative Phylogenomics Uncovers the Impact of Symbiotic Associations on Host Genome Evolution. *PLoS Genet* 2014;**10**, DOI: 10.1371/journal.pgen.1004487.
- Döbereiner J. Recent changes in concepts of plant bacteria interactions: endophytic N₂ fixing bacteria. *Cienc Cult* 1992;**44**:310–3.
- Dolezel J, Bartos J, Voglmayr H *et al.* Nuclear DNA content and genome size of

- trout and human. *Cytom Part A J Int Soc Anal Cytol* 2003;**51**:127.
- Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;**32**:1792–7.
- Elmerich C, Newton WE. *Associative and Endophytic Nitrogen-Fixing Bacteria and Cyanobacterial Associations*. Springer, 2007.
- Elmerich C. Historical perspective: from bacterization to endophytes. *Associative and Endophytic Nitrogen-Fixing Bacteria and Cyanobacterial Associations*. Springer, 2007, 1–20.
- Fisher JB, Vovides AP. Mycorrhizae are present in cycad roots. *Bot Rev* 2004;**70**:16–23.
- Futuyma DJ. *Science on Trial: The Case for Evolution*. Pantheon Books, 1983.
- Gehring MM, Pengelly J, Cuddy WS *et al*. Host selection of symbiotic cyanobacteria in 31 species of the Australian cycad genus: *Macrozamia* (Zamiaceae). *Mol Plant Microbe Interact* 2010;**23**:811–22.
- González D, Vovides AP, Bárcenas C. Phylogenetic relationships of the Neotropical genus *Dioon* (Cycadales, Zamiaceae) based on nuclear and chloroplast DNA sequence data. *Syst Bot* 2008;**33**:229–36.
- Grobbelaar N, Scott WE, Hattingh W *et al*. identification of the coralloid root endophytes of the southern African cycads and the ability of the isolates to fix dinitrogen. *South African J Bot Off J South African Assoc Bot Suid-Afrikaanse Tydskrif vir Plantkd amptelike Tydskrif van die Suid-Afrikaanse Genootskap van Plantkd* 1987.
- Hallmann J, Quadts-Hallmann A, Mahaffee WF *et al*. Bacterial endophytes in agricultural crops. *Can J Microbiol* 1997;**43**:895–914.
- Hallmann J, Quadts-Hallmann A, Mahaffee WF *et al*. Bacterial endophytes in agricultural crops. <http://dx.doi.org/10.1139/m97-131> 2011.
- Hardoim PR, van Overbeek LS, Berg G *et al*. The Hidden World within Plants: Ecological and Evolutionary Considerations for Defining Functioning of

- Microbial Endophytes. *Microbiol Mol Biol Rev* 2015;**79**:293–320.
- Haslam E. Secondary metabolism—evolution and function: Products or processes? *Chemoecology* 1994;**5**:89–95.
- Herrero A, Muro-pastor AM, Flores E *et al*. Nitrogen Control in Cyanobacteria MINIREVIEW Nitrogen Control in Cyanobacteria. 2001;**183**:411–25.
- Hilerio SLC, Ávila VMC, Zapotitla ES *et al*. Morfog{é}nesis in vitro de diomonerolae de Luca, Sabato & V{á}zquez-Torres (zamiaceae, cycadales) a partir de megagametofitos y embriones cig{ó}ticos. *Interciencia* 2008;**33**:929–34.
- Van Hove C, Lejeune A. The Azolla Anabaena symbiosis. *ProcRoyal Irish Acad* 2002;**102B**:23–6.
- Huerta-Cepas J, Dopazo J, Gabaldón T. ETE: a python Environment for Tree Exploration. *BMC Bioinformatics* 2010;**11**:24.
- Hug LA, Beiko RG, Rowe AR *et al*. Comparative metagenomics of three Dehalococcoides-containing enrichment cultures: the role of the non-dechlorinating community. *BMC Genomics* 2012;**13**:327.
- Jungblut AD, Hawes I, Mountfort D *et al*. Diversity within cyanobacterial mat communities in variable salinity meltwater ponds of McMurdo Ice Shelf, Antarctica. *Environ Microbiol* 2005;**7**:519–29.
- Kanehisa M, Goto S, Sato Y *et al*. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 2011:gkr988.
- Kearse M, Moir R, Wilson A *et al*. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012;**28**:1647–9.
- Kistner C, Parniske M. Evolution of signal transduction in intracellular symbiosis. *Trends Plant Sci* 2002;**7**:511–8.
- Kneip C, Lockhart P, Voß C *et al*. Nitrogen fixation in eukaryotes--new models for symbiosis. *BMC Evol Biol* 2007a;**7**:1.
- Kneip C, Lockhart P, Voss C *et al*. Nitrogen fixation in eukaryotes--new models for

- symbiosis. *BMC Evol Biol* 2007b;**7**:55.
- Knight CD, Adams DG. A method for studying chemotaxis in nitrogen fixing cyanobacterium-plant symbioses. *Physiol Mol Plant Pathol* 1996;**49**:73–7.
- Kuentzel LE. Bacteria, carbon dioxide, and algal blooms. *J (Water Pollut Control Fed)* 1969:1737–47.
- Lane DJ. 16S/23S rRNA sequencing. *Nucleic acid Tech Bact Syst* 1991:125–75.
- Lázaro Zermeño JM, Espinosa G, Ochoa MM *et al.* *Demografía, Conservación Y Aprovechamiento de Productos Forestales No Maderablesel Caso de Dioon Merolae (Zamiaceae) En Chiapas, México.*, 2011.
- Lederberg J, McCray AT. 'Ome sweet 'omics - A genealogical treasury of words. *Scientist* 2001;**15**:8.
- Lindblad P, Bergman B, Hofsten A V *et al.* THE CYANOBACTERIUM-ZAMIA SYMBIOSIS: AN ULTRASTRUCTURAL STUDY. *New Phytol* 1985;**101**:707–16.
- Lindblad P. Nitrogen and carbon metabolism in coralloid roots of cycads. *Adv Cycad Res I Mem New York Bot Gard* 1990;**57**:104–13.
- Lindow SE, Brandl MT. Microbiology of the Phyllosphere. *Appl Environ Microbiol* 2003;**69**:1875–83.
- Livingstone D, Jaworski GHM. The viability of akinetes of blue-green algae recovered from the sediments of Rostherne Mere. *Br Phycol J* 1980;**15**:357–64.
- Lobakova ES, Dol'nikova GA, Korzhenevskaya TG. Cyanobacterial–Bacterial Complexes in Plant Synchronoses. *Microbiology* **70**:111–6.
- Lobakova ES, Orazova MK, Dobrovol'skaya TG. The structure of cyanobacterial communities formed during the degradation of apogeotropic roots of cycads. *Microbiology* 2003;**72**:634–7.
- Losos JB, Ricklefs RE. Adaptation and diversification on islands. *Nature* 2009;**457**:830–6.

- De Luca P, Sabato S, Torres MV. *Dioon merolae* (Zamiaceae), a new species from Mexico. *Brittonia* 1981;**33**:179–85.
- Margulis L, Sagan D. *Acquiring Genomes: A Theory of the Origins of Species*. Basic Books, 2008.
- McDonald IR, Upton M, Hall G *et al*. Molecular ecological analysis of methanogens and methanotrophs in blanket bog peat. *Microb Ecol* 1999;**38**:225–33.
- Meeks JC, Campbell EL, Summers ML *et al*. Cellular differentiation in the cyanobacterium *Nostoc punctiforme*. *Arch Microbiol* 2002;**178**:395–403.
- Meeks JC, Enderlin CS, Joseph CM *et al*. Fixation of [¹³N] N₂ and transfer of fixed nitrogen in the *Anthoceros*-*Nostoc* symbiotic association. *Planta* 1985;**164**:406–14.
- Mende DR, Sunagawa S, Zeller G *et al*. Accurate and universal delineation of prokaryotic species. *Nat Methods* 2013;**10**:881–4.
- Mendes LW, Kuramae EE, Navarrete A *et al*. Taxonomical and functional microbial community selection in soybean rhizosphere. *ISME J* 2014:1577–87.
- Mereschkowsky KS. Symbiogenesis and the Origin of Species. *TH Morgan* 1926.
- Morrone JJ. Biogeographic areas and transition zones of Latin America and the Caribbean islands based on panbiogeographic and cladistic analyses of the entomofauna. *Annu Rev Entomol* 2006;**51**:467–94.
- Nagalingum NS, Marshall CR, Quental TB *et al*. Recent synchronous radiation of a living fossil. *Science* (80-) 2011;**334**:796–9.
- Narechania A, Baker RH, Sit R *et al*. Random Addition Concatenation Analysis: a novel approach to the exploration of phylogenomic signal reveals strong agreement between core and shell genomic partitions in the cyanobacteria. *Genome Biol Evol* 2012;**4**:30–43.
- Nathanielsz CP, Staff IA. A mode of entry of blue-green algae into the apogeotropic roots of *Macrozamia communis*. *Am J Bot* 1975:232–5.
- Norstog KJ, Nicholls TJ. *The biology of the cycads--Ithaca*. 1997.

- Obukowicz M, Schaller M, Kennedy GS. ULTRASTRUCTURE AND PHENOLIC HISTOCHEMISTRY OF THE CYCAS REVOLUTA-ANABAENA SYMBIOSIS. *New Phytol* 1981;**87**:751–9.
- Oksanen J, Kindt R, Legendre P *et al.* The vegan package. *Community Ecol Packag* 2007;**10**.
- Osborne R, Calonje M a, Hill KD *et al.* The World List of Cycads. *Proc 8th Int Conf Cycad Biol (CYCAD 2008)*, 2012a;**106**:480–510.
- Osborne R, Calonje MA, Hill KD *et al.* The world list of cycads. *Mem NY Bot Gard* 2012b;**106**:480–508.
- Parks DH, Beiko RG. Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* 2010a;**26**:715–21.
- Parks DH, Beiko RG. Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* 2010b;**26**:715–21.
- Pla L. Biodiversidad: Inferencia basada en el índice de Shannon y la riqueza. *Interciencia* 2006;**31**:583–90.
- Poulin R. The evolution of life history strategies in parasitic animals. *Adv Parasitol (United Kingdom)* 1996.
- Pruesse E, Quast C, Knittel K *et al.* SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 2007;**35**:7188–96.
- Quast C, Pruesse E, Yilmaz P *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013;**41**:D590–6.
- Rai AN, Bergman B, Rasmussen U *et al.* *Cyanobacteria in Symbiosis*. Springer, 2002.
- Rai AN. *Handbook of Symbiotic Cyanobacteria*. CRC Press, Inc., 1990.
- Rasmussen U, Bergman B, Johansson C. Early Communication in the Gunnera-Nostoc Symbiosis Plant-Induced Cell Differentiation and Protein Synthesis in

- the Cyanobacterium.pdf. *Mol Plant-Microbe Interact* 1994;**7**:696–702.
- Rasmussen U, Nilsson M. Cyanobacterial diversity and specificity in plant symbioses. *Cyanobacteria in Symbiosis*. Springer, 2002, 313–28.
- Reis VM, Estrada-de los Santos P, Tenorio-Salgado S *et al*. Burkholderia tropica sp. nov., a novel nitrogen-fixing, plant-associated bacterium. *Int J Syst Evol Microbiol* 2004;**54**:2155–62.
- van Rhijn P, Vanderleyden J. The Rhizobium-plant symbiosis. *Microbiol Rev* 1995;**59**:124–42.
- Rieseberg LH, Willis JH. Plant speciation. *Science* 2007;**317**:910–4.
- Rippka R, Stanier RY, Deruelles J *et al*. Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Microbiology* 1979;**111**:1–61.
- Rong X, Gurel FB, Meulia T *et al*. Draft genome sequences of the biocontrol bacterium Mitsuraria sp. strain H24L5A. *J Bacteriol* 2012;**194**:734–5.
- Ronquist F, Teslenko M, van der Mark P *et al*. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 2012;**61**:539–42.
- Rosenblueth M, Martínez-Romero E. Bacterial endophytes and their interactions with hosts. *Mol Plant Microbe Interact* 2006;**19**:827–37.
- Roughgarden J. Competition and theory in community ecology. *Am Nat* 1983;**122**:583–601.
- Schenk A, Weingart H, Ullrich MS. Extraction of high-quality bacterial RNA from infected leaf tissue for bacterial in planta gene expression analysis by multiplexed fluorescent Northern hybridization. *Mol Plant Pathol* 2008;**9**:227–35.
- Scherling C, Ulrich K, Ewald D *et al*. A Metabolic Signature of the Beneficial Interaction of the Endophyte Paenibacillus sp. Isolate and In Vitro–Grown Poplar Plants Revealed by Metabolomics. <http://dx.doi.org/101094/MPMI-22-8-1032> 2009.

- Schilling CH, Letscher D, Palsson BØ. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *J Theor Biol* 2000;**203**:229–48.
- Schneider D, Wink M, Sporer F *et al.* Cycads: their evolution, toxins, herbivores and insect pollinators. *Naturwissenschaften* 2002;**89**:281–94.
- Sessitsch a., Hardoim P, Döring J *et al.* Functional Characteristics of an Endophyte Community Colonizing Rice Roots as Revealed by Metagenomic Analysis. *Mol Plant-Microbe Interact* 2012;**25**:28–36.
- Sonnhammer ELL, Eddy SR, Durbin R. Pfam: A Comprehensive Database of Protein Domain Families Based on Seed Alignments.
- Staley JT, Konopka A. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. 1985;**39**:321–46.
- Stanier RY, Deruelles J, Rippka R *et al.* Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Microbiology* 1979;**111**:1–61.
- Sunagawa S, Mende DR, Zeller G *et al.* Metagenomic species profiling using universal phylogenetic marker genes. *Nat Methods* 2013;**10**:1196–9.
- Team Rs. RStudio: integrated development for R. *RStudio, Inc, Boston, MA URL <http://www.RStudio.com/ide>* 2014.
- Thajuddin N, Muralitharan G, Sundaramoorthy M *et al.* Morphological and genetic diversity of symbiotic cyanobacteria from cycads. *J Basic Microbiol* 2010;**50**:254–65.
- Tilman D. Diversity by default. *Science (80-)* 1999;**283**:495–6.
- Vermeij GJ. The evolutionary interaction among species: selection, escalation, and coevolution. *Annu Rev Ecol Syst* 1994:219–36.
- Vovides AP, Iglesias CG. An integrated conservation strategy for the cycad *Dioon edule* Lindl. *Biodivers Conserv* 1994;**3**:137–41.
- Vovides AP, Perez-Farrera MA, González-Astorga J *et al.* An outline of our current knowledge on Mexican cycads (Zamiaceae, Cycadales). *Plant Biol*

2003;**4**:159–74.

Ward DM, Weller R, Bateson MM. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* 1990;**345**:63–5.

Weinbauer MG, Fritz I, Wenderoth DF *et al.* Simultaneous extraction from bacterioplankton of total RNA and DNA suitable for quantitative structure and function analyses. *Appl Environ Microbiol* 2002;**68**:1082–7.

Wilke A, Glass E, Bischof J *et al.* MG-RAST Technical report and manual for version 3.3. 6–rev. 2013.

Xu Z, Malmer D, Langille MGI *et al.* Which is more important for classifying microbial communities: who's there or what they can do? *ISME J* 2014;**8**:2357–9.

Yamada S, Ohkubo S, Miyashita H *et al.* Genetic diversity of symbiotic cyanobacteria in *Cycas revoluta* (Cycadaceae). *FEMS Microbiol Ecol* 2012;**81**:696–706.

Zheng W, Song T, Bao X *et al.* High cyanobacterial diversity in coralloid roots of cycads revealed by PCR fingerprinting. *FEMS Microbiol Ecol* 2002;**40**:215–22.

Zimmerman WJ, Rosen BH. Cyanobiont diversity within and among cycads of one field site. *Can J Microbiol* 1992;**38**:1324–8.

Anexos

Tabla 2 Tabla de OTUs identificados por 16s rRNA y en la literatura se busco si estaban reportados como endofito o fijadores de nitrógeno

Phylum	Class	Order	Family	Genus	OTUs	Fijadora de Nitrógeno	Endófito	
				<i>Mucilaginibacter</i>	3			
Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Sphingobacteriaceae	<i>Sphingobium</i>	1			
				<i>Sphingomonas</i>	2			
				<i>Variovorax</i>	1		X	
	Cytophagales	Cytophagales	Cytophagaceae	<i>Dyadobacter</i>	1			
Cyanobacteria	Cyanobacteria	Nostocales	Microchaetaceae	<i>Tolypothrix</i>	6	X		
			Nostocaceae	<i>Nostoc</i>	2	X	X	
Firmicutes	Bacilli	Bacillales	Bacillaceae	<i>Bacillus</i>	16	X	X	
			Paenibacillaceae	<i>Paenibacillus</i>	2	X	X	
			Staphylococcaceae	<i>Staphylococcus</i>	1			
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	<i>Rhizobium</i>	32	X	X	
				<i>Shinella</i>	2			
			Brucellaceae	<i>Ochrobactrum</i>	1			
				Alcaligenaceae	<i>Achromobacter</i>	33	X	X
					<i>Burkholderia</i>	39	X	X
		Betaproteobacteria	Burkholderiales	Burkholderiaceae	<i>Ralstonia</i>	2	X	X
					<i>Mitsuaria</i>	8		X
				Comamonadaceae	<i>Variovorax</i>	1		
			Enterobacteriales	Enterobacteriaceae	<i>Enterobacter</i>	3	X	X
					<i>Luteibacter</i>	1		X
					<i>Pantoea</i>	1	X	X
		Gamma proteobacteria	Pseudomonadales	Pseudomonadaceae	<i>Pseudomonas</i>	21	X	X
					<i>Luteibacter</i>	2		X
		Xanthomonadales	Xanthomonadaceae	<i>Stenotrophomonas</i>	35	X	X	
				<i>Xanthomonas</i>	2		X	
Actinobacteria	Actinobacteria	Micrococcales	Microbacteriaceae	<i>Microbacterium</i>	5			
		Streptomycetales	Streptomycetaceae	<i>Streptomyces</i>	19	X	X	
Total					242			

Tabla 2 Resultados del análisis mOTUs de cada uno de los metagenomas, se resumieron los resultados a nivel de género y se utilizó un 93% de identidad.

Genero	RFPOOL	JPPOOL	RF1m	RF3m	JP2m	JP6m	RF3y1	JP6y1
<i>unidentified</i>	0	0.000203	0	0	0.002305	0	0	0.001657
<i>Achromobacter</i>	0	0	0.023756	0	0.007697	0.008468	0.006163	0.000832
<i>Acidiphilium</i>	0	0	0	0	0	0	0.000768	0
<i>Acidovorax</i>	0	0	0	0	0.004396	0	0	0
<i>Actinomyces</i>	0	0	0	0.003884	0	0	0.00073	0
<i>Actinosynnema</i>	0	0	0	0	0	0	0	0.000798
<i>Afipia</i>	0	0	0	0	0	0	0	0.000883
<i>Agrobacterium</i>	0	0	0.001236	0.036695	0.447503	0.025163	0.003141	0.001964
<i>Ahrensia</i>	0	0	0	0	0	0.001277	0	0.00075
<i>Alicyciphilus</i>	0	0	0	0	0	0	0.001522	0
<i>Alteromonas</i>	0	0	0	0	0.000138	0	0	0
<i>Amycolatopsis</i>	0	0	0	0	0	0	0.000752	0
<i>Amycolicoccus</i>	0	0	0	0	0	0	0	0.000636
<i>Anabaena</i>	0	0	0.000611	0	0.005266	0	0.001437	0.002336
<i>Anaeromyxobacter</i>	0	0.000125	0	0	0	0	0	0
<i>Aromatoleum</i>	0	0	0	0	0	0	0	0.000902
<i>Arthrobacter</i>	0	0	0	0	0	0	0.000759	0
<i>Aurantimonas</i>	0	0	0.001254	0	0	0.017031	0.006784	0.003994
<i>Azorhizobium</i>	0	0	0.003829	0	0	0	0	0
<i>Azospira</i>	0	0	0	0	0.002225	0	0	0
<i>Azospirillum</i>	0	0	0	0	0	0.0099	0	0.002272
<i>Bacillus</i>	0.867081	0.991035	0.003052	0	0.074131	0.009126	0	0
<i>Bifidobacterium</i>	0	0	0	0	0	0	0	0.000783
<i>Bilophila</i>	0	0	0	0	0	0	0.000742	0
<i>Bordetella</i>	0	0	0	0	0.005441	0	0.000745	0
<i>Brachybacterium</i>	0	0	0	0	0	0	0	0.000851
<i>Bradyrhizobiaceae</i>	0	0	0	0	0.001078	0	0.000366	0.0009
<i>Bradyrhizobium</i>	0	0	0.007857	0.00399	0.043108	0.021545	0.00353	0.002581
<i>Brevundimonas</i>	0	0	0	0	0.002182	0.019463	0	0
<i>Brucella</i>	0	0	0	0	0	0.009073	0	0
<i>Burkholderia</i>	0.000528	0.005199	0.003839	0.012114	0.041811	0.00882	0.00531	0.000838
<i>Caulobacter</i>	0	0	0	0	0.03512	0.017317	0.002245	0.008368
<i>Cellulomonas</i>	0	0	0.001249	0	0	0	0	0.000955
<i>Cellvibrio</i>	0	0	0	0	0	0	0	0.001591
<i>Chelativorans</i>	0	0	0.000593	0	0	0	0.001442	0
<i>Citricella</i>	0	0	0.000599	0	0	0	0	0
<i>Citromicrobium</i>	0	0	0	0.008534	0.002236	0	0.002552	0.001687
<i>Clavibacter</i>	0	0	0	0	0	0	0.00963	0.041487

<i>Comamonas</i>	0	0	0	0	0	0	0.000253	0
<i>Corynebacterium</i>	0	0	0	0	0	0	0	0.001418
<i>Cupriavidus</i>	0	0	0	0.003868	0.012766	0	0	0.001491
<i>Cylindrospermopsis</i>	0	0	0	0	0	0	0.000723	0
<i>Deinococcus</i>	0	0	0	0	0	0	0	0.245462
<i>Delftia</i>	0	0	0	0	0	0	0.000364	0
<i>Dermacoccus</i>	0	0	0	0	0	0	0	0.000784
<i>Dyadobacter</i>	0	0	0	0	0	0.026161	0	0.000818
<i>Enterobacter</i>	0.000742	0.002783	0	0	0	0	0	0
<i>Erwinia</i>	0.000348	0	0	0	0	0.001481	0	0
<i>Erythrobacter</i>	0	0	0	0.004121	0.002244	0.004415	0.001612	0
<i>Fulvimarina</i>	0	0	0.001216	0	0	0.008259	0.000715	0
<i>Herbaspirillum</i>	0	0	0	0	0.002208	0	0	0
<i>Hoeflea</i>	0	0	0	0	0	0	0.001423	0
<i>Hyphomicrobium</i>	0	0	0.001251	0	0.004301	0	0.000368	0.001595
<i>Idiomarina</i>	0	0	0	0	0	0.008675	0	0
<i>Intrasporangium</i>	0	0	0	0	0	0	0.000719	0.003118
<i>Isoptericola</i>	0	0	0	0	0	0	0	0.002383
<i>Janibacter</i>	0	0	0	0	0	0	0	0.000813
<i>Kineococcus</i>	0	0	0	0	0	0	0.000753	0.000816
<i>Kitasatospora</i>	0	0	0	0	0	0	0	0.003195
<i>Kocuria</i>	0	0	0	0	0	0	0	0.000827
<i>Kytococcus</i>	0	0	0	0	0	0	0.000711	0
<i>Leifsonia</i>	0	0	0	0	0	0	0.001546	0.005025
<i>Leptothrix</i>	0	0	0	0.004083	0	0.035063	0.000759	0
<i>Magnetospirillum</i>	0	0	0.001344	0	0	0	0	0.000948
<i>marine</i>	0	0	0	0	0	0	0	0.008394
<i>Marinobacter</i>	0	0	0	0	0	0	0	0.000816
<i>Meiothermus</i>	0	0	0	0	0	0	0	0.000839
<i>Mesorhizobium</i>	0	0	0.00501	0.003788	0.012885	0.010089	0.01452	0.014417
<i>Methylibium</i>	0	0	0	0.004059	0	0	0	0
<i>Methylobacterium</i>	0	0	0.004349	0.009447	0.007511	0.047431	0.016935	0.016187
<i>Methylocella</i>	0	0	0	0	0	0	0.000727	0
<i>Methylocystis</i>	0	0	0	0	0	0.008697	0.001634	0.002696
<i>Methylosinus</i>	0	0	0	0	0	0	0.000932	0.001622
<i>Methyloversatilis</i>	0	0	0	0	0	0	0.000761	0
<i>Microbacterium</i>	0	0	0.002648	0.00838	0.002282	0.008996	0.039719	0.148996
<i>Micrococcus</i>	0	0	0	0	0.006861	0	0.000766	0.012039
<i>Micromonospora</i>	0	0	0	0	0	0	0.000746	0
<i>Mycobacterium</i>	0	0	0	0	0	0	0.000741	0.000798
<i>Nitrobacter</i>	0	0	0.002499	0	0.000532	0.01282	0.001481	0.002584
<i>Nocardioideae</i>	0	0	0	0	0	0	0	0.000826
<i>Nocardiosis</i>	0	0	0	0	0	0	0	0.000895

<i>Nodularia</i>	0	0	0	0	0.01172	0	0.002225	0
<i>Nostoc</i>	0	0	0.007902	0.003857	0.044029	0.008258	0.042859	0.006976
<i>Novosphingobium</i>	0	0	0	0.00412	0.004527	0.008845	0.000766	0
<i>Ochrobactrum</i>	0	0	0	0	0	0	0	0.000781
<i>Oligotropha</i>	0	0	0	0	0.001643	0	0	0.001849
<i>Pantoea</i>	0.001328	0	0	0	0	0	0	0
<i>Paracoccus</i>	0	0	0.004853	0	0	0	0.001114	0
<i>Parvibaculum</i>	0	0	0	0.001132	0	0	0	0
<i>Pelagibacterium</i>	0	0	0	0	0	0	0.000729	0
<i>Phenylobacterium</i>	0	0	0	0	0	0	0	0.00082
<i>Plautia</i>	0.000776	0	0	0	0	0	0	0
<i>Polaromonas</i>	0	0	0	0	0	0	0.000127	0
<i>Polymorphum</i>	0	0	0.004954	0	0	0	0.001457	0.001187
<i>Pseudomonas</i>	0.129024	0.000523	0.001268	0	0.017582	0	0.001478	0.0017
<i>Pseudonocardia</i>	0	0	0	0	0	0	0.001432	0
<i>Pseudoxanthomonas</i>	0	0	0.040873	0.189189	0.001087	0.228934	0.207279	0.150427
<i>Ralstonia</i>	0	0.000132	0	0	0	0	0	0
<i>Ramlibacter</i>	0	0	0	0	0.006652	0	0	0
<i>Rhizobium</i>	0	0	0.001224	0.032659	0.041181	0.058195	0.007717	0.006005
<i>Rhodobacter</i>	0	0	0	0.003913	0	0	0.000369	0
<i>Rhodobacteraceae</i>	0	0	0	0	0	0.001412	0	0
<i>Rhodobacterales</i>	0	0	0	0	0	0.001427	0	0
<i>Rhodomicrobium</i>	0	0	0	0	0	0	0.000746	0
<i>Rhodopseudomonas</i>	0	0	0.005416	0.00066	0.008845	0.014863	0.000734	0.008959
<i>Rhodospirillum</i>	0	0	0	0	0	0	0.001481	0.000823
<i>Roseobacter</i>	0	0	0	0	0	0.002748	0	0
<i>Roseovarius</i>	0	0	0.001265	0	0	0	0	0
<i>Ruegeria</i>	0	0	0.000669	0	0	0	0	0
<i>Salmonella</i>	0.000174	0	0	0	0	0	0	0
<i>Sanguibacter</i>	0	0	0	0	0	0	0.000727	0.001614
<i>Shewanella</i>	0	0	0	0	0.000134	0	0	0
<i>Sideroxydans</i>	0	0	0	0	0	0	0	0.00166
<i>Sinorhizobium</i>	0	0	0	0.004461	0.012921	0.024923	0.008152	0.001553
<i>Sphingobium</i>	0	0	0	0.008246	0.004481	0.017549	0.004577	0.00083
<i>Sphingomonas</i>	0	0	0.000675	0.008174	0.098458	0.004451	0.011083	0.006639
<i>Sphingopyxis</i>	0	0	0	0	0	0.008739	0.006809	0.00082
<i>Spirosoma</i>	0	0	0.001288	0	0	0	0	0
<i>Starkeya</i>	0	0	0.006301	0.000669	0.006514	0.020565	0.001731	0.002571
<i>Stenotrophomonas</i>	0	0	0.827425	0.424614	0.001106	0.209418	0.330019	0.15354
<i>Streptomyces</i>	0	0	0	0	0	0	0.002198	0.003168
<i>Thermus</i>	0	0	0	0	0	0	0	0.000843
<i>Trichormus</i>	0	0	0	0	0	0	0.000725	0
<i>Variovorax</i>	0	0	0.001296	0	0.006587	0.008583	0.000743	0

<i>Vibrio</i>	0	0	0	0	0.002035	0	0	0
<i>Xanthobacter</i>	0	0	0	0	0.002083	0	0	0.00077
<i>Xanthomonas</i>	0	0	0.028399	0.215342	0	0.061823	0.225196	0.087221
<i>Xenorhabdus</i>	0	0	0	0	0.002187	0	0	0
<i>Xylanimonas</i>	0	0	0	0	0	0	0	<u>0.002373</u>

Tabla 3 Géneros identificados por mOTUs que han sido reportados como endófitos o fijadores de Nitrógeno (hecho mediante búsqueda en NCBI)

Genero	Fijador de Nitrógeno	Endófito		
<i>Achromobacter</i>	x	x	<i>Fulvimarina</i>	
<i>Acidovorax</i>		x	<i>Herbaspirillum</i>	x x
<i>Actinomyces</i>			<i>Hyphomicrobium</i>	
<i>Agrobacterium</i>	x	x	<i>Idiomarina</i>	
<i>Ahrensia</i>			<i>Leptothrix</i>	
<i>Alteromonas</i>			<i>Magnetospirillum</i>	
<i>Anabaena</i>	x	x	<i>Mesorhizobium</i>	x x
<i>Anaeromyxobacter</i>			<i>Methylibium</i>	
<i>Aurantimonas</i>			<i>Methylobacterium</i>	x
<i>Azorhizobium</i>	x	x	<i>Methylocystis</i>	x
<i>Azospira</i>	x	x	<i>Microbacterium</i>	
<i>Azospirillum</i>	x	x	<i>Micrococcus</i>	
<i>Bacillus</i>	x	x	<i>Nitrobacter</i>	
<i>Bordetella</i>			<i>Nodularia</i>	x
<i>Bradyrhizobiaceae</i>	x	x	<i>Nostoc</i>	x x
<i>Bradyrhizobium</i>	x	x	<i>Novosphingobium</i>	
<i>Brevundimonas</i>			<i>Oligotropha</i>	x x
<i>Brucella</i>			<i>Pantoea</i>	x x
<i>Burkholderia</i>	x	x	<i>Paracoccus</i>	
<i>Caulobacter</i>			<i>Parvibaculum</i>	
<i>Cellulomonas</i>		x	<i>Plautia</i>	
<i>Chelativorans</i>			<i>Polymorphum</i>	
<i>Citricella</i>			<i>Pseudomonas</i>	x x
<i>Citromicrobium</i>			<i>Pseudoxanthomonas</i>	x
<i>Cupriavidus</i>	x	x	<i>Ralstonia</i>	x x
<i>Dyadobacter</i>			<i>Ramlibacter</i>	
<i>Enterobacter</i>	x	x	<i>Rhizobium</i>	x x
<i>Erwinia</i>		x	<i>Rhodobacter</i>	x
<i>Erythrobacter</i>			<i>Rhodobacteraceae</i>	
			<i>Rhodobacterales</i>	

<i>Rhodopseudomonas</i>	x		<i>Spirosoma</i>	
<i>Roseobacter</i>		x	<i>Starkeya</i>	x
<i>Roseovarius</i>			<i>Stenotrophomonas</i>	x
<i>Ruegeria</i>			<i>Variovorax</i>	x
<i>Salmonella</i>			<i>Vibrio</i>	
<i>Shewanella</i>			<i>Xanthobacter</i>	x
<i>Sinorhizobium</i>	x	x	<i>Xanthomonas</i>	x
<i>Sphingobium</i>			<i>Xenorhabdus</i>	
<i>Sphingomonas</i>		x		
<i>Sphingopyxis</i>				

Programas para identificación taxonómica

Programa IdentificadorRF.py: Este programa está codificado en lenguaje Python y su función es encontrar la posición de las secuencias, después cambia todas las secuencias a complementario, esto es algo indispensable para hacer filogenias con las secuencias.

```
#detectar secuencias que están en sentido R y F
#convertir las secuencias en un solo sentido
#J. Antonio Corona gocanjo@hotmail.com

#importar librerías

import re
import sys
import Bio
from Bio import SeqIO
#leer archivo de entrada
try:
    blastout = sys.argv[1]

except:
    sys.exit("ERROR. Can't read supplied filename blastout ")
try:
    querysec = sys.argv[2]
except:
    sys.exit("ERROR. Can't read supplied query sequences")

# crear archivos

print (blastout)
blastoutp = open(blastout, mode='r')
print (blastoutp)
```

```

print (querysec)
querysecp = open(querysec, mode='r')
bioseq=SeqIO.parse(querysec, "fasta")
print (bioseq)
print (querysecp)

#leer blastdb

blastdb = open('SILVA_119_SSURef_Nr99_tax_silva_trunc.fasta', mode='r')
print(blastdb)

#archivos de salida

archivo= 'Fnormal-' + querysec
result = open(archivo, mode='w')

#leer datos de blastout

outdic = {}
poride = 0
claveQ = ''
for l1 in blastoutp:
    datosb = l1.split()
    #print (datosb[0])
    #print (poride)
    a6=int(datosb[6])
    a7=int(datosb[7])
    a8=int(datosb[8])
    a9=int(datosb[9])
    if claveQ == datosb[0]:
        if float(datosb[2]) > poride:
            if a6 <= a7:
                if a8 < a9:
                    outdic[datosb[0]]='F'
                    #print (datosb[0], 'F', datosb[6:10])
                elif a8 > a9:
                    #print (datosb[0], 'R', datosb[6:10])
                    outdic[datosb[0]]='R'
            else:
                print('error', datosb[0], datosb[6], datosb[7],
datosb[8], datosb[9])
                #print (poride, datosb[2])
                poride = float(datosb[2])

        else:
            claveQ = datosb[0]
            poride = 0
#print(outdic)
for sec in bioseq:
    #print(sec.id)
    if outdic[sec.id] == 'R':
        print(sec.id, 'r')
        secr=sec.reverse_complement()
        #print secr
        secr.id=sec.id
        secr.description=sec.description

```

```

        SeqIO.write(secr, result, "fasta")
    print secr
else:
    SeqIO.write(sec, result, "fasta")

result.close()

```

Programa blast_namesprocent.py: Este programa esta codificado en lenguaje Python, su función es extraer los resultados de nBlast y modificar el nombre de las secuencias fasta con el nombre del mejor hit de blast y el porcentaje de identidad.

```

#Extraer datos de blastn y anotar secuencias de 16s con porcentaje y
nombre de la especie
#J.Antonio Corona gocanjo@hotmail.com

#importar librerias

import re
import sys

#leer archivo de entrada
try:
    blastout = sys.argv[1]
except:
    sys.exit("ERROR. Can't read supplied filename blastout ")
try:
    querysec = sys.argv[2]
except:
    sys.exit("ERROR. Can't read supplied query sequences")

# crear archivos

print (blastout)
blastoutp = open(blastout, mode='r')
print (blastoutp)

print (querysec)
querysecp = open(querysec, mode='r')
print (querysecp)

#leer blastdb

blastdb = open('SILVA_119_SSURef_Nr99_tax_silva_trunc.fasta', mode='r')
print(blastdb)

#archivos de salida

archivo= 'name-' + querysec

```

```

result = open(archivo, mode='w')

#leer datos de blastout

outdic = {}
poride = 0
claveQ = ''
for l1 in blastoutp:
    datosb = l1.split()
    #print (datosb)
    #print (poride)
    if claveQ == datosb[0]:
        if datosb[2] > poride:
            outdic[datosb[0]]=datosb[1] + ' ' + datosb[2]
            poride = datosb[2]
    else:
        claveQ = datosb[0]
        poride = 0
#print(outdic)

dicdb = {}
for l2 in blastdb:
    if l2[0] == '>':
        #print(l2)
        taxa = l2
        l2 = l2.split()
        #print(l2)
        nblastdb = re.sub('^>', '', l2[0])
        nblastdb = re.sub('\s+', '', nblastdb)
        taxa = re.sub('^>\w+', '', taxa)
        dicdb[nblastdb]=taxa

fasta = {}
for key in outdic:
    claveQ= outdic[key]
    claveQ= claveQ.split()
    if claveQ[0] in dicdb:
        #print(key,dicdb[claveQ[0]],claveQ[1])
        spname=dicdb[claveQ[0]]
        spname = re.sub('^.+;', '', spname)
        spname =key+' '+claveQ[1]+' '+spname
        fasta[key]=spname
        #print(spname)
for l3 in querysecp:
    if l3[0] == '>':
        fastanm = re.sub('^>', '', l3)
        fastanm = fastanm.split()
        if fastanm[0] in fasta:
            namesp = fasta[fastanm[0]]
            namesp = '>'+namesp
            result.write(namesp)
    else:
        result.write(l3)

querysecp.seek(0)
result.close()
#meter datos en una tabla . cvs o de cualquier tipo

```

```
#print(dicdb)

blastdb.close()
```

Programa clave_intro.py: Este programa esta codificado en lenguaje Python, su función es para introducir modificaciones a los nombres de las secuencias fasta, esto se usó de manera opcional para incluir las claves del cepario en cada una de las secuencias 16s rRNA.

```
#Agregar claves a un formato fasta
#J.Antonio Corona gocanjo@hotmail.com

#importar librerias

import re
import sys

#leer archivo de entrada
try:
    fastaf = sys.argv[1]

except:
    sys.exit("ERROR. Can't read supplied filename fasta ")

# crear archivos

print (fastaf)
fastat = open(fastaf, mode='r')
print (fastat)

#leer base de datos .csv

cclave = open('cultivo_clave.csv', mode='r')
print(cclave)

#archivos de salida

archivo= 'clave-' + fastaf
result = open(archivo, mode='w')

#tabla de datos csv

tabla= fastaf +'.csv'
exel= open(tabla, mode='w')

#leer datos de base de datos
diclave = {}
for l1 in cclave:
```

```

l1=l1.strip()
clave = l1.split(',')
#print (clave)
diclave[clave[0]]=clave[1]
#print (diclave)
for l3 in fastat:
    if l3[0] == '>':
        fastanmc = re.sub('^>', '', l3)
        fastanm = fastanmc.split()
        if fastanm[0] in diclave:
            nclave = diclave[fastanm[0]]
            namesp = '>'+ nclave + ' '+fastanm[2]+' \n'
            exell= fastanm[0]+' '+fastanm[1]+' '+fastanm[2]+'
'+fastanm[3]+' \n'
            result.write(namesp)
            exel.write(exell)
        else:
            result.write(l3)

result.close()
exel.close()

```

Programa besthit_percent.py: Este programa esta codificado en lenguaje Python, su función es extraer los datos de Blast y crear una tabla para poder visualizar los datos, esta tabla es útil para ver los demás resultados de blast en caso que el mejor hit no sea el correcto.

```

#Extraer datos de blastn y crear una tabla de los mejores hits
#uso $ python besthit_namesprocent.py sequences.e1.blastout Fnormal-cut-
all-fasta-nogaps.fasta
#J.Antonio Corona gocanjo@hotmail.com

#importar librerías

import re
import sys

#leer archivo de entrada
try:
    blastout = sys.argv[1]

except:
    sys.exit("ERROR. Can't read supplied filename blastout ")
try:
    querysec = sys.argv[2]
except:
    sys.exit("ERROR. Can't read supplied query sequences")

# crear archivos

```

```

print (blastout)
blastoutp = open(blastout, mode='r')
print (blastoutp)

print (querysec)
querysecp = open(querysec, mode='r')
print (querysecp)

blastdb = open('SILVA_119_SSURef_Nr99_tax_silva_trunc.fasta', mode='r')
print(blastdb)

#archivos de salida

archivo= 'besthit-' + querysec + ".csv"
result = open(archivo, mode='w')

#Crear encabezado para resultados

encabezado ="Query"+'\t'+ "Target-
SilvaID"+'\t'+ "Percent_identity"+'\t'+ "Alignment length"+'\t'+ "#
mismatches"+'\t'+ "# gap opens"+'\t'+ "Start position Q"+'\t'+ "End position
Q"+'\t'+ "Start position T"+'\t'+ "End position T"+'\t'+ "E-value"+'\t'+ "Bit
score"+'\t'+ "Silva-name"
result.write(encabezado)

#leer blastdb
#Leer base de datos blast

dicdb = {}
for l2 in blastdb:
    if l2[0] == '>':
        #print(l2)
        taxa = l2
        l2 = l2.split()
        #print(l2)
        nblastdb = re.sub('^>', '', l2[0])
        nblastdb = re.sub('\s+', '', nblastdb)
        taxa = re.sub('^>\w+', '', taxa)
        dicdb[nblastdb]=taxa

#leer datos de blastout

outdic = {}
poride = 0
claveQ = ''
for l1 in blastoutp:
    datosb = l1.split()
    #buscar nombres
    idname=datosb[1]
    spname=dicdb[idname]
    spname = re.sub('^.+;', '', spname)
    #crear linea de resultado
    textline=l1.rstrip('\n')+'\t'+spname
    #escribir resultado
    result.write(textline)
    #print(textline)
querysecp.seek(0)

```



```

#print linecut[0]
#tf=Tree(linecut[1], format=0)
genero=linecut[0]
#print genero
genero_t[genero]=linecut[1]

print genero_t
def layout(node):
    if node.is_leaf():
        # Add node name to leaf nodes
        N = AttrFace("name", fsize=36, fgcolor="black")
        faces.add_face_to_node(N, node, 0)
        #print N
        #>>>>>>nombre del nodo
        print node.name
        nodename = node.name
        n1 =(genero_t.get(nodename))
        #tf=Tree(genero_t.get[nodename], format=0)
        #print tf
        #t=(genero_t.get[node.name])
        t = Tree(n1, format=0)
        #t.populate(10)

        T = TreeFace(t, small_ts)
        # Let's make the sphere transparent
        T.opacity = 0.8
        # And place as a float face over the tree
        faces.add_face_to_node(T, node, 1, position="aligned")
def get_example_tree():
    # Arbol base
    #t = Tree()
    t=Tree("arbol-general.tre.nwk", format = 0)
    #t.populate(20, random_branches=True)

    # Some random features in all nodes
    #for n in t.traverse():
    #    n.add_features(weight=random.randint(0, 50))

    # Create an empty TreeStyle
    ts = TreeStyle()

    # Set our custom layout function
    ts.layout_fn = layout

    # Draw a tree
    ts.mode = "c"

    # We will add node names manually
    ts.show_leaf_name = False
    # Show branch data
    ts.show_branch_length = True
    ts.show_branch_support = True
    return t, ts

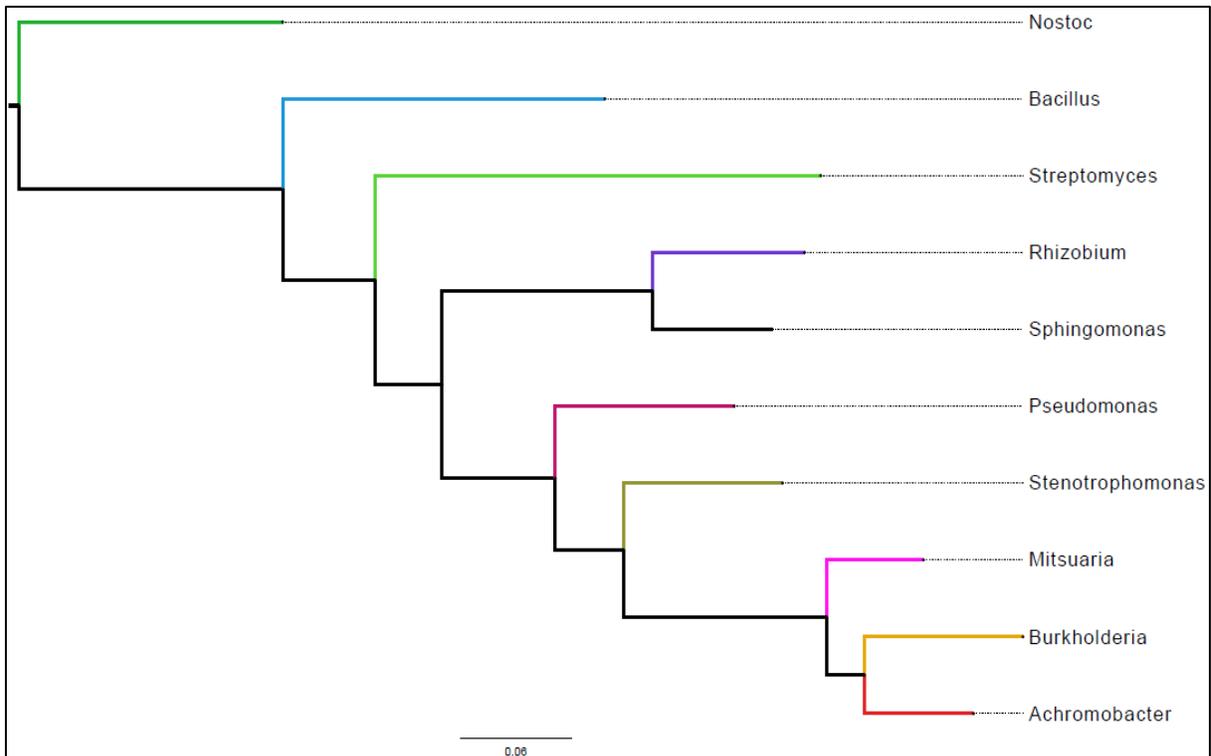
if __name__ == "__main__":
    t, ts = get_example_tree()
    #t.render("tree_faces.png", w=600, dpi=300, tree_style=ts)

```

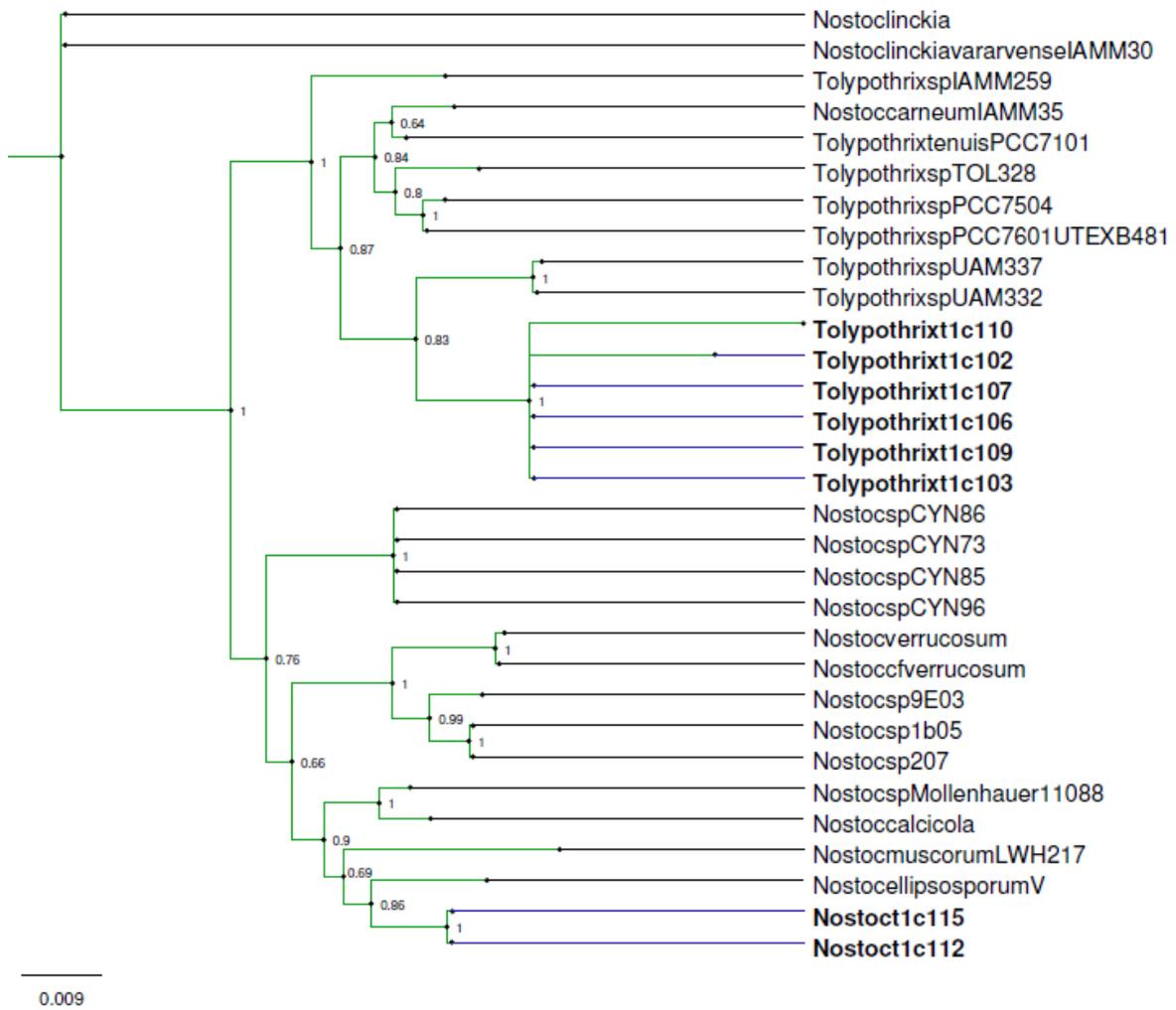
```
t.show(tree_style=ts)
```

Arboles Filogenéticos de árbol de arboles

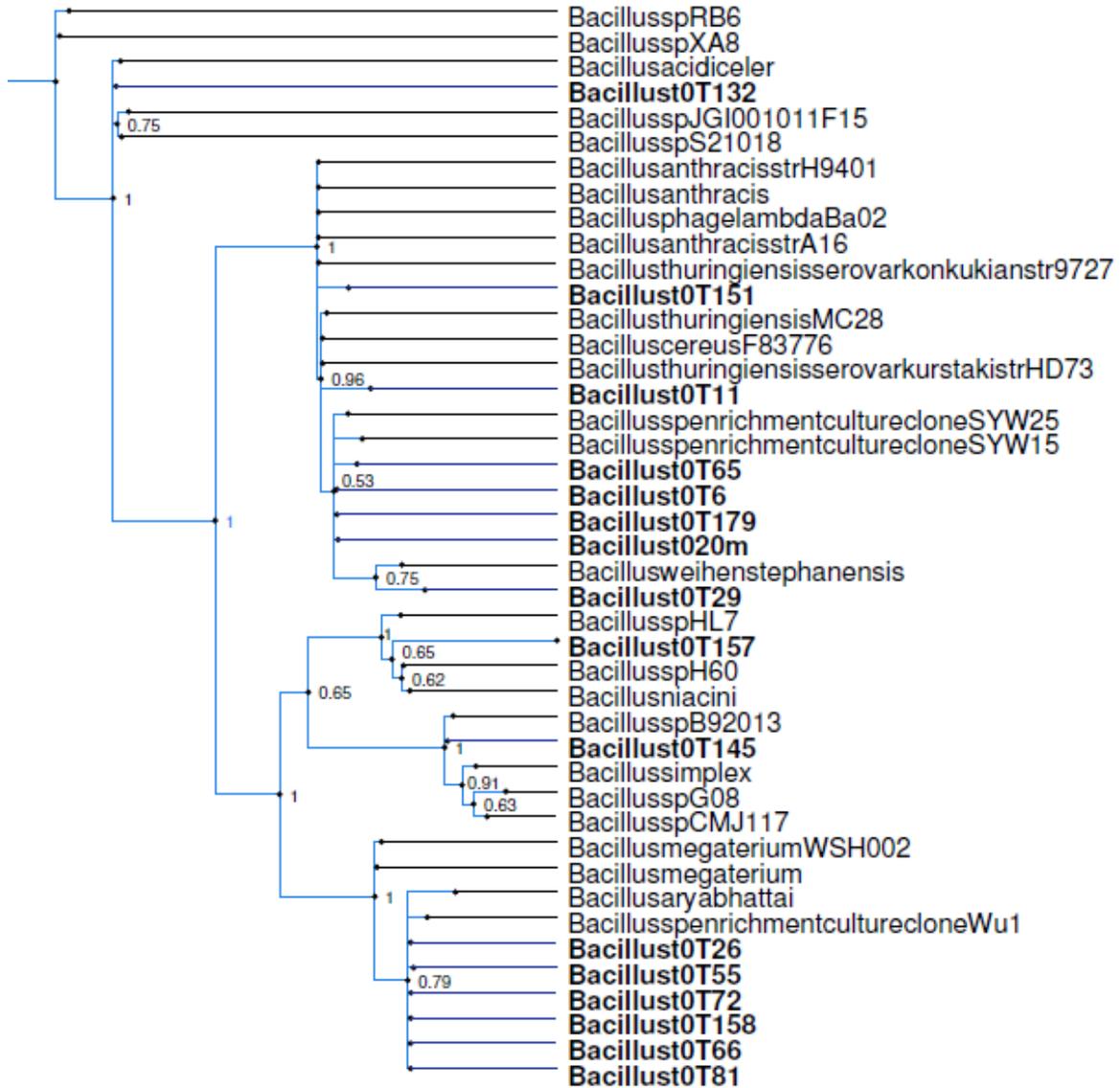
Todos los árboles se hicieron mediante inferencia bayesiana a un millón de generaciones con una distribución de tipo gama.



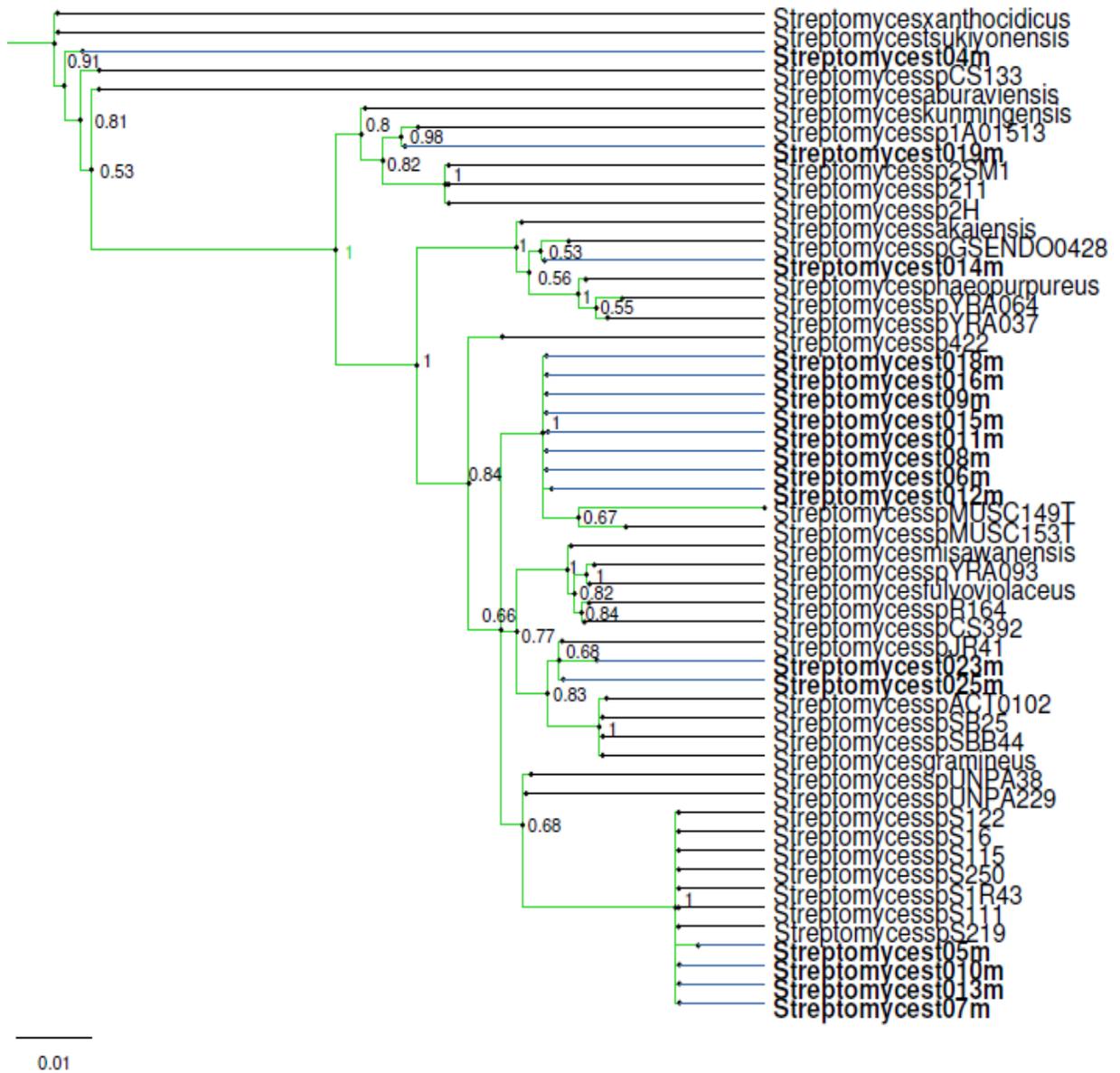
Filogenia general



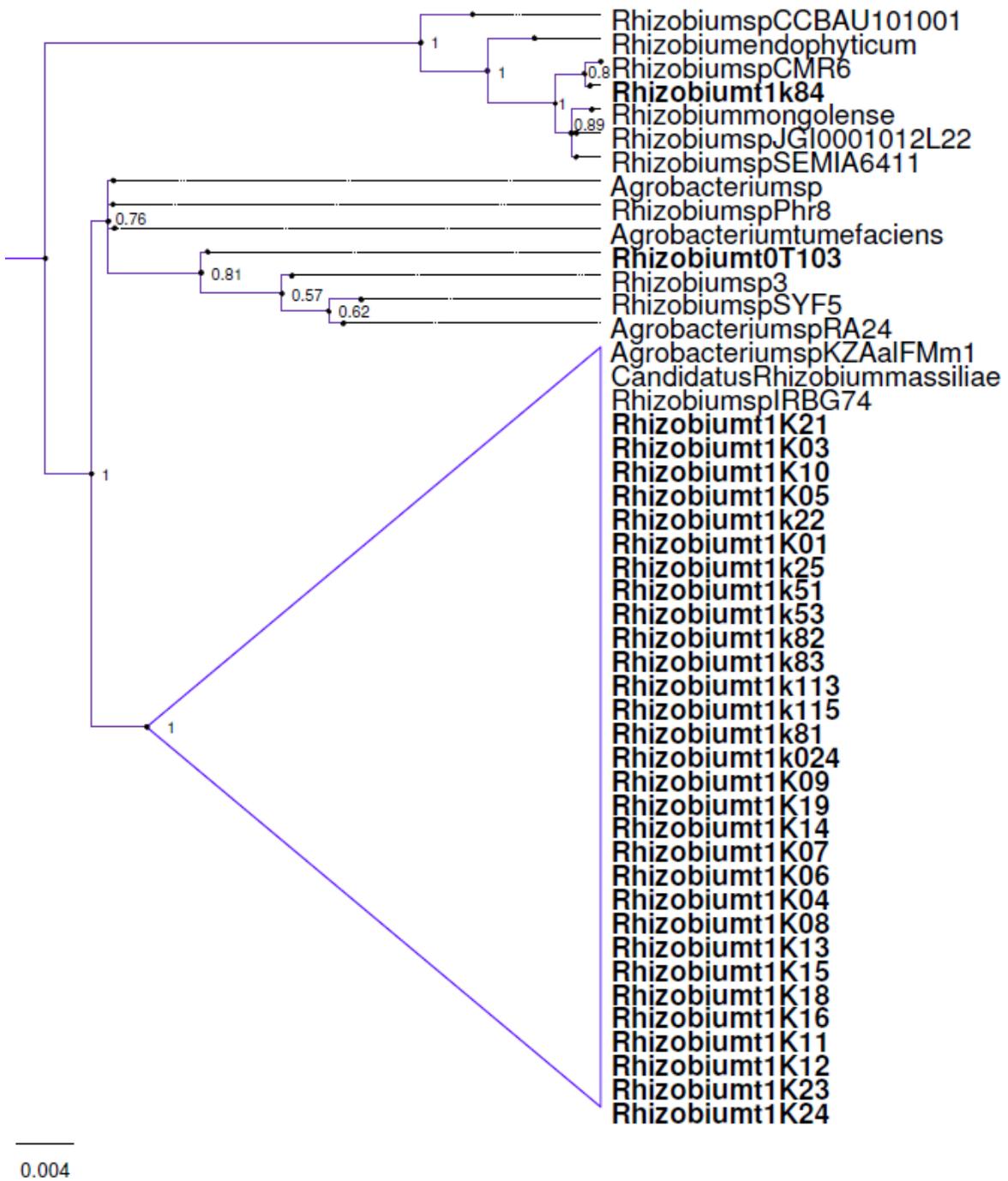
Filogenia del género *Nostoc*



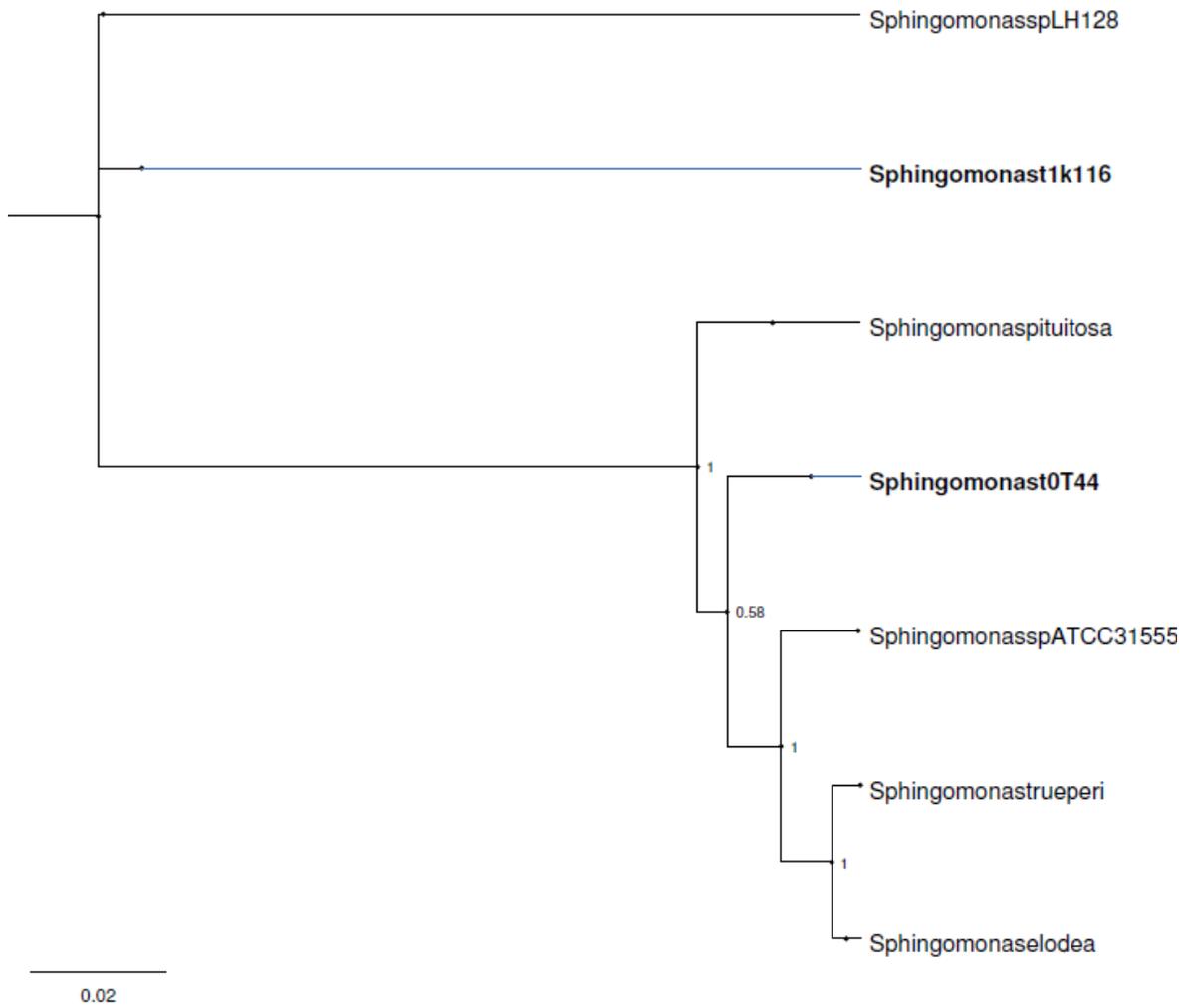
Filogenia del género *Bacillus*



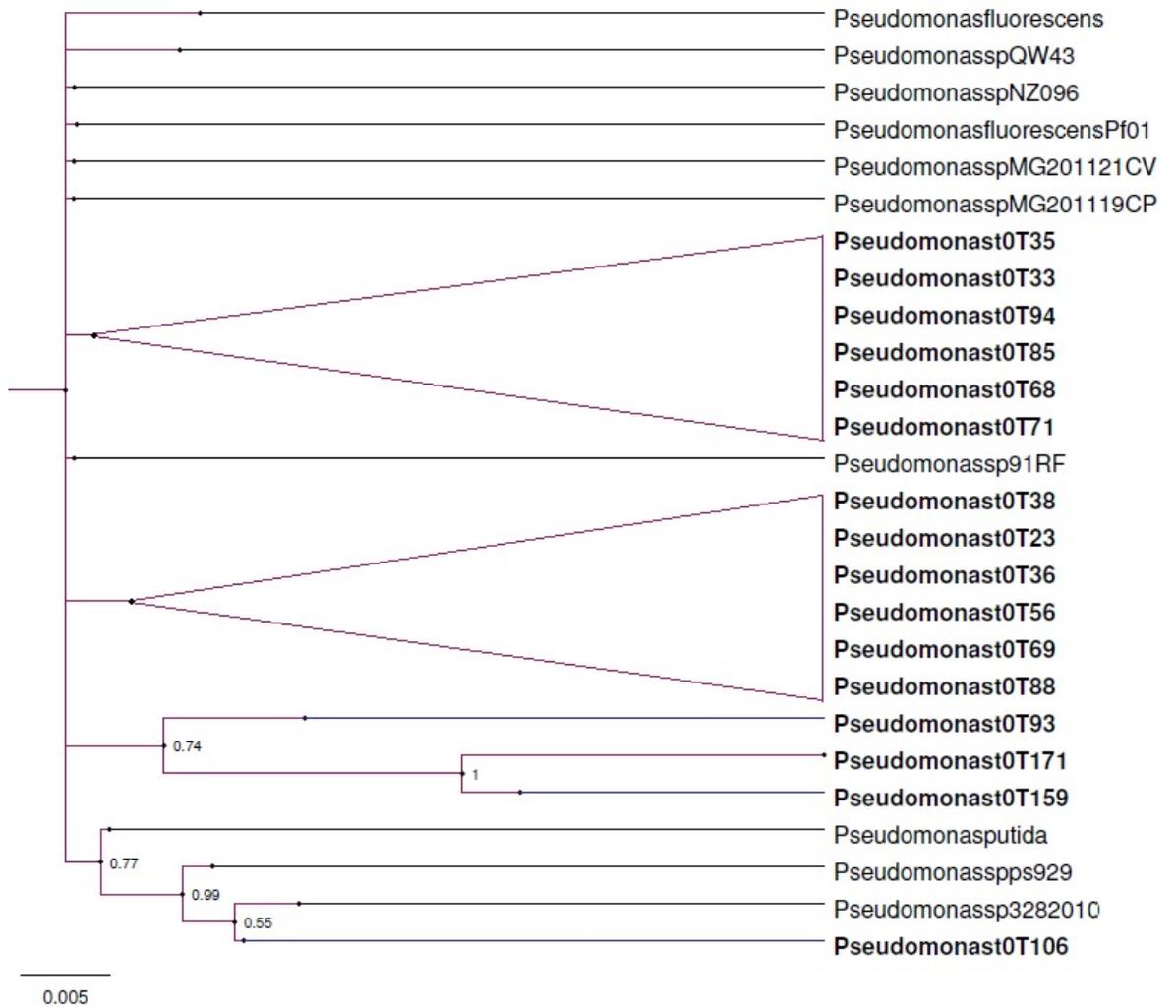
Filogenia del género *Streptomyces*



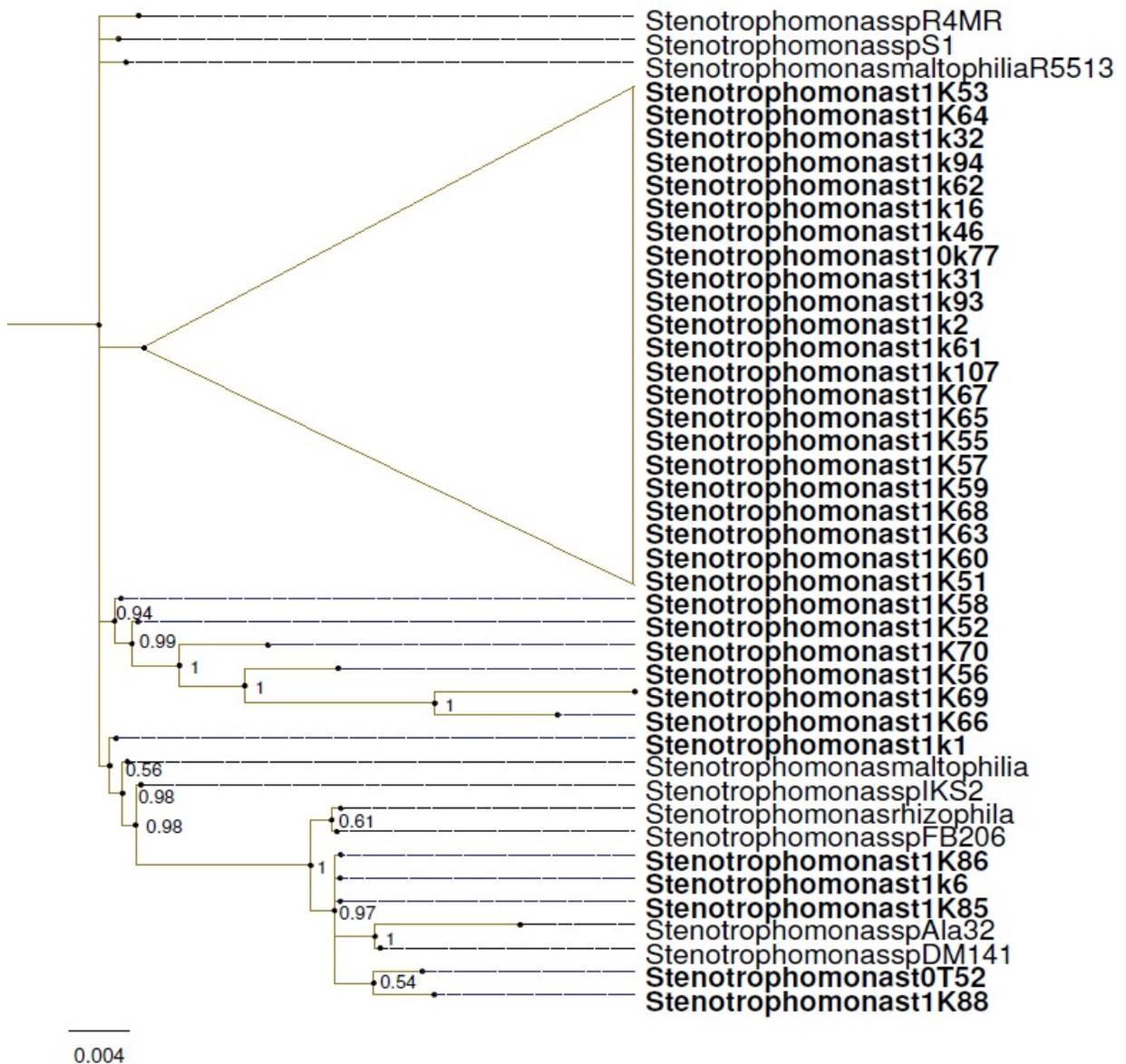
Filogenia del género *Rhizobium*



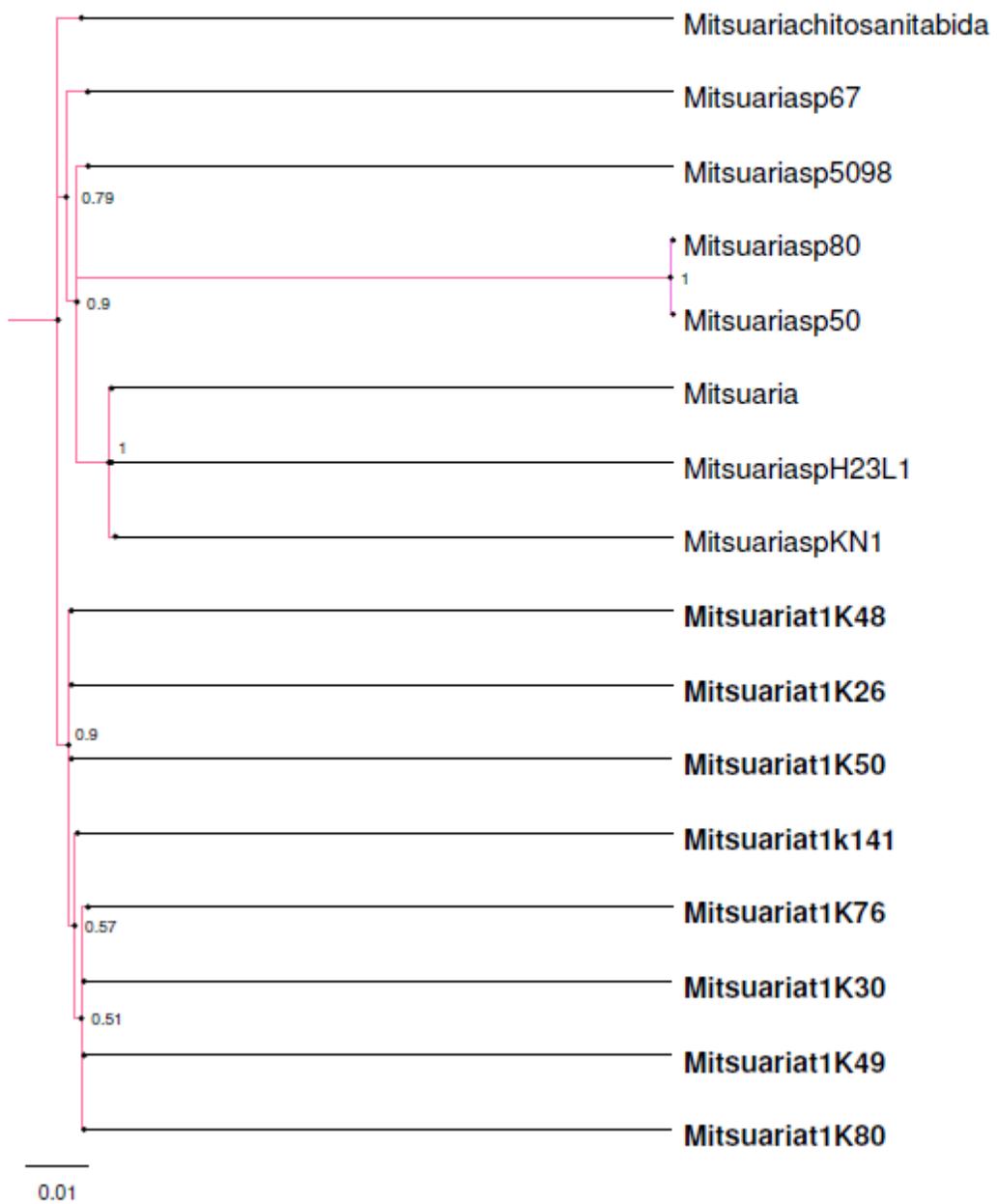
Filogenia del género *Sphingomonas*



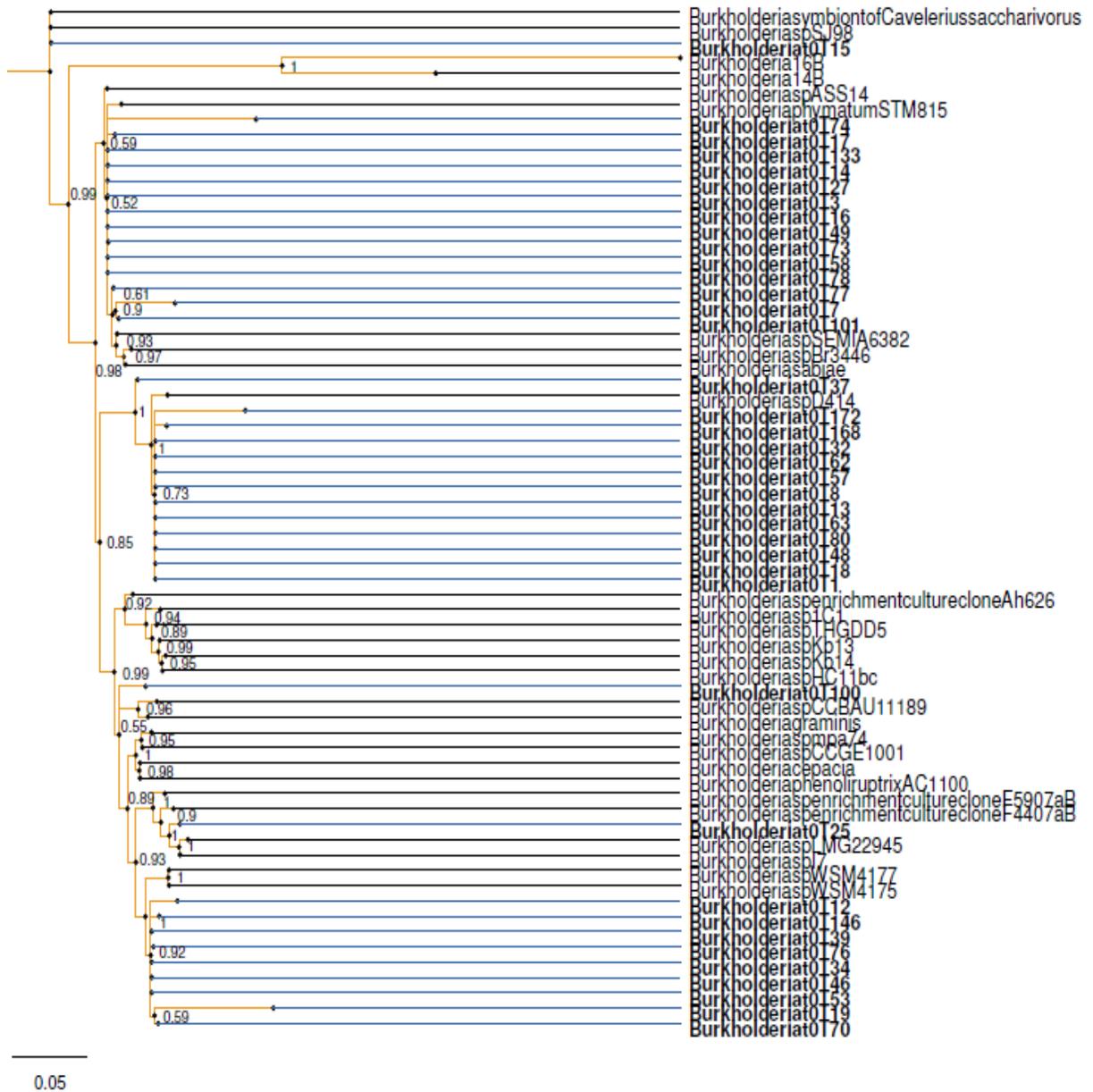
Filogenia del género *Pseudomonas*



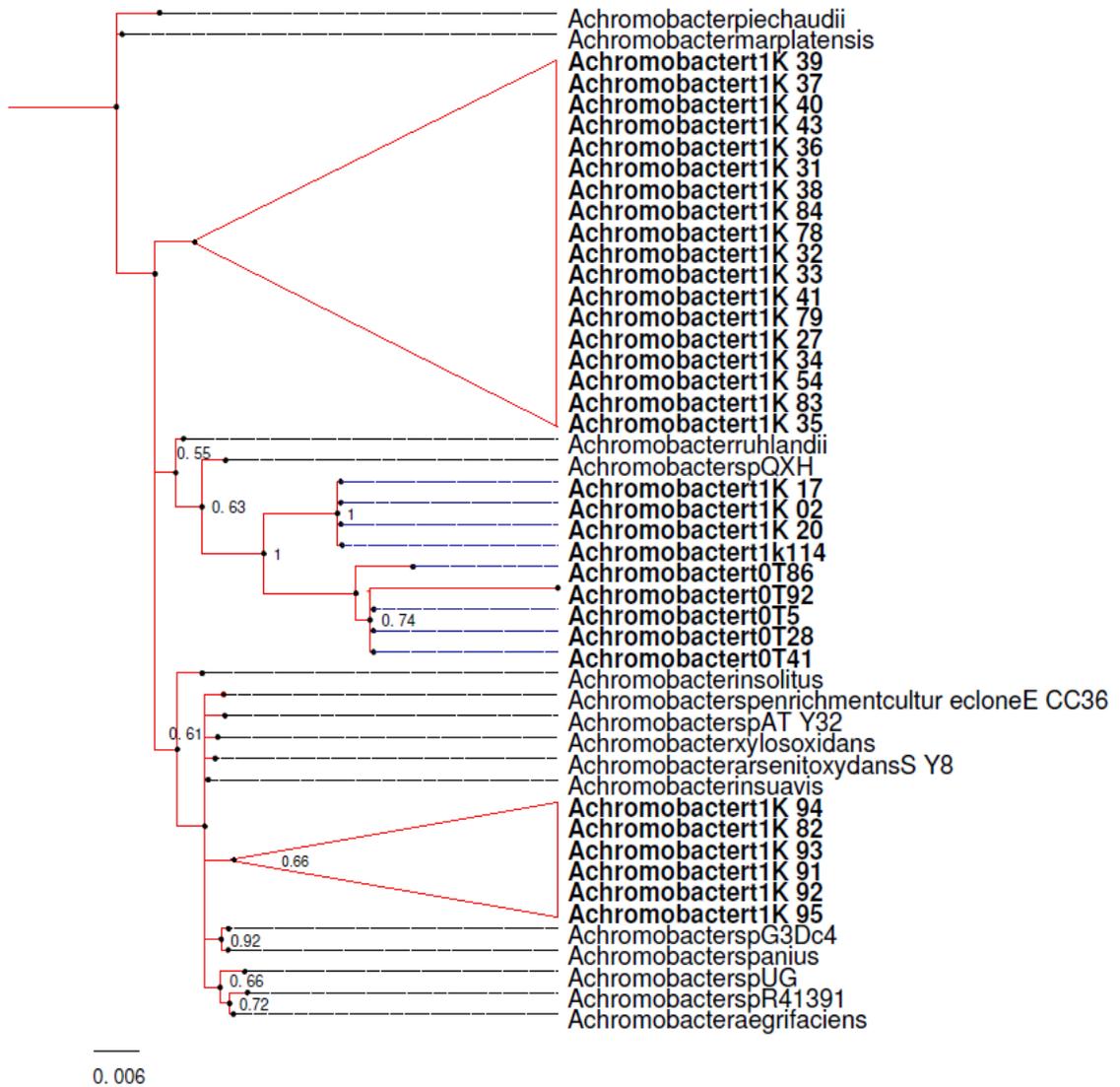
Filogenia del género *Stenotrophomonas*



Filogenia del género *Mitsuaria*



Filogenia del género *Burkholderia*



Filogenia del género *Achromobacter*

Programa ete.py: Este programa esta codificado en lenguaje Python, su función es introducir datos a un árbol filogenético utiliza la librería ete3 del programa ETE, de manera que se puedan visualizar en una sola figura, se necesita una tabla con los metadatos y el resultado es la figura 5.

```

#Hacer arboles con metadatos
#ETE: a python Environment for Tree Exploration
#ETE A Python framework for the analysis and visualization of trees.
#J.A. Corona
#gocanjo@gmail.com
## Activate ETE/Anaconda
## unset PYTHONPATH; export PATH=~/.anaconda_ete/bin:$PATH;

import re
import sys
import os

print os.listdir('.')
print os.system('unset PYTHONPATH; export
PATH=~/.anaconda_ete/bin:$PATH;')

from ete3 import Tree, ClusterTree, NodeStyle, TreeStyle, AttrFace,
ProfileFace, TextFace, faces
#from ete3.treeview.faces import add_face_to_node

#Read Tree
t1 = Tree("tree-order.nwk", format = 0)

print t1
#Read Genero color
genero_colours = {}
f = file("generocol.csv")
for line in f:
    fields=line.rstrip().split(",")
    genero_colours[fields[0]] = fields[1]

# Basic tree style
ts = TreeStyle()
ts.show_leaf_name = False
#ts.show_branch_support = True
ts.mode = "c"
#ts.arc_start = -180 # 0 degrees = 3 o'clock
#ts.arc_span = 360

# Create an independent node style for each leaf, which
# specifies the colour given in the locations.csv file
for n in t1.get_leaves():
    name = n.get_leaf_names()[0]
    nstyle = NodeStyle()
    #nstyle["fgcolor"] = genero_colours[name]
    nstyle["size"] = 5
    #nstyle["fsize"]=15
    nstyle["bgcolor"]= genero_colours[name]

```

```

        n.set_style(nstyle)
        #Node Background

#Read Tiempo text.strip
genero_t = {}
f = file("tiempo.csv")
for line in f:
    fields=line.rstrip().split(",")
    tex=TextFace(fields[1])
    color=fields[2]
    #print(tex)
    # set a border in all faces
    #tex.margin_border.width = 1
    tex.margin_bottom = 2
    tex.margin_top = 2
    tex.margin_right = 5
    tex.margin_left = 5
    tex.border.width = 1
    tex.background.color = color
    genero_t[fields[0]] = tex

for n in t1.get_leaves():
    name = n.get_leaf_names()[0]
    #print (name)
    n.add_face(genero_t[name], 0, "aligned")

#Read ambiente metadata

genero_t = {}
f = file("ambiente.csv")
for line in f:
    fields=line.rstrip().split(",")
    tex=TextFace(fields[1])
    color=fields[2]
    #print(tex)
    # set a border in all faces
    #tex.margin_border.width = 1
    tex.margin_bottom = 2
    tex.margin_top = 2
    tex.margin_right = 5
    tex.margin_left = 5
    tex.border.width = 1
    tex.background.color = color
    genero_t[fields[0]] = tex

for n in t1.get_leaves():
    name = n.get_leaf_names()[0]
    #print (name)
    n.add_face(genero_t[name], 1, "aligned")

#Read medio metadata

genero_t = {}
f = file("medios.csv")
for line in f:
    fields=line.rstrip().split(",")
    tex=TextFace(fields[1])

```

```

        color=fields[2]
        #print(tex)
        # set a border in all faces
        #tex.margin_border.width = 1
        tex.margin_bottom = 2
        tex.margin_top = 2
        tex.margin_right = 5
        tex.margin_left = 5
        tex.border.width = 1
        tex.background.color = color
        genero_t[fields[0]] = tex

for n in t1.get_leaves():
    name = n.get_leaf_names()[0]
    #print (name)
    n.add_face(genero_t[name], 2, "aligned")

#Read organismo metadata

genero_t = {}
f = file("organismo.csv")
for line in f:
    fields=line.rstrip().split(",")
    tex=TextFace(fields[1])
    color=fields[2]
    #print(tex)
    # set a border in all faces
    #tex.margin_border.width = 1
    tex.margin_bottom = 2
    tex.margin_top = 2
    tex.margin_right = 5
    tex.margin_left = 5
    tex.border.width = 1
    tex.background.color = color
    genero_t[fields[0]] = tex

for n in t1.get_leaves():
    name = n.get_leaf_names()[0]
    #print (name)
    n.add_face(genero_t[name], 3, "aligned")

#Agregar imagenes para genomas

t1.show(tree_style=ts)

```

Programa para análisis de diversidad

Programa Diversidad.R: Este programa esta codificado en lenguaje R, su función es a partir de los datos de abundancia 16s rRNA y mOTUs calcular los índices de diversidad y acumulación de especies, esto mediante el paquete vegan.

```
#install.packages("vegan")
library("vegan")
#16s-Genero.csv motus-Genero.csv iTAGr-Genero.csv
#Genero <- read.csv("/home/josian/Dropbox/Tesis_maestria/R-
Analisis/motus-Genero.csv") #linux
Genero <- read.csv("C:/Users/Josian/Dropbox/Tesis_maestria/R-
Analisis/motus-Genero.csv") #windows
A<-matrix(nrow=6)
for (i in 1:nrow(Genero)){
  spval<-Genero[i,-1]*1000
  spval<-as.integer(spval)
  spval<-matrix(data=spval)
  spnam<-matrix(data=Genero[i,1],nrow=1)
  A=cbind(A,spval)
  #print(spnam)
  #print(A)
}
A<-A[,-1]
colnames(A)<-Genero[,1]
#A<-matrix(data=A, nrow=6, +dimnames=list(c(1:6),c(Species[,1])))
BCI<-A
#calculate richness
rich<-function(x){
  x[x>0]<-1
  return(rowSums(x))
}
specnumber(BCI)
BCI_sp<-rich(BCI)
BCI_sp<-specnumber(BCI)
min(BCI_sp)
max(BCI_sp)
#calculate Margalef's or Menhinick's diversity.
D.mg<-function(x){
  r<-rich(x)
  a<-rowSums(x)
  return((r-1)/log(a))
}
#calculate BCI
BCI_Dmg<-D.mg(BCI)
#plot Dmg versus richness
plot(BCI_sp,BCI_Dmg,xlab="Species richness",ylab="Margalef's diversity")
BCI_sp
BCI_Dmg
#estimate diversity Chao's
est.all<-specpool(BCI)
est.plot<-estimateR(BCI)
```

```

par(mfrow=c(1,2))
plot(BCI_sp,BCI_Dmg,xlab="Species richness",ylab="Margalef's diversity")
plot(est.plot[1,],est.plot[2,],xlab="Observed richness",ylab="Estimated
richness (Chao)")
est.plot
min(rowSums(BCI))
rowSums(BCI)[rowSums(BCI)==340]
rare_rich<-rarefy(BCI,min(rowSums(BCI)))
#min and max values
min(rare_rich)
max(rare_rich)
sum(BCI[1,])
rowSums(BCI)[1]
#calculate richness from
#solo lo hace para la fila 1
##crear nuevo dataframe que contenga la suma de las columnas de BCI
#-----#
----#
par(mfrow=c(1,1))
sumBCI<-data.frame()
for(i in 1:ncol(BCI)){
  print(sum(BCI[,i]))
  sumBCI[1,i]<-data.frame(sum(BCI[,i]))
}
tmp.rich<-1:sum(sumBCI[1,])
est.rich<-NULL
for (i in 1:length(tmp.rich)){
  est.rich[i]<-rarefy(sumBCI[1,],i)
}
plot(tmp.rich,est.rich,xlab="Number of individuals",ylab="Estimated
richness")
rare_rich[1] #estimated richness for plot 1
#abline(v=6,lty="dashed")
abline(h=ncol(BCI),lty="dashed")
shan<-diversity(BCI,index="shannon")
simp<-diversity(BCI,index="invsimpson")
shan
simp
D.Hill<-function(x,a){
  tot<-apply(x,1,sum)
  x<-sweep(x,1,tot,"/")
  x<-x^a
  D<-apply(x,1,sum,na.rm=TRUE)
  D<-D^(1/(1-a))
}
Dhill<-D.Hill(BCI,0.4)
par(mfrow=c(1,3))
plot(BCI_sp,shan,xlab="Number of species",ylab="Shannon diversity")
plot(BCI_sp,simp,xlab="Number of species",ylab="Simpson diversity")
plot(BCI_sp,Dhill,xlab="Number of species",ylab="Diversity numbers")
Dhill
E.shan<-shan/log(BCI_sp)
E.simp<-simp/BCI_sp
E.numb<-Dhill/BCI_sp
E.shan
E.simp
#Lets plot all three on the same plot

```

```

par(mfrow=c(1,1))
plot(BCI_sp,E.shan,xlab="Number of species",ylab="Evenness",
      ylim=c(0,1), pch=19)
points(BCI_sp,E.simp,pch=19,col="red")
points(BCI_sp,E.numb,pch=19,col="blue")
#make legend
points(100,0.1,pch=19)
text(100,0.1,pos=4,"Shannon")
points(100,0.05,pch=19,col="red")
text(100,0.05,pos=4,"Simpson")
points(100,0,pch=19,col="blue")
text(100,0,pos=4,"Diversity numbers")
sor<-betadiver(BCI,index="sor")
jac<-betadiver(BCI,index="j")
BC<-vegdist(BCI, method="bray")
par(mfrow=c(1,3))
plot(jac,sor,xlab="Jaccards",ylab="Sorensens")
plot(jac,BC,xlab="Jaccards",ylab="Bray-Curtis")
plot(sor,BC,xlab="Sorensens",ylab="Bray-Curtis")
BC.dis<-1-BC
par(mfrow=c(1,3))
plot(jac,sor,xlab="Jaccards",ylab="Sorensens")
plot(jac,BC.dis,xlab="Jaccards",ylab="Bray-Curtis (dis)")
plot(sor,BC.dis,xlab="Sorensens",ylab="Bray-Curtis (dis)")
ncol(BCI)/mean(BCI_sp)-1
ncol(BCI)-mean(BCI_sp)
abc<-function(x){
  gam<-ncol(x)
  alph<-mean(specnumber(x))
  whit<-gam/alph-1
  lande<-gam-alph
  out<-c(gam,alph,whit,lande)
  names(out)<-c("Gamma", "Mean Alpha", "Whittaker's Beta", "Lande's
Beta")
  return(out)
}
abc(BCI)
#Ploteo de datos para diferentes ambientes y metodos
Humid<-shan[c(1,3,4)]
Dry<-shan[c(2,5,6)]
Shanonb<-cbind(Humid,Dry)
par(mfrow=c(2,1))
boxplot(Shanonb,horizontal = TRUE,main="Environment",xlab="Shanon
index",col = "lightgray")
BG11<-shan[3:6]
CicadBroth<-shan[1:2]
Shanonm<-list(BG11=(BG11),CicadBroth=(CicadBroth))
boxplot(Shanonm,horizontal = TRUE,main="Method",xlab="Shanon index",col =
"lightgray")
par(mfrow=c(1,1))
#Ploteo de valores de shanon
shan
colnames(Genero[-1])
#class(mtcars)
Met<-colnames(Genero[-1])
Met
Vsha<-shan

```

```

Shannon<-shan
Shannon
Esha<-E.shan
Esha
Gam<-BCI_sp
Gam
Simpson<-simp
Simpson
#Est.rich<-est.rich[-10][[-9]
#Est.rich
Rare.rich<-rare_rich
Vsha
Sha<-data.frame(Met,Vsha)
Sha
dotchart(Sha$Vsha, labels=Sha$Met, cex=.7,
          main="Divesity index",
          xlab="Shanon")
resumen<-data.frame(colnames(Genero[-
1]), Shannon, Esha, Gam, Simpson, Rare.rich)
resumen
write.table(resumen, "/home/josian/Dropbox/Tesis_maestria/R-
Analysis/motus-Genero-diver-index.csv", sep=",")

#curva de acumulacion por metodo
muestreo<-colnames(Genero[-1])
BCI[1,]
nrow(BCI)
rarefaccion<-function(x) {
  R1=1:sum(x)
  R1.rich<-NULL
  for (i in 1:length(R1)){
    R1.rich[i]<-rarefy(x,i)
  }
  R2<-list(R1,R1.rich)
  return(R2)
}
m1<-rarefaccion(BCI[1,])
m2<-rarefaccion(BCI[2,])
m3<-rarefaccion(BCI[3,])
m4<-rarefaccion(BCI[4,])
m5<-rarefaccion(BCI[5,])
m6<-rarefaccion(BCI[6,])
m7<-rarefaccion(BCI[7,])
m8<-rarefaccion(BCI[8,])
#x<-as.numeric(unlist(m[1]))
#y<-as.numeric(unlist(m[2]))

plot(tmp.rich,est.rich,type="l",col="blue",xlab="Number of
individuals",ylab="OTUs", main="Metagenome's rarefaction Curve")
#legend("topleft", names(autos_data), cex=0.8, col="blue", lty=1,
bty="n");
abline(h=ncol(BCI),lty="dashed")
lines(as.numeric(unlist(m1[2])), type="l", pch=22, lty=5, col="purple")
lines(as.numeric(unlist(m2[2])), type="l", pch=22, lty=5, col="red")
lines(as.numeric(unlist(m3[2])), type="l", pch=22, lty=5, col="green")
lines(as.numeric(unlist(m4[2])), type="l", pch=22, lty=5, col="cyan")

```

```
lines(as.numeric(unlist(m5[2])), type="l", pch=22, lty=5, col="orange")
lines(as.numeric(unlist(m6[2])), type="l", pch=22, lty=5,
col="lightblue")
lines(as.numeric(unlist(m7[2])), type="l", pch=22, lty=5, col="darkblue")
lines(as.numeric(unlist(m8[2])), type="l", pch=22, lty=5, col="pink")
abline(h=ncol(BCI),lty="dashed")
# Create a legend in the top-left corner that is slightly
# smaller and has no border
#legend("topleft", names(autos_data), cex=0.8, col=plot_colors, lty=1:3,
lwd=2, bty="n");
#help(lines)
```