



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
DOCTORADO EN CIENCIAS BIOMÉDICAS

ECOLOGÍA MOLECULAR MICROBIANA: ESTUDIOS
DE COMUNIDADES Y BIOGEOGRAFÍA EN SISTEMAS
ACUÁTICOS

TESIS

QUE PARA OPTAR POR EL GRADO DE
DOCTORA EN CIENCIAS

PRESENTA

LAURA ESPINOSA ASUAR

TUTORA: DRA. VALERIA SOUZA SALDIVAR
INSTITUTO DE ECOLOGÍA

MÉXICO D.F. MAYO DE 2015



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

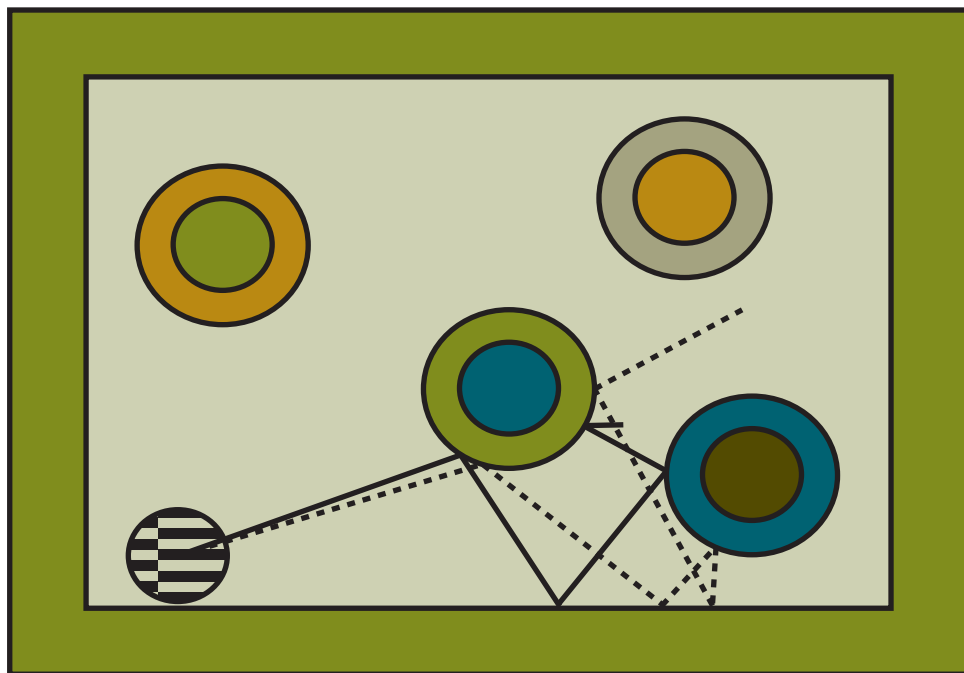
DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



ECOLOGÍA MOLECULAR MICROBIANA: ESTUDIOS DE COMUNIDADES Y BIOGEOGRAFÍA EN SISTEMAS ACUÁTICOS



**LAURA ESPINOSA ASUAR
INSTITUTO DE ECOLOGÍA, UNAM**

MÉXICO D.F.

MAYO 2015

Esta tesis se realizó bajo la asesoría de la Dra. Valeria Souza Saldivar, en el Laboratorio de Evolución Molecular y Experimental del Instituto de Ecología de la Universidad Nacional Autónoma de México, con el apoyo financiero de los proyectos Semarnat/Conacyt 2002-C01-0246, SEP-CONACyT 0237A1 y World Wildlife Found (WWF) - Alianza Carlos Slim ECOIE410.

El comité tutorial que asesoró el desarrollo de esta tesis estuvo formado por:

Dra. Valeria Souza Saldivar, Instituto de Ecología, UNAM.

Dr. Pablo Vinuesa Fleischmann, Centro de Ciencias Genómicas, UNAM.

Dr. Héctor Arita Watanabe, CIEco, UNAM.

El jurado de examen doctoral estuvo constituido por:

Dra. María del Carmen Gómez Eichelmann	Presidente
Dr. Alfonso Valiente Banuet	Vocal
Dra. Ella Vázquez Domínguez	Vocal
Dr. Luis David Alcaraz Peraza	Vocal
Dra. Valeria Souza Saldivar	Secretaria

Figura de portada: Ruelle D (2003) *Casualidad y Caos*. México. Dirección General de Divulgación de la Ciencia, UNAM.

Para NJ
A Diego y Emilio

Por eso, sólo podemos afirmar con certeza que el todo sólo es igual a la suma de las partes cuando las partes se ignoran entre sí.

Almudena Grandes

AGRADECIMIENTOS ACADÉMICOS

A la Universidad Nacional Autónoma de México

Al Instituto de Ecología

Al Doctorado en Ciencias Biomédicas

A los miembros del comité tutorial: Dr. Pablo Vinuesa Fleischmann y Dr. Héctor Arita Watanabe

A los miembros del jurado: Dra. María del Carmen Gómez Eichelmann, Dr. Alfonso Valiente Banuet, Dra. Ella Vázquez Domínguez, Dr. Luis David Alcaraz Peraza, Dra. Valeria Souza Saldivar

A todos los integrantes del Laboratorio de Evolución Molecular y Experimental

A Erika Aguirre Planter, Gabriel Manuel Rosas Barrera y Silvia Barrientos Villanueva por el apoyo técnico brindado durante el proyecto

A Erika Aguirre, a quien le debo las horas de trabajo doble

A Alejandra Vázquez-Lobo, mi mentora en R

A Valeria Souza y Luis Eguiarte, por tantos años de aprendizaje

A Ana Escalante, que es parte primordial de esta tesis

A Concepción Asuar, editora incansable

A Silvia G. Acinas por su asesoría con la técnica de T-RFLPs

A Ana María Noguez por su ayuda en el diseño experimental

A Daniel Piñero que generosamente compartió su espacio

A Jaime Gasca por la asesoría estadística

A Santiago Ramírez y Laura Figueroa, que colaboraron en las figuras

A Juan Carlos Ramírez Gloria (PRONATURA), Rodrigo González Chauvet, Luis Eguiarte, Eria Rebollar, Morena Avitia, Esmeralda López-Lozano, René Cerritos, Enrique Scheinvar y Germán Bonilla por su ayuda en el campo

A Felipe García Oliva (CIEco) que realizó el análisis de nutrientes del agua

A mis alumnas Jazmín Blaz y Lorena Peña, porque mucho aportaron a esta tesis

ÍNDICE

RESUMEN	1
ABSTRACT	2
INTRODUCCIÓN	3
1 Definiciones básicas	4
1.1 Ecología y biogeografía	4
1.2 Diversidad alfa, beta y gamma	5
1.3 Metacomunidades, otra perspectiva para estudiar a las comunidades	5
1.4 Una síntesis conceptual: cuatro procesos propuestos, cuatro modelos y cuatro aproximaciones en metacomunidades	6
1.5 Confusiones sobre algunos conceptos o términos	8
2 Algunos patrones de distribución relacionados con este trabajo	11
2.1 Decaimiento por distancia	11
2.2 Invarianza regional	12
2.3 Patrones de aislamiento geográfico, ambiente físico e interacciones bióticas en un contexto histórico	12
3 Patrones en comunidades microbianas ¿qué se sabe?	14
3.1 Teorías de macroorganismos aplicadas a procariontes	14
3.2 De dicotomías y formas de aproximarse a los mecanismos que estructuran a las comunidades microbianas	16
3.3 Sobre patrones de distribución, <i>taxa</i> y generalizaciones	19
3.4 Particularidades de los patrones de distribución y diversidad encontrados en microorganismos	20
4 Diversidad y biogeografía de sistemas acuáticos bacterianos	22
5 Los sistemas acuáticos del desierto, el caso de Cuatro Ciénegas	24
OBJETIVOS	27
CAPÍTULO I. ARTÍCULO PUBLICADO “Comparación de tres métodos moleculares para el análisis de procariontes ambientales en el mar del canal de Yucatán”	28
CAPÍTULO II. ARTÍCULO EN PROCESO DE ENVIARSE “Aquatic bacterial assemblage structure in Pozas Azules, Cuatro Ciénegas Basin in México: neutral vs. niche-based factors”	42
CAPÍTULO III. NOTAS SOBRE EL MUESTREO REGIONAL “Análisis de la distribución y diversidad de las bacterias acuáticas en el desierto Chihuahuense (Coahuila)”	83
DISCUSIÓN GENERAL	108
1 Sobre patrones y métodos moleculares	109
2 De patrones y factores	110
2.1 Sobre patrones triangulares	112
3 Patrones y escalas	116
4 De los cuatro procesos	117
5 Sobre los <i>taxa</i>	118
6 Sobre patrones en Cuatro Ciénegas	119
CONCLUSIONES	122
PERSPECTIVAS	123
BIBLIOGRAFÍA DE LA INTRODUCCIÓN Y DISCUSIÓN GENERAL	125

RESUMEN

Las causas de la distribución de la diversidad procarionte han sido objeto de investigación en ecología microbiana durante casi dos décadas. Actualmente se consideran principalmente dos grupos de factores para explicar los patrones de distribución y diversidad observados: aquellos en los que el efecto geográfico es preponderante, conocidos como factores neutros y, aquellos en los que la adaptación local a las condiciones ambientales es más influyente, llamados factores asociados a nicho. En esta tesis se usaron tres estrategias diferentes para describir la distribución y diversidad de comunidades de bacterias acuáticas. Se aplicaron distintos métodos moleculares, dirigidos en todos los casos sobre el gen 16S ADNr, aunque desde enfoques y escalas diversas. La primera estrategia fue un análisis comparativo en una muestra de agua de mar (canal de Yucatán) utilizando tres métodos (T-RFLPs, clonas y *shotgun*) (Capítulo I). La segunda estrategia fue a partir del análisis de secuencias obtenidas de librerías de clonas en un muestreo de escala local, con una distancia máxima de 1.8 km, en cuatro sistemas acuáticos cerrados (Pozas Azules, Cuatro Ciénegas, Coahuila), cuyas condiciones ambientales son similares (Capítulo II). Por último, la técnica de T-RFLPs, se aplicó a un muestreo hecho en 27 sistemas acuáticos alrededor y en la región de Cuatro Ciénegas, Coahuila, en condiciones ambientales contrastantes, a una escala regional, con una distancia máxima de 180 km (Capítulo III). De acuerdo a los resultados obtenidos, se reflexiona sobre la factibilidad y relevancia de estudiar comunidades microbianas con tecnologías sencillas y económicas, las cuales proporcionan información importante, así como patrones ecológicamente válidos. Asimismo, este trabajo permitió entender mejor la influencia de factores asociados a nicho (que ha sido considerado como el factor más importante para la distribución de la diversidad microbiana) vs. factores neutros. A mayor escala se detectó un efecto ambiental (nicho), y a menor escala se detectó un efecto geográfico (neutro); este efecto neutro no había sido reportado previamente en otros trabajos sobre bacterias en Cuatro Ciénegas. Por otra parte, las comunidades microbianas estudiadas presentaron patrones de distribución similares a comunidades de macroorganismos: efecto de distancia y distribución restringida de OTUs. Tanto el efecto geográfico (neutro) como estas evidencias, reflejan cierto grado de aislamiento en las comunidades bacterianas muestreadas. De esta forma, la generalización que existe en la literatura sobre las grandes tasas de dispersión de los microorganismos debido a su tamaño pequeño resulta imprecisa, y las evidencias actuales y de esta tesis respaldan que los microorganismos también están sujetos y responden a los mismos procesos ecológicos que se han observado en macroorganismos.

ABSTRACT

The causes for the observed microbial diversity distribution have been subject of investigation in microbial ecology for almost two decades. To date, there are mainly two groups of factors in consideration for explaining the causes of the observed patterns: those in which historical or geographical effects are preponderant, also known as neutral-based factors and, those in which local adaptation to environmental conditions are more influential, also known as niche-based factors. In this thesis, three different strategies were used to describe the distribution and diversity of aquatic bacteria communities, by using different molecular methods, all targeting the gene 16S rDNA, although at several scales and from different approaches. The first strategy was a comparative analysis on a seawater sample (Yucatan Channel) using three methods (T-RFLPs, clone libraries and metagenome shotgun sequencing) (Chapter I). The second strategy was based on the analysis of sequences obtained from clone libraries on a local sampling (with a maximum distance of 1.8 km), where similar environmental conditions are chosen in four closed aquatic systems from Pozas Azules, Cuatro Ciénegas, Coahuila, (Chapter II). Finally, the T-RFLPs technique was applied to 27 different water systems within and around Cuatro Ciénegas, Coahuila, at a regional scale (with a maximum distance of 180 km), having contrasting environmental conditions between samples (Chapter III). This thesis permitted us to examine the feasibility and relevance of studying microbial communities with simple and inexpensive technologies, which provide important information as well as ecologically valid patterns. Furthermore, the used approaches allowed us to better understand the influence of niche-associated factors (which has been considered as the most important factors for the distribution of microbial diversity) vs. neutral factors. On a larger scale, an environmental effect (niche) is detected, and at smaller scale, geographical effect (neutral) has been detected; this neutral effect had not been previously reported in other studies on bacteria in Cuatro Ciénegas. Moreover, this microbial communities study showed patterns similar to communities of macroorganisms: isolation by distance and restricted distribution of OTUs. These evidences reflect some degree of isolation in the bacterial communities sampled. Thus, the generalization in the literature affirming large dispersal rates of microorganisms due to its small size is imprecise; current evidences and this thesis support that microorganisms respond to the same ecological processes that have been observed in macroorganisms.

INTRODUCCIÓN

1 Definiciones básicas

1.1 Ecología y biogeografía

Las disciplinas principales cuyas teorías y planteamientos se han utilizado para comprender la distribución de los organismos son la biogeografía y la ecología. Aún y cuando ambas disciplinas surgen históricamente al mismo tiempo (finales del siglo XIX), en la actualidad tienen enfoques distintos. Por ejemplo, la biogeografía analiza patrones de distribución de especies espacio-temporales (Lomolino *et al.*, 2006), generalmente haciendo uso de registros fósiles y filogenias, en donde predomina una visión relacionada con la evolución de grupos, su extinción y su distribución a través del tiempo; la ecología por otra parte analiza las relaciones de los organismos con su medio ambiente, así como entre ellos (Begon *et al.*, 2006). La teoría generada por ambas ha sido importante para comprender la distribución de los organismos en general, y además resulta complementaria, ya que la distribución biogeográfica de todos los organismos es parte importante de su ecología, siendo esta distribución, causa y consecuencia de la interacción de los seres vivos con factores bióticos y abióticos.

La ecología de comunidades es el estudio de los patrones de comportamiento y estructura de ensambles biológicos, esto es, de grupos formados por diferentes especies que interactúan y concurren en espacio y tiempo (Begon *et al.*, 2006). Tiene como objetivo explicar cómo se distribuye la biodiversidad en las comunidades, si existen patrones de abundancia y cuáles son los mecanismos o procesos que mantienen a esa diversidad biológica.

Es importante resaltar que una comunidad de organismos tiene, por un lado, propiedades que derivan de la suma de las propiedades de cada uno de los individuos que la compone. Pero por otro lado, también existen interacciones entre ellos, lo que hace a una comunidad algo más que solamente la suma de sus partes (Begon *et al.*, 2006). Desde estos dos encuadres, la ecología de comunidades tiene como objetivo central entender los mecanismos por los cuales determinadas especies, y no otras, son las que forman parte de una

comunidad, comparando con la “poza regional” o el total de especies (Levins & Lewontin 1985).

1. 2 Diversidad alfa, beta y gamma

Whittaker desde el año de 1960 propuso tres niveles para aproximarse a la descripción de la distribución y abundancia de los organismos, denominadas diversidad *alfa*, *beta* y *gamma*, que desde entonces se han utilizado ampliamente y han resultado herramientas muy valiosas para la ecología de comunidades. “La diversidad *gamma* es el número de especies a nivel regional, la diversidad *alfa* se define como el número de especies a nivel local, y la diversidad *beta* es, en su definición más general, la diferencia en composición de especies entre comunidades” (Rodríguez & Vázquez-Domínguez, 2003).

Los cálculos para determinar la diversidad *alfa* y *gamma* se han hecho siempre bajo el mismo precepto. En cambio, la forma de cuantificar a la diversidad *beta* se hace desde numerosos enfoques. La primera propuesta fue la fórmula $\beta = \gamma/\alpha$, donde la diversidad β es resultado de la diversidad γ dividida entre la diversidad α promedio (Whittaker 1957). Otra manera es definir a la diversidad *beta* como el porcentaje de similitud que existe entre dos muestras (Whittaker 1957, Calderón-Patrón *et al.*, 2012), y esta ha sido la forma más común de calcularla hasta ahora. Frecuentemente se describe una “alta diversidad *beta*” para referirse a comunidades distintas entre sí. Dado que es una medida de similitud, lo correcto sería afirmar lo contrario, que la diversidad *beta* es baja (esto es, poco similar), pero es una práctica común convertir el valor de la diversidad *beta* en una medida de disimilaridad, calculada a partir del valor inverso de la similitud.

1.3 Metacomunidades, otra perspectiva para estudiar a las comunidades

El concepto de metacomunidad es de reciente adquisición en la teoría de ecología de comunidades (Logue *et al.*, 2011). Este término se refiere a “comunidades múltiples conectadas por organismos que se dispersan, y que son afectadas por factores a nivel local y regional” (Hortal *et al.*, 2014).

La teoría de metacomunidades plantea cuatro modelos o aproximaciones para entender estos factores, que brevemente consisten en (definiciones tomadas de Lee 2014, y Gravel, Poisot & Desjardins-Proulx 2014, a menos que se especifique):

i) Dinámica de parches (*Patch Dynamics* o PD por sus siglas en inglés): existen parches de hábitats idénticos, cuya diversidad está determinada por dispersión o interacción entre especies (colonización-competencia).

ii) Repartición de especies (*Species Sorting* o SS): predice que las distintas condiciones ambientales históricas producen reacomodos de especies, lo cual genera mezclas de estas especies que presumiblemente evolucionaron en distintos tiempos y escenarios, por lo que habrá diferencias en la composición de las comunidades.

iii) Efecto de masas (*Mass Effect* o ME): La inmigración a comunidades locales tiene como consecuencia que los *taxa* persistan incluso en ambientes no favorecidos.

iv) Modelo Neutro (*Neutral Model* o NM): Retoma planteamientos de la teoría neutra, asumiendo tasas de mortalidad, dispersión y especiación equivalentes entre todas las especies (Zheng *et al.*, 2013). La neutralidad del modelo consiste en suponer que las especies no interactúan directamente, por lo cual se asume que no hay interacciones de competencia o respuestas al ambiente, y el azar es el que conduce estos procesos. Ver apartado 1.5 para más detalles sobre la teoría neutra.

1.4 Una síntesis conceptual: cuatro procesos propuestos, cuatro modelos y cuatro aproximaciones en metacomunidades

Una propuesta que resulta interesante, es la planteada por Vellend (2010) y retomada por Nemergut *et al.* (2013) como una forma de utilizar sólo cuatro categorías como representantes de todos los procesos que se han definido para estructurar a las comunidades: selección, deriva, diversificación y dispersión. En la Tabla 1 se describen los cuatro procesos, que fueron tomados de la genética de poblaciones y de la ecología evolutiva, y la propuesta es utilizarlos tanto para

entender, como para describir los distintos modelos planteados por la teoría de metacomunidades, como se presenta en la Tabla 2. En esta tabla se incluyen los cuatro modelos o hipótesis que Martiny *et al.* (2006) propusieron, y que han sido ampliamente utilizados en la literatura para determinar la asociación de las comunidades microbianas a factores ambientales, a factores geográficos (un término usado frecuentemente como equivalente de distancia geográfica, lo cual refleja la dispersión de los organismos) o ambos.¹

Tabla 1. Procesos que estructuran a las comunidades (Nemergut *et al.*, 2013)

Proceso	Definición
Selección	Cambios en la estructura de comunidades causados por diferencias de adaptación entre los <i>taxa</i>
Deriva	Cambios estocásticos en la abundancia relativa de los <i>taxa</i> en la comunidades a través del tiempo
Diversificación	Generación de nueva variación genética
Dispersión	Movimiento de organismos a través del espacio

Estas cuatro alternativas de distribución de las comunidades, a su vez pueden entenderse y tienen equivalencia con los supuestos de metacomunidades, como se indica en la Tabla 2.

¹ A lo largo de la tesis, el término factores ambientales se utilizará para referirse a factores abióticos (a menos que se especifique), y el término factores neutros o geográficos se referirá a límites de dispersión o efecto de masas.

Tabla 2. Cuatro hipótesis de trabajo para explicar la distribución de los microorganismos en las comunidades (en cuanto a factores geográficos (G) o ambientales (A)), y sus equivalencias con cuatro supuestos y cuatro procesos

Alternativas para la distribución de las comunidades (Martiny <i>et al.</i> , 2006)		Equivalencia con los supuestos de metacomunidades	Procesos asociados: deriva, dispersión, selección y diversificación (Vellend, 2010; Nemergut <i>et al.</i> , 2013)
Hipótesis 1	Ninguna	Modelo neutro sin límites de dispersión, lo que equivale a un modelo nulo (propuesta de este trabajo)	Deriva, diversificación equivalente, dispersión sin límites
Hipótesis 2	A		Selección, dispersión sin límites “Todo está en todos lados, pero el ambiente selecciona” (Baas Becking, 1934; Finlay, 2002)
		Repartición de especies (asume dispersión moderada) (Ng <i>et al.</i> , 2008)	Selección, dispersión moderada
Hipótesis 3	G	Modelo neutro (Hubbell, 2001)	Deriva; diversificación y dispersión equivalentes
		Dinámica de parches (Cottenie, 2005)	Deriva, especiación y dispersión en parches equivalentes
Hipótesis 4	A y G	Efecto de masas (Baja dispersión) + Repartición de especies (Ng <i>et al.</i> , 2008)	Selección, deriva, diversificación, poca dispersión
		Efecto de masas (Alta dispersión) + Repartición de especies (Ng <i>et al.</i> , 2008)	Selección, deriva, diversificación, mucha dispersión

1.5 Confusiones sobre algunos conceptos o términos

El uso de terminología “inconsistente y a veces redundante” (Nemergut *et al.*, 2013) sobre todo en publicaciones que analizan el proceso de ensamble de comunidades microbianas, vuelve necesario mencionar algunos ejemplos para aclarar algunas confusiones comunes en la literatura.

La primera de ellas tiene que ver con el uso de los términos “local” y “regional” equivalentes en muchos textos a “enfoque ecológico” o “enfoque

biogeográfico”, respectivamente. La escala local, o *alfa*, se ha reportado como influenciada principalmente por factores ambientales (bióticos y abióticos, característicos de cada lugar), a diferencia de la escala regional, relacionada con la diversidad total, influenciada principalmente por eventos históricos (Rodríguez & Vázquez-Domínguez, 2003).

De esta forma el término “factores locales” se ha vuelto sinónimo de factores abióticos, con lo cual se asume que en escalas locales el proceso estructurador de las comunidades es únicamente la selección, lo cual no es correcto, ya que en esta escala también puede haber influencia de otros procesos, así como efecto de las interacciones bióticas. Lo mismo ocurre para el término “factores regionales”, que se utiliza como sinónimo de algunos procesos asociados a escala regional como la deriva, pero que no necesariamente siempre ocurren a escala regional. En este sentido, Martiny y colaboradores (2011) describen asociación de las comunidades de Nitrosomonadales con distancia geográfica, y esta asociación la encuentran únicamente a escala local y no para escala regional, por lo que proponen que los límites de dispersión, en escala ecológica o local, podrían facultar que también ocurra deriva a esta escala. Por esta razón, es importante no utilizar los términos factores locales o regionales como sinónimos de selección o deriva respectivamente.

Por otro lado, y en este mismo sentido, el término “procesos históricos” también se ha utilizado como sinónimo de factores relacionados con la historia geográfica, por ejemplo la deriva o los límites de dispersión. Anderson y colaboradores (Andersson *et al.*, 2014) resaltan que los procesos históricos, al ser eventos evolutivos y ecológicos que ocurren en el pasado, también involucran –además de límites de dispersión o deriva– selección por condiciones ambientales pasadas, un fenómeno que no ha sido muy estudiado en comunidades bacterianas. Por lo tanto debe evitarse utilizar el término “histórico” como sinónimo de factores geográficos o de aquéllos asociados con la historia geológica de las comunidades, teniendo en mente que es un término que abarca

cualquier proceso pasado que haya tenido efecto en la estructura de las comunidades².

Por otra parte, es muy común encontrar referencias hacia la teoría neutra como “teoría estocástica” o “teoría de los ensambles por dispersión”, aunque ni la dispersión o la estocasticidad son características necesariamente neutras (Harpole, 2012). La neutralidad del modelo consiste en suponer que las especies no interactúan directamente y que la distribución de la diversidad no depende de competencia o respuestas al ambiente. El presuponer que la dispersión y la competitividad son equivalentes es lo que fundamenta la neutralidad de esta teoría, y el azar es el que conduce estos procesos definidos como neutros, *desde el supuesto de la equivalencia*. Se afirma frecuentemente que la dispersión o la competitividad son al azar o estocásticas, cuando lo importante es que el azar operará a partir de la equivalencia de estas características.

Por último, como una forma de evitar confusiones en el uso de conceptos frecuentes para el campo de estudio (e.g. factor vs. proceso), se elaboró la Tabla 3, en la que se muestran ejemplos de la forma en que se nombrarán algunos conceptos y términos en este trabajo, para evitar confusiones y que la terminología sea clara.

Tabla 3. Ejemplos de conceptos y términos relevantes para la ecología de comunidades usados en este trabajo

Concepto	Ejemplos
Factores y mecanismos	Competencia, coexistencia, depredación, asociados a nicho, factores neutros, factores ambientales
Patrones	Decaimiento por distancia, especies-área, gradiente latitudinal

² El término estructura de comunidades en este trabajo se refiere a la abundancia y distribución de especies.

Procesos	Selección, deriva, diversificación, dispersión
Perspectivas o aproximaciones en metacomunidades	Modelo neutro, repartición de especies, dinámica de parches, efecto de masas

2 Algunos patrones de distribución relacionados con este trabajo

La biología ha reconocido patrones de distribución de comunidades desde inicios del siglo XIX (*e.g.* gradiente latitudinal, propuesto por Von Humboldt en 1799), pero no es hasta la década de los 60s que los ecólogos comenzaron a interpretarlos, cuando los procesos de poblaciones se integraron a la disciplina de ecología de comunidades (Ricklefs, 2004), y la variación en el recambio de especies o diversidad *beta* se comenzó a asociar con factores locales, regionales o con las propiedades emergentes derivadas de las interacciones en la comunidad.

Algunos de estos patrones de distribución, en particular los relacionados con este trabajo, se describirán brevemente en los siguientes apartados.

2.1 Decaimiento por distancia

Un patrón importante que propone Whittaker (1960), derivado del concepto de diversidad *beta*, es el llamado “decaimiento de la similitud con la distancia”, el cual se ha encontrado en la mayoría de las comunidades de macroorganismos (Calderón-Patrón *et al.*, 2012): las comunidades se parecen cada vez menos mientras mayor es la distancia que las separa. Se ha reportado que este patrón está controlado conjuntamente por procesos relacionados con factores ambientales (físicos y bióticos), configuración espacial y procesos neutros (Nekola & White, 1999; Soininen *et al.*, 2007), aunque los factores ecológicos que generan este patrón no se han entendido del todo (Nemergut *et al.*, 2013) y se han encontrado otros mecanismos que lo generan, como por ejemplo perturbaciones y comportamiento animal (Morlon *et al.*, 2008). Nekola y White

(1999) reportan que esta relación es más fuerte (la pendiente es más pronunciada) en organismos con menor habilidad de dispersión.

2.2 Invarianza regional

Esta propuesta de estructura de la comunidad se deriva del modelo neutro y de los trabajos clásicos de biogeografía de islas (MacArthur & Wilson, 1967): la abundancia relativa de una especie dentro de una comunidad (isla) depende de su abundancia relativa en la fuente de colonización (poza regional o metacomunidad) desde la que migra. Este patrón se espera en comunidades que se encuentran bajo procesos neutros (Sloan *et al.*, 2006), aunque existen estudios que reportan comunidades que se no ajustan a este patrón, a pesar de encontrarse dentro de los supuestos que se plantean. Por ejemplo, Langenheder & Székely (2011) en su experimento con microcosmos hecho en comunidades acuáticas de bacterias que habitan en pequeños huecos de rocas, encuentran que en algunos casos la estructura de estas comunidades es similar a la del agua de lluvia (fuente de colonización), pero en otros casos no lo es, y concluyen que la importancia de este proceso depende de la cantidad de generalistas o especialistas en la comunidad. Al respecto, se ha reportado que otros factores como es el orden de colonización (Gleason, 1927) y el tamaño del hábitat (Woodcock *et al.*, 2007) pueden resultar en comunidades divergentes a la de la fuente de colonización.

2.3 Patrones de aislamiento geográfico, ambiente físico e interacciones bióticas en un contexto histórico

“Los debates filosóficos que han acompañado el desarrollo de la ciencia frecuentemente han sido expresados en términos de dicotomías entre puntos de vista sobre la estructura de la naturaleza, la explicación de procesos naturales, y los métodos apropiados para investigar” (Levins & Lewontin 1985). Un ejemplo clásico de una dicotomía en ecología de comunidades ha sido el debate entre el planteamiento determinista de Clements (1916, 1936), con una visión de las comunidades como “superorganismos”, y en el otro extremo el punto de vista

reduccionista y azaroso, que reconoce sólo el nivel de individuo y niega a las comunidades, planteado por Gleason (1927), que aseguraba que la distribución de las especies respondía individualmente a factores ambientales. El planteamiento de Gleason se hace en el contexto de las ideas de Descartes, en las que el todo es la suma de sus partes, y en este sentido se considera que son las propiedades individuales las que determinan el comportamiento del sistema. Estas ideas denegaron durante mucho tiempo la posibilidad de propiedades emergentes en las comunidades a consecuencia de la interacción de sus partes (Valiente B. *et al.*, 2014), esto es, no reconocían la importancia de las interacciones bióticas en la estructuración de las comunidades.

Hoy en día esta dicotomía aún se cita y se discute sobre todo en la literatura que estudia comunidades microbianas (e.g. Schmidt *et al.*, 2014), expresada como procesos determinísticos vs. procesos estocásticos, o factores ambientales (nicho) vs. factores neutros.

El fondo de la discusión entre Clements y Gleason, es el reconocimiento (o negación) de ese “ente organísmico” que Clements llamó comunidad. Era un planteamiento conceptual muy distinto a lo que hoy se reconoce como una comunidad, pero fue el planteamiento formal que permitió que las comunidades se volvieran campo de estudio, y también permitió que se reconociera el valor del papel de las interacciones en las comunidades. Por su parte Gleason propone que hay un mayor papel del azar y las especies se distribuyen de acuerdo a sus propiedades individuales, lo que produce que una respuesta a factores abióticos. Es importante resaltar que la dicotomía planteada en la literatura actualmente no debate entre la existencia o no de comunidades. Es en todo caso una mezcla entre los planteamientos de Clements y Gleason, ya que el debate actual se centra entre el papel que tiene el azar en la estructura de las comunidades (Gleason) vs. el papel de los factores ambientales, lo que implica factores bióticos / interacciones (Clements) o factores abióticos (reconocidos por Gleason). Por esta razón los nombres de Clements y Gleason deberían mencionarse como un viejo debate clásico, interesante desde el punto de vista

histórico, pero se sugiere no utilizar estos nombres como representantes de este otro debate actual entre factores neutros vs. factores de nicho.

No es tarea simple distinguir estos dos factores (neutro vs. nicho) (Ricklefs, 2004), entre otras razones porque frecuentemente la distancia ambiental (referida en la gran mayor parte de los reportes de comunidades microbianas como factores abióticos) y la geográfica están intercorrelacionadas (Gilbert & Lechowicz, 2004), o porque fuertes gradientes ambientales favorecen un filtro ambiental sobre dinámicas neutras (Östman *et al.*, 2010; Östman *et al.*, 2012). Actualmente se propone un nuevo enfoque que plantea sintetizar ambas perspectivas (Leibold & McPeck, 2006; Ricklefs, 2004), el cual propone que las comunidades se sitúan entre algún lugar entre estos dos extremos teóricos (Gravel *et al.*, 2006; Thompson & Townsend, 2006).

Dicotomías como las mencionadas, históricas o actuales, han sido consecuencia de las filosofías dominantes y reduccionistas que han predominado en la ciencia de los países orientales, y producen que el estudio de sistemas complejos, como es el caso de las comunidades, no sea el adecuado (Levins y Lewontin 1985). Los enfoques que sugieren un punto intermedio para los extremos teóricos, apoyan la idea de la “existencia de muchas posibles configuraciones de poblaciones que preservan las mismas propiedades cualitativas a nivel de un todo”. Esto produce una visión más allá de un extremo u otro, planteando un gradiente o combinación de posibilidades. Por ejemplo Levins y Lewontin (1985) afirman que “los procesos estocásticos podrían ser las bases de los procesos determinísticos, y los determinísticos las bases de los estocásticos, no se excluyen unos a otros”.

3 Patrones en comunidades microbianas ¿qué se sabe?

3.1 Teorías de macroorganismos aplicadas a procariontes

Debido a su tamaño, los microorganismos sólo pueden ser estudiados utilizando herramientas específicas. Por ejemplo, el microscopio ha sido la herramienta clásica que desde hace más de un siglo permitió aproximarse a su morfología.

Hoy en día las técnicas moleculares que continúan perfeccionándose han sido claves para analizar su metabolismo y rasgos genéticos (Green *et al.*, 2008; Williamson & Yooseph, 2012; Zarraonaindia *et al.*, 2013; Eguiarte *et al.*, 2013), datos que han sido útiles para describir la distribución y diversidad microbiana.

Algunos autores aseguran que “no existe teoría en ecología microbiana” (Prosser *et al.*, 2007), o se argumenta que, a pesar de la revolución darwinista, los avances de la ecología microbiana son equivalentes a los de la ecología de los macroorganismos antes de que Darwin zarpara en el Beagle. Por ejemplo, los microbiólogos aún no pueden acordar cuántas especies de bacterias existen: se calculan estimados con desviaciones entre ellos de hasta 5 a 6 órdenes de magnitud (Torsvik & Øvreås, 2002; Lee, 2014). Estas afirmaciones se deben a que el desarrollo reciente de esta área (ca. 25 años) no ha permitido aún que la información sobre estos organismos sea tan abundante como es hoy en día la descripción de la historia natural de los macroorganismos.

Sin embargo, existe información generada a partir de algunos microorganismos modelo, por ejemplo en cuanto a las relaciones entre genotipo y fenotipo así como la comprensión profunda de los mecanismos moleculares detrás de su metabolismo. Esta información ha llevado a algunos investigadores a generalizar a todas las bacterias a partir de unos cuantos ejemplos muy conocidos. Por ejemplo, muchos autores afirman que el tamaño pequeño de los microorganismos, así como su enorme tamaño poblacional y sus rápidas tasas de crecimiento (como es el caso de *Escherichia coli*) implican necesariamente una gran dispersión (eg. Soininen, 2014) , aunque otros estudios descartan esta idea (Zinger *et al.*, 2014), ya que reportan que la pendiente de las curvas taxa-área (TAR) es la misma que para organismos grandes, lo que significa que la capacidad de dispersión de los microorganismos es limitada, a pesar de su tamaño. Otra similitud de los microorganismos con las comunidades ecológicas en general, es la distribución de las abundancias (Lee, 2014), con dominancia de pocas especies, y a su vez con un gran número de especies que están representadas por pocos individuos (Pedrós-Alió, 2012).

La comprensión de los procesos detrás de los patrones biogeográficos ha resultado una tarea compleja, incluso para macroorganismos (Ricklefs, 2004). Nemergut *et al.* (2013) plantea la necesidad de reconocer y definir características particulares de los microorganismos que permitan plantear hipótesis y alternativas que se ajusten a distribución. El tamaño pequeño no es una característica útil que permita inferir su distribución, ya que no resulta informativa *per se* (Nemergut *et al.* 2013). En cambio lo que sí se sabe y se ha estudiado, es que los eventos de dispersión activa en microorganismos son raros, por lo que, si la forma de dispersión en los microorganismos es pasiva, resulta poco probable que resulte en múltiples eventos a grandes distancias. Latencia y plasticidad fenotípica, así como la relación entre caracteres y taxonomía son características que también pueden resultar útiles para entender su distribución (Nemergut *et al.*, 2013).

En resumen, podemos concluir que las particularidades de los microorganismos no los excluyen de las teorías ecológicas. Mientras mejor se defina y se conozca la ecología de los microbios, sus patrones de distribución espacial podrán asociarse con mayor certeza a los procesos que los estructuran.

3.2 De dicotomías y formas de aproximarse a los mecanismos que estructuran a las comunidades microbianas

Aún siendo una disciplina muy joven, no han faltado dicotomías para la ecología microbiana. Hace unos 20 años, los ecólogos microbianos se dividieron en aquéllos que aseguran grandes tasas de dispersión en los microorganismos (como se describe en el apartado anterior), lo cual implica asumir que la dispersión de los microorganismos es al azar, y se aseguraba que el ambiente era el responsable de la estructura de sus comunidades (abiótico en la mayoría de los reportes), lo que corresponde a la hipótesis 2 de la Tabla 2. Este planteamiento se basó en la citadísima frase que BaasbeKing (1934) declaró a inicios del siglo pasado: “todo está en todos lados pero la naturaleza selecciona”, una idea descartesiana similar a las ideas de Gleason (ver apartado 2.3) en contexto con los planteamientos de esa época. Finlay (2002) es uno de los

representantes modernos y defensor de ese planteamiento. Por otro lado algunos grupos de trabajo se dieron a la tarea de demostrar límites de dispersión para microorganismos, por lo que se utilizaron herramientas de la teoría neutra que permitieron probarlo, por ejemplo con reportes de decaimiento por distancia (e.g. Martiny *et al.*, 2011, Zinger *et al.*, 2014), correspondiente a la hipótesis 3 de la Tabla 2, y también por medio de patrones *taxa-área* (e.g. Bell *et al.*, 2005, Noguez *et al.*, 2005). La dificultad para distinguir un factor de otro (factores abióticos vs. distancia geográfica) ya se ha mencionado previamente (apartado 2.3), y hoy en día se asegura con mayor frecuencia en la literatura, que ambos factores pueden estructurar a la vez a las comunidades bacterianas. Esta aproximación para estudiar a las comunidades, por medio de la asociación a factores abióticos, distancia genética o ambos (Tabla 2), es la que ha predominado en la literatura en los últimos años, y es la manera en la que se ha trabajado en esta tesis.

En este sentido, el término “biogeográfico” (equivalente a regional) ha sido ampliamente utilizado en muchos artículos que estudian a las comunidades microbianas, para referirse a la manera en la que se distribuyen las bacterias, haciendo referencia a que no se toma en cuenta la ecología (interacciones bióticas) en esta distribución. Pero muchos de estos “patrones biogeográficos” descritos, ocurren a escalas locales, y es en esta escala que se estudia la asociación a factores abióticos o a distancia genética como se comentó en el párrafo anterior.

¿Pero qué sucede entonces con las interacciones bióticas y cuál es el efecto que tienen en las comunidades de microbios? Se ha asumido que las bacterias y otros microorganismos tienen vidas unicelulares independientes, sin el comportamiento cooperativo que se ha descrito para mamíferos, pájaros o insectos, pero esta idea ha cambiado en años recientes, y se ha descubierto que los microorganismos también presentan comportamiento social en complejos sistemas de cooperación, comunicación y sincronización (para una revisión, ver West 2007), una historia que recuerda a la ya descrita dicotomía entre Clements y Gleason (apartado 2.3). El papel de los microorganismos en los ecosistemas

está empezando a comprenderse, su participación en ciclos biogeoquímicos y su aportación en la productividad de suelos o en el mar (Horner-Devine *et al.*, 2004, van der Heijden *et al.*, 2008) ejemplifican que las interacciones entre ellos y con macroorganismos son muy frecuentes (West 2007).

Explorar las interacciones de microorganismos no es una tarea fácil. La forma más común de hacerlo es a través del cultivo de bacterias, ya que este método permite caracterizarlas en forma precisa en el laboratorio, y es posible generar redes de interacciones a partir de aquéllos rasgos que se han determinado con certeza (*eg.* Aguirre-von-Wobeser *et al.*, 2014). Pero las bacterias de cultivo representan una pequeña fracción de la diversidad total, porque son tan sólo el 1% es cultivable (Vartoukian *et al.*, 2010). Así, el estudio del material genético obtenido directamente del ambiente (metagenómica), además de la secuenciación de un número cada vez mayor de genomas bacterianos, ha permitido una nueva aproximación a la diversidad y las funciones de las comunidades de bacterias. En la literatura ya existen algunos reportes que plantean analizar la estructura de estas comunidades a nivel funcional (*e.g.* Goberna *et al.*, 2014, Barberán *et al.*, 2014), por medio de inferencias a partir del gen 16S ADN_r, bajo el supuesto de que las funciones en microorganismos (aquéllas que son más complejas) se encuentran conservadas para ciertos *taxa* (Martiny *et al.*, 2013). De esta forma, por medio de comparaciones con los genomas microbianos que están disponibles en las bases de datos mundiales, se infiere la presencia o ausencia de caracteres en las comunidades estudiadas. La interpretación de este tipo de datos resulta controversial, ya que por un lado las tasas de transferencia horizontal de genes en bacterias (para una revisión, ver Boto *et al.*, 2010) se ha reportado que es muy alta, aunque sobre todo y con mayor frecuencia, entre linajes o *taxa* similares, y por otra parte no todos los atributos funcionales se encuentran conservados filogenéticamente, por ejemplo aquellos atributos que no son complejos no mantienen este patrón (Martiny *et al.*, 2013). En un futuro, cuando sea posible ensamblar la mayoría de los genomas que hay en una muestra de ADN ambiental, podrán determinarse en forma directa los caracteres que existen

en las comunidades. Por lo pronto, la forma de aproximarse a la estructuración de las comunidades bacterianas desde las interacciones bióticas, tendrá que continuar realizándose de manera indirecta.

3.3 Sobre patrones de distribución, *taxa* y generalizaciones

Existen algunas características de los microorganismos, reconocidas y mencionadas en la literatura como determinantes en la definición de patrones, como lo es la baja abundancia, que define a los llamados *taxa* “raros”. Se ha reportado que estos *taxa* tienen actividades y funciones metabólicas características y/o particulares (Pedrós-Alió, 2012), cuyos patrones de distribución son diferentes al compararlos con los de los *taxa* abundantes (Galand *et al.*, 2009). En relación con la historia de vida, los *taxa* son generalmente clasificados en especialistas y generalistas. Frecuentemente se menciona en la literatura que los *taxa* generalistas son dominantes en las comunidades (Székely & Langenheder, 2014), y los especialistas son *taxa* raros. Además se ha sugerido que los *taxa* generalistas presentan una distribución neutra, dado que son *taxas* que se ensamblan por mecanismos relacionados a la dispersión, debido a su capacidad de ocurrir en un amplio rango de hábitats, mientras que los especialistas, restringidos a ciertos hábitats, estarán asociados principalmente a factores abióticos (Pandit *et al.*, 2009). En este mismo sentido, Lepère *et al.* (2013) plantean que los *taxa* raros también son susceptibles a límites de dispersión, porque la probabilidad de que un propágulo se disperse es reducida, ya que la posibilidad de colonización y establecimiento de este tipo de *taxa* es muy baja. Por otra parte, Logares *et al.* (2013) plantean lo contrario, argumentando que los *taxa* raros se apartan del modelo nulo porque tienen crecimiento lento y son *taxa* latentes, lo cual conlleva a no cumplir con la condición neutra de tener crecimiento equivalente para todos los *taxa*.

Estas dos últimas generalizaciones sobre los *taxa* poco abundantes, describiéndolos como de lento crecimiento o malos dispersores, se basa en suposiciones vagas, lo que lleva a conclusiones contradictorias. Evidencia experimental en cepas silvestres aún dentro de organismos modelo como

Bacillus (Valdivia *et al.*, sometido) demuestran que cada genotipo, incluso dentro de una misma especie, tiene diferentes estrategias de crecimiento. Por otra parte, la dispersión no sólo depende de la capacidad de movimiento intrínseca de cada especie, sino de la capacidad particular de cada cepa migrante para adaptarse a nuevas condiciones. De lo anterior se desprende que puede haber organismos especialistas asociados con una alta abundancia o generalistas con baja dominancia en las comunidades microbianas.

3.4 Particularidades de los patrones de distribución y diversidad encontrados en microorganismos

En cuanto a los factores ambientales o de distancia, parece ser que las comunidades bacterianas están gobernadas principalmente por factores ambientales (Wang *et al.*, 2013), porque son más frecuentemente detectados, con un 94% de las publicaciones analizadas, contra un 69% para efecto de distancia (Hanson *et al.*, 2012). Por otra parte, en estudios que analizan mesocosmos, se ha encontrado que los factores neutros son importantes, sobre todo al inicio de los ensamblajes, por ejemplo en las comunidades acuáticas bacterianas descritas previamente, que habitan en pequeños huecos de rocas, las cuales son similares en etapas iniciales a la fuente que las coloniza (agua de lluvia) (Langenheder & Székely, 2011), o las comunidades bacterianas en el suelo 4 semanas después de un incendio, que tienen menores desviaciones nulas que aquéllas en las que el suelo no sufrió alteraciones (Ferrenberg *et al.*, 2013).

Respecto a las escalas, las temporales son de gran importancia. Por ejemplo, se ha reportado que una heterogeneidad ambiental que es estable en el tiempo determina que las comunidades presenten fuertes procesos de selección, ya que esa estabilidad temporal permite que las comunidades se adapten a los distintos ambientes (Rainey & Travisano, 1998). En cambio, si la heterogeneidad ambiental está dada por modificaciones en una temporalidad corta, esto genera comunidades con *taxa* adaptados a estas grandes variaciones (generalistas) y con predominancia de procesos neutros, por

ejemplo en los suelos, recientemente descritos por Barberán *et al.* (2014) o en la superficie del mar (Ghiglione *et al.*, 2012), que presentan comunidades compuestas principalmente por *taxa* generalistas.

Por otra parte, los patrones no han resultado tan claros para el caso de la escala espacial. Hanson *et al.* (2012), encuentra que patrones biogeográficos estructurados por factores neutros han sido reportados tanto a escalas grandes (más de 5000 km) como pequeñas (0-1 km). Martiny *et al.* (2011) también ha reportado efecto geográfico a escala local, argumentando que a este nivel las bacterias tienen dispersión limitada, por lo que el efecto geográfico se hace evidente. Pero a escalas mayores estos autores han encontrado que los factores ambientales son los que determinan la estructura de las comunidades, aunque no descartan que ambos factores, tanto neutros como ambientales, pueden operar en todas las escalas.

Respecto a la amplitud de la distribución de los *taxa*, Barberán *et al.* (2014) reporta que genomas grandes, que son los que tienen mayor versatilidad metabólica –esto es, mayor potencial genómico funcional– presentan distribuciones ambientales y geográficas más amplias. Asimismo, estos autores encuentran que *taxa* con intervalos de distribución similares pueden no compartir historias evolutivas o filogenéticas, pero sí atributos funcionales. Barberán *et al.* (2014) sugieren que estas comunidades de grandes genomas versátiles tienen mayores tasas de transferencia horizontal, lo que produce que los atributos funcionales no se encuentren conservados filogenéticamente.

Estos atributos funcionales han sido también explorados como un mecanismo determinante en los patrones de distribución bacterianos, y se ha propuesto que un marco teórico basado en atributos puede usarse para entender ensamblajes de comunidades (Burke *et al.*, 2011; Barberán *et al.*, 2012) y para explicar cambios de composición en gradientes ambientales (Edwards *et al.*, 2013). Respecto a la distribución de estos caracteres en la filogenia de los microorganismos, Martiny *et al.* (2013) reportan que caracteres complejos (e.g. fotosíntesis o metanogénesis) se distribuyen en pocos *clusters* filogenéticos asociados a largas ramas en la filogenia, lo que refleja *taxa* diferenciados en

mayor tiempo evolutivo, mientras que caracteres fisiológicos adaptativos más simples, como el uso de sustratos de carbono, se encuentran filogenéticamente dispersos, tal como sucede en las comunidades de suelo de Barberán *et al.* (2014), descritas en el párrafo anterior. Wang *et al.* (2013) por otra parte, encuentra *clusters* filogenéticos (en este estudio se reportan los *clusters* en ramas cortas, que corresponden a tiempos evolutivos recientes) asociados a nichos ecológicos, por lo que concluye que esta asociación a hábitat ha emergido a través de adaptaciones evolutivas de ciertos *taxa* en las comunidades microbianas.

4 Diversidad y biogeografía de sistemas acuáticos bacterianos

Los ecosistemas acuáticos ocupan más del 70% de la superficie de la tierra, y están dominados por microorganismos planctónicos que los habitan (para una revisión detallada, ver Zinger, Gobet & Pommier, 2012). Los ecólogos microbianos especializados en estos sistemas contaban, hasta hace muy poco tiempo, con escasa información sobre los procesos y factores que determinan la estructura de estas comunidades acuáticas (Jones *et al.*, 2012). Dada la importancia de estas comunidades bacterianas en la transformación de materia orgánica, flujos de energía y ciclos biogeoquímicos (Brown *et al.*, 2014) (por mencionar algunos ejemplos), recientemente se ha comenzado a publicar cada vez más información al respecto.

Sobre la distribución y la estructura de las comunidades en el mar, se sabe que es compleja y está determinada entre otros factores, por corrientes de agua sujetas a fenómenos como efecto coriolis, viento, gravedad, y calor del sol, o en el caso de la zona profunda, por la circulación de termoclinas (para una revisión consultar Brown *et al.*, 2014). Todos estos factores producen desde corrientes que pueden transportar a las bacterias a grandes distancias hasta fuertes cambios ambientales. Se han observado patrones de diversidad beta para bacterias marinas, aunque no siempre son equivalentes a los patrones que existen para estudios macroecológicos (Brown *et al.*, 2014). Por ejemplo, se han reportado patrones cíclicos estacionales (Jones *et al.*, 2012), también una fuerte

asociación entre variación ambiental y diversidad beta (Hatosy *et al.*, 2013), pero además se describen patrones de decaimiento por distancia (Jones *et al.*, 2012; Zinger *et al.*, 2014). La estructura de las comunidades en la superficie es distinta a las comunidades de muestras profundas, con mucho recambio en la superficie y menor diversidad filogenética (ramas más cortas en la filogenia), debido a que las condiciones ambientales tan cambiantes en poco tiempo no producen mucha divergencia entre las especies. Las poblaciones que se encuentran a grandes profundidades tienen mayor diversidad filogenética, porque las condiciones ambientales son más estables, y permiten que el tiempo para adaptarse a ellas sea mayor (Ghiglione *et al.*, 2012).

Los lagos y estanques, por otra parte, son particularmente útiles para realizar estudios macroecológicos, debido a su naturaleza discreta y su estructura geográfica estable (Hortal *et al.*, 2014). Estos hábitats son equivalentes a islas en un mar de tierra, y al igual que una verdadera isla, poseen tamaño, forma, diversidad e historia particulares. Factores ambientales y de distancia han sido estudiados y reportados como estructuradores de comunidades en estos sistemas (Lindström & Langenheder, 2012). Östman *et al.* (2012) han sugerido que la importancia relativa de estos dos factores en estos hábitats depende de la heterogeneidad ambiental entre los sitios estudiados, y que una gran variación ambiental promoverá que factores ambientales sean los que determinen la estructura de las comunidades. Estos autores reportan que las comunidades acuáticas de lagos en Suecia mostraron una asociación significativa entre el grado de invarianza regional y la variación entre variables ambientales.

Los datos metagenómicos han permitido una comparación más precisa entre las comunidades acuáticas. Se ha determinado que las comunidades del mar, comparadas con las de lagos, sí comparten algunas funciones, pero existen grandes diferencias. Estas diferencias se deben a adaptaciones muy específicas para cada tipo de hábitat (Eiler *et al.*, 2014), como genes asociados a la degradación de polisacáridos de las plantas presentes en los sistemas terrestres, para el caso de los lagos, o a la degradación de material proteico del

plancton, en el caso del mar. Respecto al promedio del tamaño efectivo del genoma (Eiler *et al.*, 2014), se ha definido un mayor tamaño en lagos, de 1.91 mega (millones) pares de bases, y de 1.33 mega pares de bases para sistemas marinos, posiblemente porque el tamaño pequeño permite una mejor adaptación a sistemas oligotróficos (Giovannoni *et al.*, 2005, Grote *et al.*, 2012). Los metagenomas marinos son los que poseen más genes metabólicos (comparados con genomas de bacterias de lagos), lo que probablemente se debe, como Ghiglione *et al.*, (2012) lo reporta, a que en la superficie del mar habitan bacterias adaptadas a sobrevivir a los fuertes cambios ambientales que se dan en breves lapsos de tiempo, por lo que son capaces de ocupar hábitats más amplios (para lo que se requiere mayor cantidad de genes metabólicos), similar a lo que se ha reportado para bacterias de suelo (Barberán *et al.*, 2014). En cambio los lagos constituyen sistemas cerrados y se han descrito como ambientes que pueden producir fuertes cuellos de botella en algunas *taxa*, así como menores tasas de recombinación comparados con *taxa* marinos (Zaremba-Niedzwiedzka *et al.*, 2013), reflejando una mayor adaptación a condiciones ambientales particulares.

5 Los sistemas acuáticos del desierto, el caso de Cuatro Ciénegas

Los sistemas acuáticos del desierto son poco comunes. Conocidos como “archipiélagos acuáticos”, son semejantes a las islas oceánicas que han permitido la sobrevivencia de plantas y animales no acuáticos, cuya colonización se ha debido a la dispersión de organismos de tierras adyacentes; a diferencia de la biota de las pozas en el desierto, ecosistemas en los que habitan organismos remanentes y relictos de otros tiempos (Unmack & Minckley, 2008).

El valle de Cuatro Ciénegas (CCC) es una área húmeda de gran importancia, que se ubica en el desierto Chihuahuense. Se encuentra delimitado por grandes cadenas montañosas formadas a partir de una compleja historia geológica (Souza *et al.*, 2006; Wolaver *et al.*, 2008). Esta área natural protegida alberga más de 70 especies de plantas y animales endémicas. Por otra parte, CCC posee un gran cantidad de sistemas acuáticos con condiciones

fisicoquímicas contrastantes (Elser *et al.*, 2005). El sistema hidrogeológico es complejo; se propone que los sistemas acuáticos más cercanos a la base de la montaña, con menor conductividad y mayor temperatura, descargan directamente de la montaña (la sierra de San Marcos y Pinos), y el resto de los sistemas están modificados por evapotranspiración, interacción con depósitos superficiales y/o enfriamiento (Johannesson *et al.*, 2004, Evans 2005). Se calcula que las aguas tardan más de 20 años en atravesar esta sierra, para emerger formando algunos de los numerosos cuerpos de agua del valle (Forti *et al.*, 2004), y más recientemente se reporta que el agua en estos sistemas es una mezcla de agua subterránea muy antigua (calculado que puede tener hasta 60 años de residencia) que proviene de fallas profundas, caracterizados por la presencia de sulfato de calcio (Wolaver *et al.*, 2013). Otros cuerpos de agua se componen de una mezcla de agua de lluvia con la del acuífero regional, y se identifican por tener bicarbonatos en lugar de sulfato de calcio (Wolaver *et al.*, 2013). Existe un gran número de canales de drenaje subterráneo que atraviesan los sedimentos del valle y son los que dan origen a estos sistemas acuáticos (Forti *et al.*, 2004), de los cuales no hay información precisa sobre la conexión subterránea entre ellos.

Es un sitio oligotrófico, en el que las concentraciones de fósforo (P) son bajas tanto en suelo como en agua (Elser *et al.*, 2005; Breitbart *et al.*, 2009; Peimbert *et al.*, 2012; López-Lozano *et al.*, 2012). En Pozas Azules por ejemplo, se reportan medidas de hasta 0.1 mg/L de fósforo total (Pajares *et al.*, 2012), o una relación N:P en promedio de 157:1 (Souza *et al.*, 2012) reportado en Río Mezquites (Elser *et al.*, 2005). A pesar de esta limitante, las comunidades bacterianas son muy abundantes y diversas, con numerosas especies endémicas (Souza *et al.*, 2006; Cerritos *et al.*, 2011; Escalante *et al.*, 2008). Estas bacterias presentan además una gran diversidad de funciones metabólicas (Breitbart *et al.*, 2009; Peimbert *et al.*, 2012; Bonilla-Rosso *et al.*, 2012).

En cuanto a la distribución de esta gran diversidad en el valle, las evidencias sugieren fuertes procesos de selección ambiental y adaptación local

en las poblaciones o comunidades de bacterias, ya que se ha reportado desde una alta especialización de genes en el género *Bacillus* (Alcaraz *et al.*, 2010), hasta asociación a factores ambientales como conductividad (una medida relacionada a la salinidad del agua) (Cerritos *et al.*, 2011), estacionalidad (Rodríguez-Verdugo *et al.*, 2012), hábitat (Rebollar *et al.*, 2012) o humedad (López-Lozano *et al.*, 2012). Se ha propuesto además que los niveles de transferencia horizontal de genes (HGT por sus siglas en inglés) son limitados (Souza *et al.*, 2008), por lo que la HGT no es suficiente como para romper el desequilibrio de ligamiento, ni por lo tanto, para alterar los patrones de diferenciación microbiana, siendo esta la explicación que se ha propuesto sobre la clonalidad de las poblaciones reportadas para distintas especies de bacterias (e.g. *Pseudomonas* (Escalante, 2008), *Exiguobacterium* (Rebollar *et al.*, 2012) o *Bacillus* (Cerritos *et al.*, 2011)); para estas tres especies se reportan evidencias históricas (cuellos de botella) detectadas en un mismo período de tiempo (Avitia *et al.*, 2014).

Por otro lado los análisis genómicos reportan un tamaño de genoma reducido en el género *Bacillus* (Alcaraz *et al.*, 2008) y en diversos phyla que coexisten en las comunidades de tapetes microbianos (Peimbert *et al.*, 2012), lo cual es similar a lo reportado para otras bacterias que habitan ambientes con poca cantidad de nutrientes, como son las bacterias en el mar (Eiler *et al.*, 2014), pero la baja proporción de genes para metabolismo secundario detectados para los genomas de *Bacillus* de CCC (Alcaraz *et al.*, 2010), sugiere que se trata de bacterias con poca amplitud de nicho, de acuerdo a la propuesta de Barberán *et al.* (2014), aunque esta propuesta no es más que una aproximación, ya que la adaptación a diferentes nichos no sólo depende de estos genes.

Por último, en cuanto a los estudios de las comunidades bacterianas del valle en el contexto de mesocosmos en suelo (López-Lozano *et al.*, 2012, 2013) y microcosmos en agua (Pajares *et al.*, 2012, 2013), ambos trabajos describen comunidades diversas de bacterias en las que los eventos de perturbación produjeron cambios drásticos en su composición, lo que refleja la fragilidad de estos sistemas.

OBJETIVOS

1. Describir la distribución y diversidad de comunidades de bacterias acuáticas, utilizando el gen que codifica para el 16S ARNr.
2. Presentar los alcances y limitaciones de tres métodos moleculares (T-RFLPs, biblioteca de clones y secuenciación *shotgun* de metagenoma).
3. Reflexionar sobre la estructuración de comunidades bacterianas acuáticas bajo distintas condiciones de muestreo (escala de muestreo local o regional, en condiciones abióticas similares o contrastantes).

CAPÍTULO I

ARTÍCULO PUBLICADO

Comparación de tres métodos moleculares para el análisis de procariontes ambientales en el mar del canal de Yucatán.

Hidrobiológica. **24**: 167-179

Comparación de tres métodos moleculares para el análisis de procariontes ambientales en el mar del canal de Yucatán

Analysis of three molecular methods to describe sea Yucatan channel environmental prokaryotic taxa

Laura Espinosa-Asuar,¹ Ana E. Escalante,²
Luisa I. Falcón,¹ Germán Bonilla-Rosso,¹
Santiago Ramírez-Barahona,¹ Luis E. Eguiarte¹
y Valeria Souza^{1,*}

¹ Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, AP 70-275, México 04510

² Departamento de Ecología de la Biodiversidad, Instituto de Ecología, Universidad Nacional Autónoma de México, AP 70-275, México 04510

e-mail: souza@unam.mx

Espinosa-Asuar L., A. E. Escalante, L. I. Falcón, G. Bonilla-Rosso, S. Ramírez-Barahona, L. E. Eguiarte y V. Souza. 2014. Comparación de tres métodos moleculares para el análisis de procariontes ambientales en el mar del canal de Yucatán. *Hidrobiológica* 24 (3): XX-XX.

RESUMEN

En este trabajo describimos la comunidad procarionte de una muestra de agua marina del canal de Yucatán. Para la determinación de la diversidad microbiana se usaron tres métodos moleculares: 1) T-RFLPs. 2) Secuenciación de amplicones (bibliotecas de clones). 3) Secuenciación *shotgun* de un metagenoma. Como un segundo objetivo, se presenta una comparación de los alcances y los límites de cada uno de estos tres métodos. Para esta comparación, se tomaron en cuenta tres criterios: el número de unidades taxonómicas detectadas, la precisión en la asignación taxonómica y el costo del estudio. Los *taxa* más abundantes fueron *Candidatus* Portiera OTU 3744 (equivalente al clado SAR86) y *Candidatus* Pelagibacter. Los resultados mostraron que la estrategia de secuenciación *shotgun* de ADN es la más poderosa en términos de unidades taxonómicas detectadas, mientras que los datos que se obtuvieron por T-RFLPs y con la biblioteca de clones representan sólo una submuestra de la biblioteca de fragmentos generados mediante *shotgun*. En cuanto a resolución filogenética (determinación taxonómica), la aproximación más precisa es la secuenciación de bibliotecas de clones. Los costos de las tres estrategias varían considerablemente, pero sus alcances también lo hacen. Por lo tanto, es importante tomar en consideración que algunas preguntas ecológicas y evolutivas sólo pueden ser contestadas específicamente por una u otra metodología.

Palabras clave: Clonación, diversidad procarionte, NGS, *shotgun*, T-RFLPs.

ABSTRACT

In this study we describe a prokaryotic community from a seawater sample obtained in the Yucatan Channel, using for this purpose three molecular methods: 1) T-RFLPs. 2) Sequencing of amplicons (clone libraries). 3) Metagenome *shotgun* sequencing; the three are useful for the determination of microbial diversity. We also present a comparison of the scope and limits of each method. The comparison took into account three criteria: the number of taxonomic units detected, the taxonomic assignment accuracy and the cost of the study. The most abundant taxa were *Candidatus* Portiera OTU 3744 (equivalent to SAR86 clade) and *Candidatus* Pelagibacter. The results showed that the *shotgun* se-

quencing strategy is the most powerful in terms of detected taxonomic units, while the data obtained by T-RFLPs and clone library methods represent only a subsample of the shotgun fragment library. Regarding phylogenetic resolution (taxonomical determination), the more accurate approach is the sequencing of clone libraries. The costs of the three strategies vary considerably, but so does its scope. Therefore, it is important to consider that one, or another methodology, can only specifically answer some ecological and evolutionary questions.

Key words: Cloning, prokaryotic diversity, NGS, *shotgun*, T-RFLPs.

INTRODUCCIÓN

La diversidad microbiana es una pieza de conocimiento indispensable para comprender el funcionamiento de los ecosistemas, ya que los microbios son un componente fundamental en los procesos biogeoquímicos (Fierer & Lennon, 2011; Nazaries *et al.* 2013). Sin embargo, este componente fue pasado por alto durante años en la literatura ecológica, debido en parte a la falta de herramientas técnicas que permitieran analizar y modelar el papel de los microorganismos en los ecosistemas (Hughes *et al.*, 2006). En las dos últimas décadas ha surgido una gran variedad de herramientas moleculares para el estudio de la composición de comunidades microbianas enteras (Green *et al.*, 2008; Segata *et al.*, 2013; Zarraonaindia *et al.* 2013). Estos avances tecnológicos han abierto la puerta al crecimiento sin precedentes de estudios ecológicos de los microorganismos en sus ambientes naturales (Allison *et al.* 2013; Bonilla-Rosso *et al.* 2008, 2012; Peimbert *et al.* 2012).

Los estudios modernos sobre diversidad microbiana inician con la extracción de ADN total de una muestra ambiental. Posteriormente puede trabajarse con técnicas que analizan los genomas completos de la muestra o con otras que se enfocan en la amplificación de un solo gen. Dentro de esta segunda opción existen marcadores moleculares con los que se analiza la diversidad de fragmentos de genes o de regiones específicas (*fingerprint*) o marcadores que se utilizan para analizar secuencias individuales de genes, que generalmente involucran la clonación de los amplicones. Por otro lado, están las técnicas que secuencian directamente el ADN total de la muestra, como es el caso de la secuenciación *shotgun* (secuenciación de fragmentos al azar, la única de este tipo que no precisa de amplificación de regiones particulares) o más recientemente, las técnicas conocidas como secuenciación de nueva generación (*Next Generation Sequencing* o NGS por sus siglas en inglés) (Eguarte *et al.* 2013; Williamson & Yooseph, 2012). En estas últimas (por ejemplo, pirosecuenciación 454, SOLiD o Illumina/Solexa), sí existe un paso de amplificación al enriquecer las bibliotecas; la amplificación puede hacerse en todos los fragmentos que se generan o también es posible amplificar regiones específicas.

América Latina es una de las regiones a nivel mundial con mayor biodiversidad en macroorganismos (CONABIO, 1988; Rzedowski, 2006). Sin embargo, de acuerdo con una revisión realizada en ISI Web of Knowledge (rango: 1980-2009, palabra clave:

microbial diversity), esta región del mundo ha publicado el menor número de reportes en comparación con EE UU, Europa y Asia. La revisión realizada también reflejó un incremento de publicaciones en todo el mundo en la última década analizada (2000-2009). Una probable causa de estas observaciones es que los análisis moleculares se han vuelto más accesibles en los últimos años, por lo que se ha hecho posible describir y estudiar la diversidad de microorganismos en cualquier lugar del mundo (para una revisión véase Lindstrom & Langenheder, 2012).

El objetivo de este trabajo es la evaluación de tres métodos moleculares: *fingerprint*, en particular Fragmentos Terminales de Restricción Polimórficos por Tamaño (T-RFLPs por sus siglas en inglés), biblioteca de clones derivadas de PCR y biblioteca de secuencias generadas por *shotgun*, utilizados para analizar la diversidad procarionte de una muestra ambiental de agua marina del canal de Yucatán (muestra GS017 reportada por Rusch *et al.* 2007). La composición taxonómica de la comunidad microbiana fue analizada aplicando el marcador del gen 16S ADNr. Los tres criterios de análisis para evaluar las metodologías empleadas fueron: 1) la comparación del número de unidades taxonómicas identificadas por cada método, 2) la precisión en la determinación de los grupos taxonómicos y 3) los costos para cada una de las metodologías.

MATERIALES Y MÉTODOS

Se analizó la composición microbiana en una muestra de agua marina (GS017) que fue tomada a 2 m de profundidad en el mar abierto del canal de Yucatán (20°31'21"N; 85°24'49"O), a unos 150 km al este de la isla de Cozumel, México (Rusch *et al.* 2007).

Extracción de ADN: El ADN se extrajo de una porción del filtro obtenido en la expedición del grupo de Rusch (Rusch *et al.*, 2007), utilizando el método reportado en ese mismo trabajo.

Métodos asociados a PCR (T-RFLPs y biblioteca de clones): El gen que codifica para el ARN ribosomal 16S (gen 16S ADNr) fue amplificado por PCR con los oligonucleótidos universales para bacterias y arqueas F515 y R1492 (Turner *et al.* 1999), cuyas secuencias son: F515 (5'-GCGGATCCTCTAGACTGCAGTGCCAGCAGCCGCGGTAA-3') R1492 (5'-GGCTCGAGCGCCGCCGGGTTACCTGTTCAGACTT-3').

Para la amplificación de T-RFLPs, el oligonucleótido F515 estaba marcado con fluorescencia VIC en el extremo 5' (Applied

Biosystems). Las condiciones de la reacción de PCR fueron las siguientes: *buffer* para PCR 1X, 1.65 mM de MgCl₂, 0.2 mM de dNTPs, 0.06 mM de cada oligonucleótido, 1 U de Taq polimerasa (Applied Biosystems, Foster City, CA) y 5 % de DMSO. Todas las reacciones se llevaron a cabo en un termociclador MJ Research (Watertown, MA) con los siguientes ciclos de reacción: 94 °C 4 min, 35 ciclos de 92 °C 1.5 min, 50 °C 1.5 min, 72 °C 2 min, completando con 72 °C 10 min. Para tener una mejor representación de la muestra, se hicieron tres reacciones de PCR por separado (tanto para clones como para T-RFLPs). El amplificado (aprox. 977 pb) de las tres reacciones se mezcló y purificó a partir de un gel de agarosa al 2 %, con el kit QIAquick (Qiagen Inc., Netherlands).

La técnica de T-RFLPs se fundamenta en una reacción de restricción, con la que se generan distintos fragmentos de ADN que componen un patrón característico (denominado "la huella" o *fingerprint*) de la comunidad analizada. El tamaño de cada fragmento es considerado como equivalente a una unidad taxonómica operativa (OTU, por sus siglas en inglés), y corresponde al sitio de restricción más cercano al oligonucleótido fluorescente que se utilizó en la reacción de PCR (Osborn *et al.* 2000). El producto obtenido por PCR (fragmento del gen 16S ADNr) fue digerido enzimáticamente con *Alu I* (Promega, Madison, WI) en reacciones de 20 µl. Para cada reacción se utilizaron 1 unidad de *Alu I* y 50 ng de producto de PCR. Las reacciones fueron incubadas en un termociclador MJ Research durante 3 h a 37 °C, seguidas de la inactivación de la enzima durante 30 min a 65 °C. Se realizaron tres corridas independientes de la muestra, usando un secuenciador automático ABI Avant 3100 PRISM® (Applied Biosystems, Foster City, CA). Con esto se obtuvieron los tamaños y alturas de los Fragmentos de Restricción Terminales (TRFs, por sus siglas en inglés), en forma de picos en un cromatograma. Para evitar fluorescencia inespecífica, sólo los picos que estuvieron presentes en las tres réplicas, y que presentaban alturas > 10 unidades de fluorescencia, fueron considerados en el análisis. Cada TRF obtenido en este análisis se consideró una OTU.

Biblioteca de clones: La mezcla de amplificados del gen 16S ADNr ya purificada se clonó en el vector plasmídico pCR2.1, de acuerdo con las instrucciones del fabricante (Invitrogen Corp., San Diego, CA). Los plásmidos se aislaron con el *kit* de extracción SNAP® Miniprep Kit (Invitrogen Corp.) y aquellos que contenían el inserto fueron secuenciados usando los oligonucleótidos M13F y M13R, en un secuenciador ABI Avant 3100 PRISM® (Applied Biosystems, Foster City, CA). Todas las secuencias fueron revisadas manualmente en BioEdit (Atschul *et al.* 1990). Para descartar quimeras (artefacto producido por la técnica de PCR, que origina fragmentos formados con secuencias de distinta procedencia), las secuencias obtenidas fueron analizadas con Bellerophon (Huber *et al.* 2004) y Chimera-Check en RDP (Cole *et al.* 2003). Quimera-Check es una herramienta que actualmente no se encuentra en línea, pero en su lugar puede utilizarse UCHIME (Edgar *et al.*, 2011) o DECIPHER (Wright *et al.*, 2012), disponibles también

en RDP. También se consideraron quimeras aquellas secuencias que produjeron discrepancias en el orden de las ramas de dos árboles filogenéticos generados con regiones distintas del gen 16S ADNr (533-873 y 874-1215, de acuerdo con la numeración de este gen en *Escherichia coli*). El tamaño final de la biblioteca, después de eliminar las quimeras, fue de 93 secuencias. Se hizo una selección de una sola secuencia para cada filotipo (se define filotipo como secuencias distintas). El corte utilizado para diferenciar estos filotipos o secuencias tipo (con una diferencia mayor al 98% entre cada filotipo) fue sugerido por Stackebrandt y Ebers en 2006. De esta manera, las 93 secuencias que se obtuvieron en la biblioteca quedaron representadas por un total de 42 secuencias, las cuales fueron registradas en GenBank con los números de acceso de KF303144 a KF303185.

Método independiente de PCR (secuenciación genómica por *shotgun*): Las secuencias del gen 16S ADNr en el metagenoma de la muestra GS017 correspondiente al Canal de Yucatán (Rusch *et al.* 2007) fueron detectadas mediante rRNA-HMM (Huang *et al.* 2009) y sólo se retuvieron aquellas mayores que 200 pb. Estas secuencias fueron inspeccionadas para la detección de quimeras mediante Mallard (Ashelford *et al.* 2006) y aquellas que pasaron la prueba fueron alineadas y clasificadas en Greengenes (DeSantis *et al.* 2006b). La alineación fue inspeccionada manualmente. Dada la naturaleza de la base de datos metagenómica obtenida por *shotgun*, los fragmentos recuperados no son superpuestos a lo largo de regiones extensas, lo que previene la construcción de OTUs mediante agrupación por similitud a nivel de secuencia.

Determinación taxonómica: Para determinar las categorías taxonómicas a nivel *phylum* de ambas bibliotecas (clones y *shotgun*), se utilizó la herramienta Classify de Greengenes en octubre del 2011 (DeSantis *et al.* 2006a, 2006b). La asignación taxonómica a nivel de orden, sólo para la biblioteca de clones, se determinó con Simrank (también herramienta de Greengenes), utilizando las filogenias de referencia que recientemente se actualizaron en esta base de datos (McDonald *et al.* 2012) y se clasificaron aquellas secuencias con > 98 % de similitud para una asignación taxonómica más confiable (Stackebrandt & Ebers, 2006). En caso de haber obtenido < 98 %, el índice de similitud correspondiente se especifica en el texto.

Análisis filogenético: Se realizó un análisis filogenético con las secuencias del gen 16S ADNr obtenidas en la biblioteca de clones con el fin de corroborar la identidad taxonómica generada en Greengenes, utilizando las 42 secuencias tipo o filotipos. Para cada una se descargaron las dos secuencias con mayor similitud de las bases de datos públicas (una obtenida mediante BLAST (Altschul *et al.*, 1990), en junio de 2011, y otra en Greengenes en esa misma fecha) y se usaron como referencia para la filogenia. Todas estas secuencias fueron alineadas en Greengenes (DeSantis *et al.* 2006a) y el alineamiento se revisó manualmente en BioEdit (Hall 1999). Para determinar la posición filogenética de

estas secuencias, se realizaron análisis de Neighbor-Joining en MEGA5 (Tamura *et al.*, 2011), de máxima parsimonia en PAUP 4.0 (Swofford, 2002) y de máxima verosimilitud en phyML 3.0 (Guindon *et al.*, 2010), usando las secuencias de arquea como grupo externo. El mejor modelo de sustitución se obtuvo con Model Generator 0.85 (Keane *et al.*, 2006).

Análisis de TRFs con la biblioteca de clones: Para la clasificación taxonómica de los TRFs obtenidos con T-RFLPs y para la comparación de estos datos con la biblioteca de clones, se utilizaron 59 secuencias de esta biblioteca seleccionadas en función de la calidad que tenían en el extremo 5'. Se generó un mapa de restricción (digestión *in silico*) con *Alu I* en BioEdit (Altschul *et al.* 1990) para cada una de estas secuencias, cuya determinación taxonómica ya se había hecho previamente, con el fin de identificar el tamaño de TRF correspondiente a cada taxón.

Comparación de costos de las estrategias moleculares: Los costos para cada estrategia se calcularon tomando en cuenta los precios en promedio de acuerdo con las cotizaciones de distintas compañías. En el caso de las secuencias obtenidas por *shotgun*, se tomó como referencia a Thomas *et al.* (2012). Para clones y T-RFLPs, el costo fue calculado enlistando el material necesario requerido para ambas técnicas (Taq polimerasa, dNTPs, materiales), añadiéndose un costo aproximado para la extracción de ADN. Para los T-RFLPs se consideró el costo por corrida en el secuenciador automático de capilares (rendimiento de capilares, materiales utilizados, POP6, formamida, *buffer* y marcador de peso molecular fluorescente) y para clonación se consideró el costo de los *kits* de clonación y extracción de plásmidos. Se tomó en cuenta el menor precio de secuenciación por muestra, comparando dos unidades de secuenciación: 1) Washington University High Throughput Sequencing (cotización) y 2) la Unidad de Síntesis y Secuenciación del Instituto de Biotecnología (Universidad Nacional Autónoma de México, disponible en línea, consultado 2 noviembre 2013). Con fines comparativos, también se calcularon costos de técnicas de secuenciación de nueva generación (NGS), consultando para ello las siguientes páginas: Langebio (cotización), UC Davis, University of Tennessee y University of Virginia (disponibles en línea, consultados 2 noviembre 2013).

RESULTADOS

Diversidad, medida como número de unidades taxonómicas detectada: La *Tabla 1* resume el número de grupos taxonómicos que fue posible determinar para cada técnica, y en cada apartado se discutirán estos datos con detalle.

T-RFLPs versus biblioteca de clones : Con la técnica de T-RFLPs se detectaron 39 OTUs (fragmentos); la mayoría de ellos resultó con poca abundancia y sólo algunos fueron muy abundantes (*Fig. 1*). Para comparar estos datos con la biblioteca de clones, se utilizaron las secuencias de esa biblioteca ($n = 59$) cuyos cromatogramas mostraron buena calidad en el extremo 5' (con picos bien

Tabla 1. Número de unidades taxonómicas detectadas por las tres metodologías utilizadas: biblioteca de clones, *shotgun* y T-RFLPs.

	Biblioteca de clones	<i>Shotgun</i>	T-RFLPs
OTUs (fragmentos)	17	ND	39
OTUs (secuencias tipo)	42	ND	ND
<i>Phyla</i>	7	34	ND

definidos, separados entre sí y con poco ruido de fondo). Con estas secuencias, se realizó una restricción *in silico* que generó 17 fragmentos distintos (OTUs), cuya clasificación taxonómica (correspondiente a cada una de las 59 secuencias seleccionadas) se especifica en la *Tabla 2*. Sólo nueve OTUs coincidieron por tamaño en ambas metodologías (*Fig. 1*) y el OTU más abundante obtenido para ambas metodologías fue el fragmento de 123 pares de bases. Este fragmento puede provenir de una gran variedad de secuencias de la biblioteca de clones correspondientes a distintos *taxa* de bacterias: *Candidatus Pelagibacter* y Rhodospirillaceae, que son Alfacaproteobacterias. *Vibrionales* y el OTU 3370, en el caso de Gammaproteobacteria, y por último del orden Sphingobacteriales que pertenece al *phylum Bacteroidetes*.

Debido a que en la mayoría de los casos los fragmentos de *shotgun* no tienen la misma región terminal que los productos de PCR (con los cuales se generan los fragmentos de restricción terminales), no fue posible realizar una digestión *in silico* para compararlos con las otras dos metodologías.

Biblioteca de clones versus *shotgun*: El tamaño promedio de las 860 secuencias del gen 16S ADNr de la secuenciación por *shotgun* fue de 400 pares de bases y se detectaron 34 *phyla*, mientras que en la biblioteca de clones para este mismo gen sólo se pudieron detectar siete *phyla* (*Fig. 2*), con una extensión de cerca de 700 pares de bases por secuencia. Todos los *taxa* que componen la biblioteca de clones fueron detectados en la biblioteca de *shotgun*. Alfacaproteobacteria y Gammaproteobacteria fueron los *phyla* más abundantes recuperados por ambas técnicas (20.4 % Alfacaproteobacteria y 18.8 % Gammaproteobacteria en la biblioteca por *shotgun* y 53.7 % y 34.4 % en la biblioteca de clones, respectivamente). Bacteroidetes fue uno de los *phyla* más abundantes (15.11 %) en las secuencias obtenidas por *shotgun*, aunque fue detectado en menor proporción en la biblioteca de clones (4.3 %). Representantes poco abundantes, que resultaron únicos o con dos secuencias en la biblioteca de *shotgun* (con 0.11 % o con 22 % del total) fueron los *phyla* Acidobacteria, Chlorobi, Fusobacteria, Nitrospirae, OP11, SR1, Thermi, TM6 y WS3. En el caso de la biblioteca de clones, los *phyla* con un solo representante (Euryarchaeota, Archaea), Cyanobacteria y *Marine group A* (o Grupo A marino), con ~1 % del total, tuvieron una abundancia intermedia en la biblioteca obtenida por *shotgun*.

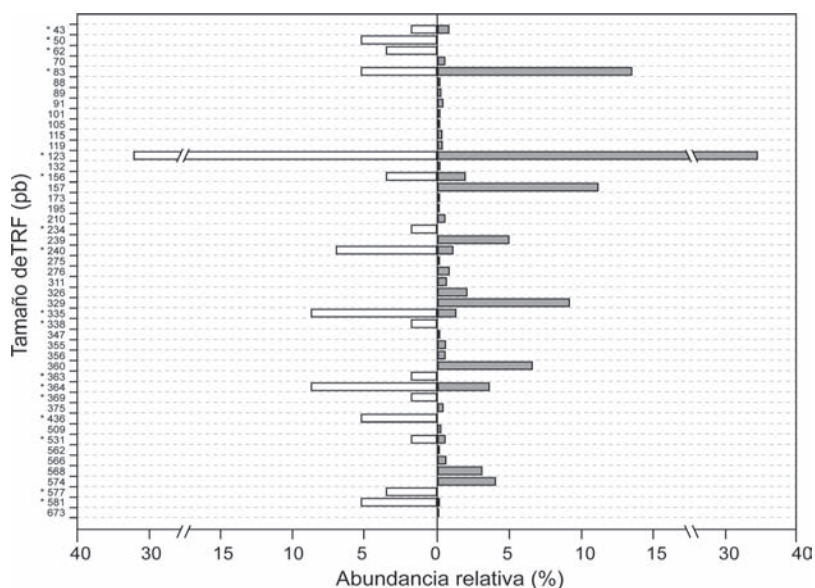


Figura 1. Fragmentos de restricción del gen 16S ADN_r generados con T-RFLPs y con la biblioteca de clones. El largo de las barras representa la abundancia relativa de cada fragmento en la muestra total. Las barras blancas corresponden a fragmentos de restricción obtenidos *in silico* sobre las secuencias provenientes de la biblioteca de clones y las barras grises corresponden a fragmentos de restricción generados con T-RFLPs. Con asterisco se resaltan los fragmentos a los que fue posible asignarles clasificación taxonómica, la cual se especifica en la Tabla 2.

Tabla 2. Clasificación taxonómica de los OTUs que fue posible designar, utilizando los fragmentos *in silico* provenientes de la biblioteca de clones. Todas las secuencias se clasificaron utilizando un índice de similitud (Simrank) > 98 %, exceptuando los OTUs especificados en el paréntesis.

TRF (pb)	Clasificación Simrank de acuerdo con los fragmentos <i>in silico</i>
43	<i>Candidatus</i> Portiera (OTU 3744)
50	<i>Candidatus</i> Pelagibacter (OTU 2827)
62	<i>Candidatus</i> Portiera (OTU 3744)
83	<i>Candidatus</i> Pelagibacter (OTU 2828)
123	<i>Candidatus</i> Pelagibacter (OTU 2827), Rhodospirillaceae (OTU 2793 sc), Sphingobacteriales (OTU1170), <i>Photobacterium angustum</i> (OTU 3974, 93 %), Gammaproteobacteria (OTU 3370)
156	<i>Candidatus</i> Portiera (OTU 3744), Rhodospirillaceae (OTU 2793, 87.74 %)
234	<i>Candidatus</i> Portiera (OTU 3744)
240	<i>Candidatus</i> Portiera (OTU 3744)
335	<i>Candidatus</i> Portiera (OTU 3744)
338	<i>Pseudoalteromonas</i> (OTU 3440)
363	<i>Candidatus</i> Portiera (OTU 3744, 94.69 %)
364	<i>Candidatus</i> Pelagibacter (OTU 2827, 89.29 %), Rhodospirillaceae (OTU 2793, 94.14 %)
369	Vibrionales (swat-3)
436	<i>Candidatus</i> Portiera (OTU 3744), OM60 (OTU 3802)
531	<i>Marine group</i> II (OTU 198)
577	Deltaproteobacteria (OTU 3167)
581	<i>Candidatus</i> Pelagibacter (OTU 2827, 94.34 %)

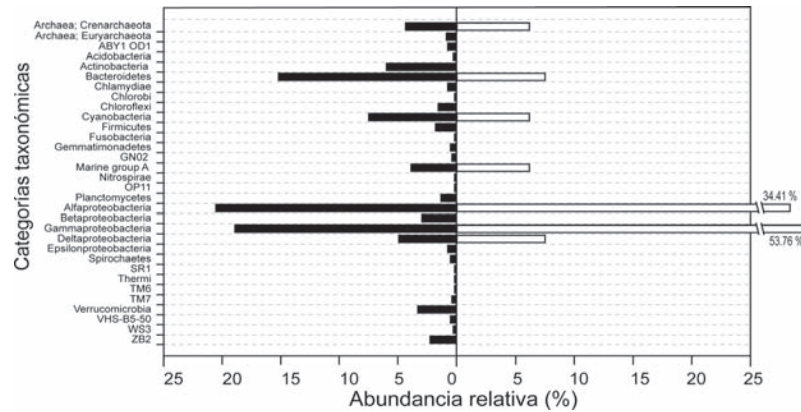


Figura 2. Diversidad de *phyla* detectados con biblioteca de clones y secuenciación genómica por *shotgun* del gen 16S ADNr. El largo de las barras representa la abundancia relativa de cada *phylum* en la muestra total. Las barras negras corresponden a *phyla* identificados en la base de datos de la secuenciación de ADN por *shotgun* y las barras blancas corresponden a *phyla* identificados en la biblioteca de clones.

Por otra parte, la biblioteca de clones se analizó a un nivel jerárquico más restringido, para lo cual, las 93 secuencias fueron agrupadas con un 98 % de similitud (Stackebrandt & Ebers, 2006) en 42 tipos de secuencias (filotipos o SecTipo en Fig. 3). Estas 42 secuencias son equivalentes a 42 unidades taxonómicas (OTUs). Veinticuatro de ellas resultaron únicas y cuatro resultaron ser las más abundantes, con más de cinco clones por tipo, asociadas a los géneros marinos no cultivables *Candidatus Pelagibacter* y *Candidatus Portiera* OTU 3744. El resto tuvo menor abundancia (de dos a cuatro clones por tipo). Sólo 17 de los 42 filotipos pudieron ser clasificados en estos niveles taxonómicos más bajos (Fig. 3).

Grupos taxonómicos: El árbol filogenético generado con los 42 filotipos de la biblioteca de clones muestra que la clasificación de la base de datos de Greengenes es congruente con la distribución filogenética de las ramas (Fig. 3) a los dos niveles analizados (*phylum* y orden). Se muestra la clasificación obtenida a nivel orden o menor sólo para aquellas secuencias que dieron > 98 % de similitud en Simrank. En contraste, con las secuencias de *shotgun* no fue posible hacer una filogenia, ya que corresponden a fragmentos que cubren distintas regiones del gen 16S ADNr, lo cual no permite un alineamiento adecuado.

Costos: El costo aproximado actual de una biblioteca (una sola muestra), obtenida utilizando secuenciación genómica por *shotgun*, es de 400,000 dólares (EE UU) (Thomas et al. 2012).

Para hacer una biblioteca de 96 clones se necesita una inversión de 465 dólares y, por último, una muestra de ADN analizada con T-RFLPs tiene un costo de 25 dólares. La Tabla 3 contiene los detalles de estos estimados y los costos actuales para metodologías de nueva generación (NGS).

DISCUSIÓN

En este trabajo presentamos la comparación de tres metodologías para el estudio de diversidad de comunidades microbianas para lo que se usó como referencia una muestra ambiental del canal de Yucatán. Debido a la naturaleza de este estudio, la interpretación ecológica de los datos es necesariamente limitada en cuanto a los aspectos ecológicos de la muestra analizada. Sin embargo, es de gran utilidad al comparar detalladamente los alcances de las tres metodologías utilizadas: T-RFLPs, bibliotecas de clones y *shotgun*.

Diversidad de la muestra: Independientemente del método empleado, la distribución de la estructura de la comunidad observada fue similar a la que se ha descrito para la mayoría de las comunidades bacterianas en diversos ambientes (Woodcock et al. 2007; Pedrós-Alió, 2012): pocos *taxa* muy abundantes y muchos poco abundantes (Figs. 1 y 2). Adicionalmente, se observó gran diversidad de fragmentos (T-RFLPs) y *phyla* (*shotgun*), lo cual podría explicarse por la alta productividad primaria que ha sido reportada en la zona de donde proviene la muestra (CONABIO, disponible en línea, consultado 2 noviembre 2013; Barton et al. 2010).

Más allá de los patrones generales de diversidad, la información taxonómica fina sólo pudo lograrse con las secuencias de la biblioteca de clones al detectar *taxa* a niveles jerárquicos más restringidos, incluso a nivel de género (Fig. 3). Las secuencias corresponden a bacterias que han sido reportadas como parte de comunidades oceánicas en todo el mundo. Los dos *taxa* más abundantes fueron *Candidatus Pelagibacter* (Alfaproteobacteria) y *Candidatus Portiera* OTU 3744 (Gammaproteobacteria, equivalente al clado SAR86). Son bacterias heterótrofas aeróbicas que se han especializado en utilizar distintas fuentes de carbono,

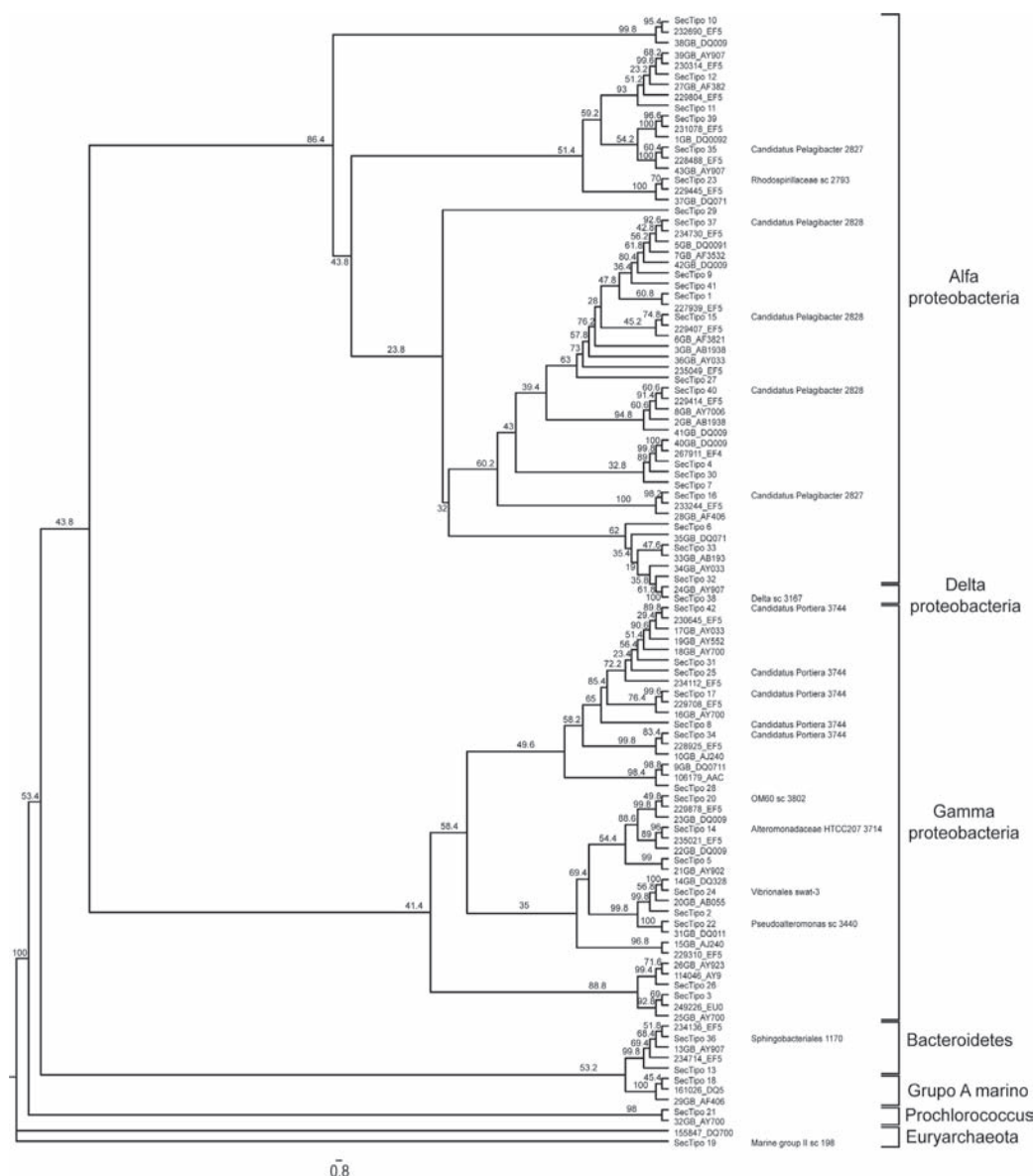


Figura 3. Representación filogenética de la diversidad del gen 16S ribosomal encontrada en la biblioteca de clones de una muestra de agua del canal de Yucatán. El árbol fue construido con un algoritmo de máxima verosimilitud (no se muestran los árboles construidos con los otros dos algoritmos, ya que generaron una topología similar). Las secuencias SecTipo corresponden a las 42 secuencias tipo generadas en la biblioteca de clones, y el resto son secuencias obtenidas de bases de datos públicas. La clasificación correspondiente (utilizando Simrank), a nivel orden o menor, se asignó sólo en las secuencias tipo con similitud > 98 % a la base de datos. El número en cada rama corresponde al valor de *bootstrap* (%) sobre 500 réplicas.

aparentemente para reducir la competencia (Dupont *et al.* 2011). En contraste también se encontraron potenciales competidores como *Pseudoalteromonas* 3440 y Vibrionales (*swat3*), capaces de inhibir el crecimiento de muchas especies de bacterias (Skovhus *et al.* 2007, J. Craig Venter Institute, disponible en línea, consultado 2 noviembre 2013). Otro grupo detectado fue el de las bacterias fotótrofas de la familia Rhodospirillaceae (e.g. OTU 2793 y

OM60; Madigan *et al.* 2003; Zheng *et al.* 2011) y las arqueas del grupo marino II (OTU 198), de las que recientemente se ha reportado una proteorodopsina que es ancestral a la de las bacterias (Iverson *et al.* 2012). Adicionalmente fueron identificados algunos *taxa* bacterianos con utilidad industrial o ecológica, como Alteromonadaceae (HTCC207 OTU 3714), el cual es importante para el reciclaje y biodegradación de materia orgánica, en particular de

Tabla 3. Detalles de los costos en USD estimados para los tres métodos moleculares utilizados, así como el costo de algunas técnicas de secuenciación de nueva generación.

	Métodos en este trabajo							Plataformas de secuenciación de Nueva Generación (NGS)			
	T-RFLPs	Shotgun	Clonas	454 GS Junior Langebío	454 GSxL Langebío	454 GS Titanium Langebío	454 GS Titanium U. de Tennessee	Illumina UCDavis	Ion Proton Langebío		
Extracción ADN	22	22	22	22	22	22	22	22	22		
PCR	1		1								
Clonación			15								
Extracción plásmido			152								
Secuenciación	4		275								
Precio reportado		400 000 (Thomas <i>et al.</i> , 2012)		2 772	5 343	5 300 y 50 900*	16 000	3 553	2 825		
Total por muestra	25	400 022	465	2 794	5 365	5 322 y 50 922*	16 022	3 575	2 847		
Tipo de datos obtenidos	Un perfil de fragmentos por muestra	1 millón y medio de secuencias mayores que 700 pb	96 secuencias	92 000 secuencias de 31 pb	1 millón de secuencias de 500 pb	1 millón de secuencias de 380 pb	1 millón de secuencias de 380 pb	90 millones de secuencias de 100 pb	20 millones de secuencias de 200 pb		

•Costo total de una placa que incluye la preparación de 192 bibliotecas (para más información, consultar el texto).

hidrocarburos (Jin *et al.* 2012). También encontramos a las Delta-proteobacteria (OTU 3167), cuyos linajes de mares fríos (equivalentes al clado SAR324 en la taxonomía de Silva (Pruesse *et al.* 2007)) desempeñan un papel importante en el reciclaje de fosfato orgánico disuelto (Brown & Donachie, 2007). Por último, se identificaron secuencias del orden Sphingobacteriales (OTU 1170), cuya capacidad de adherencia con algas ha sido recientemente reportada (Gómez-Pereira *et al.* 2012).

El inventario de la diversidad de microorganismos en distintos hábitats abre la posibilidad de investigar la relación entre ciertos grupos microbianos (algas, protozoarios, bacterias, hongos, virus) y aspectos funcionales de los ecosistemas. La caracterización de tal diversidad de organismos a través de estrategias genómicas independientes de cultivo permite el monitoreo de ambientes naturales y el efecto de perturbaciones en la diversidad como uno de los controladores de las funciones. Por ejemplo, el clado SAR11, al que pertenece *Candidatus Pelagibacter*, ha sido utilizado para monitorear los patrones de calentamiento en los mares, lo que potencialmente permitirá un mejor entendimiento de los patrones de respuesta de éstas y otras especies de bacterioplancton al cambio climático (Giovannoni & Vergin, 2012). Por otra parte, existe una estrecha relación entre la estructura de las comunidades bacterianas en el océano con la productividad primaria, medida como clorofila A (Falcón *et al.* 2008, Yilmaz *et al.* 2012) que podría determinarse en el canal de Yucatán con estudios estacionales en esta zona. Sin embargo, dependiendo del objetivo con el que se realice el estudio, el diseño experimental, número de muestras y número de secuencias o datos genómicos obtenidos son cruciales en la interpretación ecológica de la información. De esta manera, los datos aquí presentados son ilustrativos del tipo de información que puede obtenerse, aunque no nos permiten por el momento llegar a conclusiones ecológicas definitivas.

Comparación de metodologías para el análisis de diversidad procarionte:

La comparación de los patrones obtenidos con los tres métodos arrojó claras diferencias en la caracterización de la diversidad de la muestra analizada (i.e. número de *taxa* observados, Tabla 1). En cuanto a las diferencias entre T-RFLPs y la biblioteca de clones, la observación más importante es la evidencia de un muestreo incompleto en ambas técnicas, reflejado en la falta de coincidencias en la presencia y/o ausencia de OTUs. Esta discrepancia es producida principalmente por aquellos fragmentos con menor abundancia que pueden o no ser detectados por la técnica utilizada, efecto característico de una intensidad de muestreo moderada. A pesar de que en el análisis por T-RFLPs un OTU realmente no representa un taxón particular (ya que el fragmento puede ser generado por distintos *taxa* que tengan un mismo sitio de restricción), la caracterización de una muestra a partir del análisis de la diversidad de fragmentos sí permite hacer inferencias indirectas sobre la diversidad de la muestra, como ha sido reportado en análisis de comunidades de suelos por Noguez

et al. (2005), por Fierer & Jackson (2006) o recientemente por Reed & Martiny (2013). Así, el patrón de fragmentos obtenidos por T-RFLPs para esta muestra indica que la diversidad de *taxa* en el canal de Yucatán es alta.

Por otra parte, la diferencia entre la biblioteca de clones y la de *shotgun* hizo evidente que esta última estrategia de secuenciación permite un muestreo más exhaustivo. La secuenciación tipo Sanger asociada a *shotgun* tiene bajas tasas de error y lecturas largas (hasta 700 pb) (Thomas *et al.* 2012) lo que permite que la clasificación de las secuencias que se obtienen sea precisa (Hao & Chen, 2012). Además, la secuenciación por *shotgun*, al no tener el sesgo de amplificación (PCR) que otras técnicas sí tienen, ha resultado un excelente modelo para comparar con las dos metodologías que se llevaron a cabo en el presente trabajo. De esta forma, es posible afirmar que la abundancia relativa que se obtuvo con la base de datos de la biblioteca por *shotgun* representa con mayor precisión la estructura de la ensamble procarionte del canal de Yucatán. La abundancia relativa de las Alfa-proteobacteria y Gammaproteobacteria en la biblioteca de clones (53.7 % y 34.4 %, respectivamente), por ejemplo, hace pensar que durante la PCR se incorpora un sesgo hacia estos grupos, ya que la abundancia obtenida para estos mismos grupos por *shotgun* fue mucho menor (20.4 % y 18.8 %). Este sesgo por PCR se da por amplificación preferencial de ciertos grupos filogenéticos y ha sido reportado previamente (ver por ejemplo Frank *et al.* 2008; Bogert *et al.* 2011), un efecto que se agudiza si consideramos que el tamaño de muestra (número de clonas analizadas) no es muy grande. Aún con las ventajas y precisión que tiene la secuenciación de ADN por *shotgun* sobre otras, actualmente está utilizándose cada vez menos en estudios de diversidad metagenómica, ya que la secuenciación por el método Sanger que emplea es muy costosa y, además, requiere mayor tiempo que las técnicas de nueva generación (NGS) para el procesamiento de las muestras. Un ejemplo de estas nuevas técnicas de secuenciación es la pirosecuenciación por *barcoding* (amplicones provenientes de un solo gen, Huse *et al.*, 2008), que además de tener un menor costo, produce un mayor número de secuencias (Suenaga, 2012; Williamson & Joosephs, 2012). En promedio (dependiendo de la plataforma utilizada y el tipo de corrida que se elija) se pueden obtener bibliotecas desde unas 1500 hasta un millón de secuencias (Tabla 3) con una longitud más corta que las de clones (desde 100 hasta 500 pb hoy en día). Posibilidades como ésta actualmente han mejorado en gran medida el problema del número de secuencias obtenidas en el análisis de las comunidades microbianas.

La biblioteca de clones aporta información más precisa en cuanto a resolución filogenética, ya que produce secuencias más largas, aunque en menor número. El largo de las secuencias obtenidas permite una detallada asignación taxonómica de los integrantes de la comunidad incluso a nivel de especie. La secuenciación genómica por *shotgun* en cambio tiene una resolución taxonómica menor, debido a la reducida longitud de los frag-

mentos. Además, estas fracciones de ADN provienen de distintas regiones del gen, por lo que realizar una alineación o un análisis filogenético es casi imposible.

De los tres métodos utilizados, no es difícil ver que la aproximación por T-RFLPs es la más pobre en su resolución taxonómica, ya que se obtienen fragmentos del mismo tamaño, que pueden resultar de grupos filogenéticos muy distintos (Tabla 2). Esta limitación metodológica es evidenciada por la reducida asignación taxonómica que puede hacerse de los fragmentos (Fig. 1 y Tabla 2), algunos de ellos posiblemente producidos por distintos *taxa*, o por grupos filogenéticos que tienen más de un fragmento característico. Sin embargo, es necesario considerar que para laboratorios con bajo presupuesto, la técnica de T-RFLPs representa una buena opción, ya que es metodológicamente sencilla, pero poderosa en cuanto al conocimiento que puede generarse. Además de su bajo costo y sencillez, publicaciones recientes han demostrado la utilidad de los T-RFLPs en investigaciones ecológicas que requieran seguimiento de comunidades en el tiempo o en tratamientos experimentales, en las cuales uno de los requisitos es la identificación de cambios generales en la composición de las comunidades (ver por ejemplo Erwin *et al.* 2011; Escalante *et al.* 2008; Lavin *et al.* 2008). También se han publicado patrones de diversidad o composición microbiana de T-RFLPs que están asociados a factores ambientales (Fierer & Jackson, 2006; Noguez *et al.* 2008; Reed & Martiny, 2013; Zhao *et al.* 2012), lo que permite hacer importantes inferencias ecológicas sobre la estructura de las comunidades de bacterias. Esta técnica permite igualmente, de una manera rápida y precisa, comparar *grosso modo* varias localidades al mismo tiempo, así como estimar la diversidad de cada una, utilizando herramientas clásicas de la Ecología (Noguez *et al.* 2005; Reed & Martiny, 2013).

La elección de uno u otro método para la caracterización de diversidad microbiana requiere, necesariamente, de la consideración simultánea de los costos y de las capacidades de cada uno. En este sentido, ya se han mencionado detalladamente alcances y limitaciones de los tres métodos en cuanto a diversidad (número de *taxa* observados) y asignación taxonómica, que junto con los datos de costo, proporcionan los argumentos necesarios para sugerir qué tipo de estrategia es más adecuada para distintas investigaciones de orden ecológico o evolutivo. Ya se ha indicado también que con el desarrollo de tecnologías NGS ha sido posible incrementar el número de secuencias por muestra a costos cada vez más bajos, por lo cual también es necesario considerarlas. En términos generales, obtener datos con una biblioteca *shotgun* constituye la aproximación más costosa (400,022 dólares), seguida por la construcción de bibliotecas de clones (465 dólares) y por último la de los T-RFLPs (25 dólares) (precios considerando la extracción de ADN, Tabla 3). En cuanto a NGS, el mejor precio que se encontró lo ofrece Langebio, en Irapuato, México. Por muestra se cobran ≈265 dólares, cifra menor que el precio estimado por biblioteca de clones (460 dólares); para este ejemplo en particular

cada muestra generaría, en promedio, unas 2000 secuencias de 380 pb de largo con la plataforma GS Titanium. Este cálculo se obtuvo del presupuesto para la corrida de una placa, dividida en 16 regiones, y utilizando 12 marcas (*barcoding*) distintas en cada región, para un total de 192 muestras. Ésta y otras compañías generalmente no ofrecen el servicio por muestra, por lo que es necesario pagar la placa entera (≈3000 dólares). Si a un laboratorio le interesan 192 bibliotecas con estas características, es posible juntar todas las muestras en una sola placa o corrida (192 es sólo un ejemplo para obtener el mejor precio por muestra, aunque hay opciones para menor número de muestras), lo cual significa una inversión total de 50,943 dólares. La diferencia de precio entre secuenciar una sola muestra o varias (en este caso, 192) radica en el trabajo de preparación de las muestras individuales (*barcoding* y bibliotecas). Ya sea una sola muestra o múltiples (multiplex), con este tipo de tecnología (pirosecuenciación) se obtienen desde 92 000 hasta un millón de secuencias (divididas entre el número de muestras), sin embargo, la variación de precios es grande (Tabla 3). Por ejemplo el costo de la plataforma GS Titanium es de 5322 dólares en Langebio, y el precio por este mismo tipo de secuenciación en la Universidad de Tennessee es de 16,000 dólares. Con otras tecnologías NGS (Illumina, por ejemplo) el precio es menor por muestra: 3553 dólares para una biblioteca de 90 millones de fragmentos de 100 pb en UCDavis, y es aún menor utilizando Ion Proton, por ejemplo en Langebio el precio por muestra de esta tecnología es de 2825 dólares por 20 millones de fragmentos de 200 pb.

Estas comparaciones son considerando únicamente escenarios en donde el objetivo sea analizar la diversidad de amplicones específicos (e.g. 16S rDNA, ITS1), sin embargo, el potencial de estrategias como *shotgun* e incluso estas nuevas tecnologías de secuenciación incluye secuenciar genomas completos de organismos únicos o metagenomas (genomas de comunidades completas). Con ello, en ambos esquemas (*shotgun* o NGS), se genera información que permite hacer inferencias sobre el funcionamiento de la comunidad microbiana, además de conocer la composición taxonómica (Bonilla-Rosso *et al.* 2008). Con este enfoque se han reportado un número cada vez mayor de estudios evolutivos que analizan la relación entre la diversidad taxonómica de los microorganismos con la función ecosistémica (Bonilla-Rosso *et al.* 2012; Peimbert *et al.* 2012).

Aunque puede ser obvio para este momento, cabe mencionar que la mayor demanda de trabajo al utilizar una técnica como la secuenciación genómica por *shotgun* (o métodos de NGS) se da en la parte de análisis informático de las extensas bases de datos generadas en el proceso de secuenciación, por lo que tendría que considerarse una inversión en equipos de cómputo en los costos finales. Muchos programas y plataformas de programas fueron desarrollados en los últimos años, disponibles de manera gratuita para los investigadores (por ejemplo, la plataforma QIIME, Caporaso *et al.* 2010), pero aquellos análisis en los que

se requiera hacer modificaciones de los *pipelines* (relacionados con el proceso de programación) requerirán de ciertas habilidades para programar que también es necesario considerar. De esta forma, la inversión del tiempo necesario para los análisis, así como los costos que ello genere y el *software* son factores importantes para tomar en cuenta (Yoccoz, 2012).

En resumen, los estudios de tipo experimental, en donde sean requeridas muchas réplicas o seguimiento temporal, así como el estudio de patrones de diversidad asociados a factores ambientales, podrían verse beneficiados por los bajos costos y la alta reproducibilidad de los T-RFLPs, considerando por supuesto la limitación de la asignación taxonómica (Erwin *et al.* 2011; Fierer & Jackson, 2006; Reed & Martiny, 2013; Zhao *et al.* 2012). En contraste, estudios que requieran una caracterización minuciosa de la diversidad y abundancia de grupos taxonómicos con precisión filogenética deberán apostar por una estrategia de bibliotecas de clones o pirosecuenciación con *barcoding* (Huse *et al.* 2008), dependiendo del presupuesto con que se cuente. En cuanto a la estrategia por *shotgun*, ésta es ideal en estudios exploratorios, de descripción gruesa de diversidad y de bioprospección (de genes y funciones), en donde no se requiere el análisis de múltiples muestras o réplicas experimentales (Simon & Daniel, 2011).

AGRADECIMIENTOS

Esta investigación fue financiada por Semarnat/Conacyt 2002-C01-0246 y además constituye un requisito parcial del programa del Posgrado en Ciencias Biomédicas de la Universidad Nacional Autónoma de México (UNAM). Laura Espinosa-Asuar agradece el apoyo de la beca proporcionada por el Consejo Nacional de Ciencia y Tecnología (Conacyt) no. 113 997, y de la UNAM. Agradecemos a la doctora Erika Aguirre Planter por su apoyo técnico, a la doctora Karla Heidelberg por facilitarnos la porción del filtro y su colaboración en la extracción del ADN, al biólogo Jaime Gasca Pineda por su asesoría en las filogenias y a la M. en Edición Concepción Asuar por su ayuda en la edición y corrección de estilo de este artículo.

REFERENCIAS

- ALLISON, S. D., Y. LU, C. WEIHE, M. L. GOULDEN, A. C. MARTINY, K. K. TRESSEDER & J. B. MARTINY. 2013. Microbial abundance and composition influence litter decomposition response to environmental change. *Ecology* 94: 714-725.
- ALTSCHUL, S. F., W. GISH, W. MILLER, E. MYERS & D. J. LIPMAN. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 522-529.
- ASHFELFORD, K. E., N. A. CHUZHANOVA, J. C. FRY, A. J. JONES & A. J. WEIGHTMAN. 2006. New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. *Applied Environmental Microbiology* 72: 5734-5741.
- BARTON, A. D., S. DUTKIEWICZ, G. FLIERL, J. BRAGG & J. F. FOLLOWS. 2010. Patterns of diversity in marine phytoplankton. *Science* 327: 1509-1511.
- BOGERT, B. VAN DEN, M. DE VOS WILLEM, E. G. ZOETENDAL & M. KLEEREBEZEM. 2011. Microarray analysis and barcoded pyrosequencing provide consistent microbial profiles depending on the source of human intestinal samples. *Applied and Environmental Microbiology* 77: 2071-2080.
- BONILLA-ROSSO, G., V. SOUZA & L. E. EGUIARTE. 2008. Metagenómica, genómica y ecología molecular: la nueva ecología en el bicentenario de Darwin. *TIP Revista Especializada en Ciencias Químico-Biológicas* 11: 41-51.
- BONILLA-ROSSO, G., M. PEIMBERT, L. D. ALCARAZ, I. HERNÁNDEZ, L. E. EGUIARTE, G. OLMEDO-ÁLVAREZ & V. SOUZA. 2012. Comparative metagenomics of two microbial mats at Cuatro Ciénegas Basin II: community structure and composition in oligotrophic environments. *Astrobiology* 12: 659-673.
- BROWN, M. V. & S. P. DONACHIE. 2007. Evidence for tropical endemism in the Deltaproteobacteria Marine Group B/SAR324 bacterioplankton clade. *Aquatic Microbial Ecology* 46: 107-115.
- CAPORASO, J. G., J. KUCZYNSKI, J. STOMBAUGH, K. BITTINGER & F. D. BUSHMAN. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7: 335-336.
- COLE, J. R., B. CHAI, T. L. MARSH, R. J. FARRIS, Q. WANG. 2003. The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Research* 31: 442-443.
- CONABIO. 1988. *La diversidad biológica de México: Estudio de país*. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México.
- CONABIO (disponible en): http://www.conabio.gob.mx/gap/index.php/Procesos_oceanograficos (consultado 2 noviembre 2013).
- DESANTIS, T. Z., P. HUGENHOLTZ, N. LARSEN, M. ROJAS & E. L. BRODIE. 2006a. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology* 72: 5069-5072.
- DESANTIS, T. Z., P. HUGENHOLTS, K. KELLER, E. L. BRODIE, N. LARSEN, Y. M. PICE-NO, R. PHAN & G. L. ANDERSEN. 2006b. Nast: a multiple sequence alignment server for comparative analysis of 16s rRNA genes. *Nucleic Acids Research* 34: W394-W399.
- DUPONT, C. L., D. B. RUSCH, S. YOOSEPH, M. J. LOMBARDO & R. A. RICHTER. 2011. Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *The ISME Journal* 6: 1186-1199.
- EDGAR, R. C., B. J. HAAS, J. C. CLEMENTE, C. QUINCE & R. KNIGHT. 2011. Uchime improves sensitivity and speed of chimera detection. *Bioinformatics* 27: 2194-2200.
- EGUIARTE, L. E., J. A. AGUIRRE-LIGUORI, L. JARDÓN-BARBOLLA, E. AGUIRRE-PLANTER & V. SOUZA. 2013. Genómica de poblaciones: nada en evolución va a tener sentido si no es a la luz de la genómica, y nada en genómica tendrá sentido si no es a la luz de la evolución. *TIP Revista Especializada en Ciencias Químico-Biológicas* 16: 5-19.

- ERWIN, P. M., J. B. OLSON & R. W. THACKER. 2011. Phylogenetic diversity, host-specificity and community profiling of sponge-associated bacteria in the Northern Gulf of Mexico. *Plos One* 6: 1-16.
- ESCALANTE, A. E., L. E. EGUIARTE, L. ESPINOSA-ASUAR, L. J. FORNEY, A. M. NOGUEZ & V. SOUZA SALDÍVAR. 2008. Diversity of aquatic prokaryotic communities in the Cuatro Ciénegas Basin. *FEMS Microbiology Ecology* 65: 50-60.
- FALCÓN, L. I., A. M. NOGUEZ, L. ESPINOSA-ASUAR, L. E. EGUIARTE & V. SOUZA. 2008. Evidence of biogeography in surface ocean bacterioplankton assemblages. *Marine Genomics* 1: 55-61.
- FIERER, N. & R. B. JACKSON. 2006. The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America* 103: 626-631.
- FIERER, N. & J. T. LENNON. 2011. The generation and maintenance of diversity in microbial communities. *American Journal of Botany* 98: 439-448.
- FRANK, J. A., C. I. REICH, S. SHARMA, J. S. WEISBAUM, B. A. WILSON & G. J. OLSEN. 2008. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Applied and Environmental Microbiology* 74: 2461-2470.
- GIOVANNONI, S. J. & K. L. VERGIN. 2012. Seasonality in ocean microbial communities. *Science* 335: 671-676.
- GÓMEZ-PEREIRA, P. R., M. SCHÜLER, B. M. FUCHS, C. BENNKE & H. TEELING. 2012. Genomic content of uncultured Bacteroidetes from contrasting oceanic provinces in the North Atlantic Ocean. *Environmental Microbiology* 14: 52-66.
- GREEN, J. L., B. J. M. BOHANNAN & R. J. WHITAKER. 2008. Microbial Biogeography: from taxonomy to traits. *Science* 320: 1039-1043.
- GUINDON, S., J. F. DUFAYARD, V. LEFORT, M. ANISIMOVA, W. HORDIJK & O. GASCUÉL. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology* 59: 307-321.
- HALL, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series* 41: 95-98.
- HAO, X. & T. CHEN. 2012. OTU analysis using metagenomic shotgun sequencing data. *PLoS ONE* 7: e49785.
- HUANG, Y., P. GILNA & W. LI. 2009. Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* 25: 1338-1340.
- HUBER, T., G. FAULKNER & P. HUGENHOLTZ. 2004. Bellerophon; a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* 20: 2317-2319.
- HUGHES, J. B., B. J. BOHANNAN, J. H. BROWN, R. K. COLWELL, J. A. FUHRMAN & J. L. GREEN. 2006. Microbial biogeography: putting microorganisms on the map. *Natural Review of Microbiology* 4: 102-112.
- HUSE, S. M., D. LES, J. A. HUBER, D. M. WELCH, D. A. RELMAN & M. L. SOGIN. 2008. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLOS Genetics* 4: 1-10.
- IVERSON, V., R. M. MORRIS, C. D. FRAZAR, C. T. BERTHIAUME, R. L. MORALES & E. V. ARMBRUST. 2012. Untangling genomes from metagenomes: revealing an uncultured class of marine Euryarchaeota. *Science* 335: 587-590.
- J. CRAIG VENTER INSTITUTE (disponible en línea en): <https://moore.jcvi.org/moore/SingleOrganism.do?speciesTag = VSWAT3> (consultado 2 noviembre 2013).
- JIN, H. M., J. M. KIM, H. J. LEE, E. L. MADSEN & C. O. JEON. 2012. Alteromonas as a key agent of polycyclic aromatic hydrocarbon biodegradation in crude oil-contaminated coastal sediment. *Environmental Scientific Technology* 46: 7731-7740.
- KEANE, T. M., C. J. CREEVEY, M. M. PENTONY, T. J. NAUGHTON & J. O. MCINERNEY. 2006. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evolution Biology* 6: 29.
- LANGEBIO (disponible en línea en): <http://www.Langebio.cinvestav.mx/lab-sergen/> (consultado 2 noviembre 2013).
- LAVIN, P., P. GÓMEZ, B. GONZÁLEZ & O. ULLOA. 2008. Diversity of the marine picocyanobacteria Prochlorococcus and Synechococcus assessed by terminal restriction fragment length polymorphisms of 16S-23S rRNA internal transcribed spacer sequences. *Revista Chilena de Historia Natural* 81: 515-531.
- LINDSTROM E. S. & S. LANGENHEDER. 2012. Local and regional factors influencing bacterial community assembly. *Environmental Microbiology Reports* 4: 1-9.
- MADIGAN, M. T., J. M. MARTINKO & J. PARKER. 2003. Brock Biology of Microorganisms. Prentice-Hall: New Jersey, decimal edición.
- MCDONALD, D., M. N. PRICE, J. GOODRICH, E. P. NAWROCKI & T. Z. DESANTIS. 2012. An improved Greenegenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *The ISME Journal* 6: 610-618.
- NAZARIES, L., Y. PAN, L. BODROSSY, E. M. BAGGS, P. MILLARD, J. C. MURRELL & B. K. SINGH. 2013. Microbial regulation of biogeochemical cycles: evidence from a study on methane flux and land-use change. *Applied and Environmental Microbiology* 79: 4031-4040.
- NOGUEZ, A. M., H. T. ARITA, A. E. ESCALANTE, L. J. FORNEY, F. GARCÍA-OLIVA & V. SOUZA. 2005. Microbial macroecology: highly structured prokaryotic soil assemblages in a tropical deciduous forest. *Global Ecology Biogeography* 14: 241-248.
- NOGUEZ, A. M., A. E. ESCALANTE, L. J. FORNEY, M. NAVA-MENDOZA, I. ROSAS, V. SOUZA & F. GARCÍA-OLIVA. 2008. Soil aggregates in a tropical deciduous forest: effects on C and N dynamics, and microbial communities as determined by t-RFLPs. *Biogeochemistry* 89: 209-220.
- OSBORN, A. M., E. R. MOORE & K. N. TIMMIS. 2000. An evaluation of terminal-restriction fragment length polymorphism. T-RFLP analysis for the

- study of microbial community structure and dynamics. *Environmental Microbiology* 2: 39-50.
- PEDRÓS-ALIÓ, C. 2012. The rare bacterial biosphere. *Annual Review of Marine Science* 4: 449-466.
- PEIMBERT, M., L. D. ALCARAZ, G. BONILLA-ROSSO, G. OLMEDO-ÁLVAREZ, F. GARCÍA-OLIVA, L. SEGOVIA, L. E. EGUIARTE & V. SOUZA. 2012. Comparative metagenomics of two microbial mats at Cuatro Ciénegas Basin I: Ancient lessons on how to cope with an environment under severe nutrient stress. *Astrobiology* 12: 648-657.
- PRUESSE, E., C. QUAST, K. KNITTEL, B. FUCHS, W. LUDWIG, J. PEPLIES & F. O. GLÖCKNER. 2007. Silva: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research* 35: 7188-7196.
- REED, H. E. & J. B. MARTINY. 2013. Microbial composition affects the functioning of estuarine sediments. *The ISME Journal* 7: 868-879.
- RUSCH, D. B., A. L. HALPERN, G. SUTTON, K. B. HEIDELBERG & S. WILLIAMSON. 2007. The sorcerer II global ocean sampling expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLoS Biology* 5: e77.
- RZEDOWSKI, J. 2006. *Vegetación de México*. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad: México. Primera edición digital, 504 págs.
- SEGATA, N., D. BOERNIGEN, T. L. TICKLE, X. C. MORGAN, W. S. GARRETT & C. HUTTENHOWER. 2013. Computational meta'omics for microbial community studies. *Molecular Systematic Biology* 9: 666.
- SIMON, C. & R. DANIEL. 2011. Metagenomic analyses: past and future trends. *Applied and Environmental Microbiology* 77: 1153-1161.
- SKOVHUS, T. L., C. HOLMSTRÖM, S. KJELLEBERG & I. DAHLÖF. 2007. Molecular investigation of the distribution, abundance and diversity of the genus *Pseudoalteromonas* in marine samples. *FEMS Microbiology Ecology* 61: 348-61.
- STACKEBRANDT, E. & J. EBERS. 2006. Taxonomic parameters revisited: tarnished gold standards. *Microbiology Today* 33: 152-155.
- SUENAGA, H. 2012. Targeted metagenomics: a high-resolution metagenomics approach for specific gene clusters in complex microbial communities. *Environmental Microbiology* 14: 13-22.
- SWOFFORD, D. L. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts, USA (disponible en línea en): [http://www.tsu.edu/PDFFiles/CBER/Miranda/PAUP %20Manual.pdf](http://www.tsu.edu/PDFFiles/CBER/Miranda/PAUP%20Manual.pdf) (consultado 2 noviembre 2013).
- TAMURA, K. D. PETERSON, N. PETERSON, G. STECHER, M. NEI & S. KUMA. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28: 2731-2739.
- THOMAS, T., J. GILBERT & F. MEYER. 2012. Metagenomics - a guide from sampling to data analysis. *Microbial Informatics Experimentation* 2: 3.
- TURNER, S., K. M. PRYER, V. P. W. MIAO, & J. D. PALMER. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *Journal of Eukaryotic Microbiology* 46: 327-338.
- UCDAVIS GENOME CENTER (disponible en línea en): <http://dnatech.genome-center.ucdavis.edu/prices.html> (consultado 2 noviembre 2013).
- UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO (Unidad de Síntesis y Secuenciación, Instituto de Biotecnología) (disponible en línea en): <http://www.ibt.unam.mx/sintesis/secuenciacion.html> (consultado 2 noviembre 2013).
- UNIVERSITY OF TENNESSEE (Center for Environmental Biotechnology) (disponible en línea en): <http://www.ceb.utk.edu/dnasequence.html> (consultado 2 noviembre 2013).
- UNIVERSITY OF VIRGINIA MEDICAL SCHOOL (Center for Public Health Genomics) (disponible en línea en): <http://cphg.virginia.edu/mackey/projects/sequencing-pipelines/sweetseq/next-gen-sequencing-costs-at-ua/> (consultado 2 noviembre 2013).
- WILLIAMSON, S. J. & S. YOOSEPH. 2012. From bacterial to microbial ecosystems (metagenomics). In Helden J van, A. Toussaint & D. Thieffry (eds.) *Bacterial Molecular Networks: Methods and Protocols. Series Methods in Molecular Biology* 804: 35-55.
- WOODCOCK, S., C. J. VAN DER GAST, T. BELL, M. LUNN, T.P. CURTIS, I. M. HEAD & W. T. SLOAN. 2007. Neutral assembly of bacterial communities. *FEMS Microbiology Ecology* 62: 171-180.
- WRIGHT, E. S., L. S. YILMAZ & D. R. NOGUERA. 2012. DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Applied and Environmental Microbiology* 78: 717-725.
- YILMAZ, P., M. H. IVERSEN, W. HANKELN, R. KOTTMANN, C. QUAST & F. O. GLÖCKNER. 2012. Ecological structuring of bacterial and archaeal taxa in surface ocean waters. *FEMS Microbiology Ecology* 81: 373-385.
- YOCOZ, N. G. 2012. The future of environmental DNA in ecology. *Molecular Ecology* 21(8) (Special Issue environmental DNA): 2031-2038.
- ZARRAONAINDIA, I., D. P. SMITH & J. A. GILBERT. 2013. Beyond the genome: community-level analysis of the microbial world. *Biology & Philosophy* 28: 261-282.
- ZHAO, D., R. HUANG, J. ZENG, W. YAN, J. WANG, T. MA, M. WANG & Q. L. WU. 2012. Diversity analysis of bacterial community compositions in sediments of urban lakes by terminal restriction fragment length polymorphism (T-RFLP). *World Journal of Microbiology and Biotechnology* 28: 3159-3170.
- ZHENG, Q., R. ZHANG, M. KOBLÍZEK, E. N. BOLDAREVA, V. YURKOV, S. YAN & N. JIAO. 2011. Diverse arrangement of photosynthetic gene clusters in aerobic anoxygenic phototrophic bacteria. *PLoS One* 6: e25050.

Recibido: 12 de noviembre del 2013.

Aceptado: 31 de marzo de 2014.

CAPÍTULO II

ARTÍCULO EN PROCESO DE ENVIARSE

Aquatic bacterial assemblage structure in Pozas Azules,
Cuatro Ciénegas Basin in México: neutral vs. niche-based
factors

Aquatic bacterial assemblage structure in Pozas Azules, Cuatro Ciénegas Basin in México: neutral vs. niche-based factors

Running title: Aquatic bacterial assemblage structure: neutral vs. niche-based factors

Espinosa-Asuar L.,¹ Escalante A.E.,² Gasca-Pineda J.,¹ Blaz J.,¹ Peña L.,¹ Eguiarte L. E.,¹ and Souza V.^{1*}

¹Departamento de Ecología Evolutiva, and ²Laboratorio Nacional de Ciencias de la Sostenibilidad of Instituto de Ecología, Universidad Nacional Autónoma de México, México City, México.

*For correspondence. E-mail: souza@servidor.unam.mx; Tel. (+52) 55 5622 9006; Fax (+52) 55 5622 8995.

Keywords: bacteria, biogeography, Cuatro Ciénegas, inland water

Abstract

The causes for the observed microbial diversity distribution have been subject of investigation in microbial ecology for almost two decades. To date, there are mainly two groups of factors in consideration for explaining the causes of the observed patterns: those in which historical or geographical effects are preponderant, also known as neutral-based factors and, those in which local

adaptation to environmental conditions are more influential, also known as niche-based factors. In this study we aim to distinguish the influence of neutral-based vs. niche-based processes in the distribution of microbial diversity in four ponds within a stable over-time aquatic system known as Pozas Azules in the Cuatro Ciénegas Basin (CCB), México. We followed a sampling strategy of sites that were geographically delimited, and have low environmental variation to avoid obscuring distance (neutral) effects. Aquatic bacterial diversity was characterized following a culture independent approach (16S sequencing of clone libraries). Results show a correlation between bacterial beta diversity (1-Sorensen) and geographic distance, which indicates influence of neutral factors in the bacterial communities assembly. Our findings show, for the first time in the CCB area, the influence of neutral-based factors (history) in prokaryotic diversity distribution.

Introduction

Biogeography studies the distribution of organisms in space and time. This discipline provides understanding of the underlying causes of the distribution patterns, and ultimately of the drivers that generate and maintain biodiversity (Martiny et al. 2006, Lindström & Langenheder 2012, Hanson et al. 2012, Nemergut et al. 2013). Traditionally, biogeographical studies have been conducted in macro organisms, but in recent years, distribution patterns of microorganisms have been investigated through culture-independent approaches. These studies have advanced our understanding of the

mechanisms that generate and maintain microbial diversity (Noguez et al. 2005, Ragon et al. 2012, Zheng et al. 2013, Logares et al. 2013, Hatosy et al. 2013).

Causes of biogeographical patterns have remained controversial for a long time; specifically, there is an ongoing debate regarding the processes that lead to the assembly of communities, either being neutral (stochastic) or niche-based (deterministic). These two processes are frequently described using different terminologies, leading to confusion (Nemergut et al. 2013). The neutral and niche-based processes (used by Wang et al. 2013 for example), are terms that correspond, respectively, to dispersal history and environmental heterogeneity (dispersal history is also used as historical contingency or limitation processes, e.g. Ragon et al. 2012) and have been also equated to the regional and local (habitat) factors (e.g. Lindström and Langenheder, 2012). Finally, the most utilized terminologies are geographical or environmental association.

In order to investigate the biogeographical patterns in microorganisms, Martiny and colleagues (2006) proposed a framework to evidence the potential causes of microbial diversity. Martiny's model considers four scenarios of microbial distribution, distinguishing neutral from niche-based assembly processes. 1) A random scenario as the null hypothesis, in which microorganisms are arbitrarily distributed over space. 2) A deterministic hypothesis, in which the observed distribution reflects differences in the contemporary environment. 3) The existence of dispersal limits in bacteria, in which dispersal history is reflected in beta diversity (dissimilarity in composition) coupled with geographic distances. 4) An association between both historical factors and contemporary environmental

conditions. To properly test these four hypotheses, sampling design is critical, requiring samples from several locations that are geographically delimited, in systems with low environmental variation (to avoid obscuring distance effects) (Bissett et al. 2010, Logares et al. 2013).

The Cuatro Ciénegas Basin (CCB) is part of the Chihuahuan Desert (Coahuila, México) (Fig. 1) and considered a protected wetland since 1994 (APFF according to Mexican federal law). Despite the extreme oligotrophy (Elser et al. 2005, Souza et al. 2008, 2012), this place harbors high microbial diversity in different environments, including soil (López-Lozano et al. 2012, 2013), microbial mats (Bonilla-Rosso et al. 2012, Peimbert et al. 2012) and water (Souza et al. 2006, Pajares et al. 2012, 2013). Other studies conducted in CCB have investigated associations of microbial diversity distribution with spatial and environmental parameters. These studies did not find geographical association, neither at the community (Escalante et al. 2008, López-Lozano et al. 2012, 2013) or at population levels (Cerritos et al. 2011, Rodríguez-Verdugo et al. 2012, Rebollar et al. 2012). However, sampling design in these studies was performed in contrasting environmental systems that probably interfered with the statistical ability to detect geographic association (Bissett et al. 2010, Logares et al. 2013). Consequently, in this paper, we used a sampling design in non-contrasting and equivalent environments, in order to distinguish the influence of neutral-based vs. niche-based processes in the observed microbial diversity distribution. In particular, we studied microbial diversity in an aquatic system known as Pozas Azules, using a culture-independent approach. This system is characterized by

many isolated small circular ponds of different diameter (ranging from 6 to 50 m. in diameter and with maximum depths of 10m (Forti et al. 2004) locally named “pozas” (Fig. 1b), that are stable in temperature and in chemical properties over time (Johannesson et al. 2004, Peimbert et al. 2012). We considered that the spatial distribution of the ponds in Pozas Azules system is ideal for biogeographic exploration, as the ponds can represent aquatic islands in a sea of land (Dodson 1992). Because many of the reported spatial scales revealing historically contingent assembly (geographical association) in bacterial communities have been small (near to one km) (Hanson et al. 2012), we selected four nearby ponds (the maximum distance is ca. 1.8 km) with equivalent temperature and low variation on pH and conductivity measures.

Using the four scenarios proposed by Martiny (Martiny et al. 2006) we present evidence supporting biogeographic patterns associated with geographic distance of bacterial assemblages in the aquatic system of Pozas Azules from CCB. This result indicates that there are limits for dispersion of bacteria even at a scale smaller than two kilometers, consistent with the third hypothesis that explains community structure patterns based on historical contingencies or mainly influenced by neutral-based factors.

Material and Methods

Study site and sampling

On February of 2007 we sampled four ponds of the Pozas Azules aquatic system, a private preserve ranch since 2000 (Pronatura A.C. and The Nature

Conservancy property), that is located in the southwest flank of the San Marcos Sierra within the CCB (Fig. 1a). A recent general description of the hydrology of the CCB basin described the area of Pozas Azules as a karstic system, rich in CaHCO_3 , where water is a mixture of mountain recharge and carbonate aquifer groundwater (Wolaver et al. 2012). A preliminary exploration of the system ponds by our group showed a range of conductivity among them from 1900 to 4500 uS. We selected four neighboring ponds that were GPS referenced (Table 1). The chosen ponds were visually similar (clear and blue water, Fig. 1b), they were not visibly connected (maximum separation was 1.8 km, Fig. 1c), and were 10 to 30 meters in diameter. The selected ponds ranged from 7.5 to 7.8 in pH, and from 1900 to 2800 uS in conductivity.

Preliminary analysis in the CCB using TRFLPs showed that different samples taken from the same pond at the same time were very similar (data not shown). Ponds in this area are relatively homogeneous systems due to the presence of a spring stream that mixes water. This circumstance reduces spatial structuring as suggested by Lindström (2006) for aquatic environments, due to the small size of the ponds and the mixing effect of the spring flow and the wind. With this preliminary data, we concluded that a composite sample per pond gives an adequate representation of the bacterial community. In each pond we sampled five points, four of them near the edge (equidistant) and one at the center. All samples consisted of 2.5L of water collected at 50 cm below the surface. The five samples from each pond were mixed in one clean and sterile container (composite sample), and preserved at 4° C degrees, until processing (filtered)

few hours later. Composite samples were divided in subsamples for DNA (three 4L subsamples) and nutrient analyses (three 100ml subsamples). The water samples for DNA analysis were filtered first with a 10µm Durapore™ filter, and next with a 0.22µm Durapore™ filter (Millipore, USA) using a field vacuum pump. Filters were placed into sterile 2 ml eppendorf tubes, and were preserved in liquid nitrogen until brought to the laboratory, where they were stored at -80°C until DNA extraction. Water for nutrient analysis was filtered using a 0.22µm nitrocellulose filter, and the water was preserved at 4°C until processing.

Physico-chemical and nutrient analyses

Measurements of pH, conductivity (Total Dissolved Solids (TDS), salinity and conductivity), temperature, and dissolved oxygen were registered in the field for all five sampling points using a hydrolab YSI600QS (YSI Inc., Yellow Springs, OH).

Total carbon (TC) and inorganic carbon (IC), total nitrogen (TN) and dissolved inorganic nitrogen (DIN or NH₄), nitrite (NO₂), total phosphorus (TP) and inorganic phosphorous (PO₃) were determined in the laboratory following standard procedures (Peimbert et al. 2012), using triplicates of 100 ml from the composite samples. Dissolved organic carbon (DOC), phosphorus (DOP) and nitrogen (DON) were determined as the difference of the total minus inorganic measures.

DNA isolation and sequencing of 16S rRNA clone libraries

Genomic DNA was extracted from the filtered samples (three replicates per pond) using an UltraClean Water DNA Kit (MoBio, Carlsbad, CA, USA), including

a final cleaning step of phenol-chloroform procedure (Sambrook & Russell 2006). PCR amplification was performed independently on each one of the three DNA replicates per pond, by targeting 16S rRNA gene fragments of the bacterial domain using the universal primers 27F (5'-AGA GTT TGA TCC TGG CTC AG-3') and 1492R (5'-GGT TAC CTT GTT ACG ACT T-3') (Lane 1991). The three PCR products for each pond were mixed prior to cloning reaction. The cloning reaction, plasmid extraction, sequencing and PCR conditions were performed as previously reported (Pajares et al. 2012). We obtained one clone library per pond ($n=80\pm 10$) and partial sequences of the 16S rRNA gene were obtained using the primer 27F.

Sequence analysis

Sequences were trimmed and manually checked using BioEdit software 7.0.9.0 (Hall 1999). Sequences with average length of 700 bp were aligned with SILVA database and NAST aligner (DeSantis et al. 2006) using mothur v.1.31.2 (Schloss et al. 2009) and manually checked. Chimera sequences were identified using Bellerophon software (Huber et al. 2004), UCHIME v4.2.40 (Edgar et al. 2011), and DECIPHER (Wright et al. 2012) and were discarded from the dataset. Unique Operational Taxonomic Units (OTUs) were defined with a threshold of 97% sequence similarity and were identified using mothur v.1.31.2 (Schloss et al. 2009).

Diversity and statistical analyses

OTUs phylogenetic identity was determined using the classifier tool implemented in the Ribosomal Database Project (RDP) website (Wang et al. 2007). Once

unique sequences (OTUs) were determined, phylogenetic identity of OTUs was confirmed following a Neighbor-Joining analysis (500 bootstrap) in MEGA v.5. (Tamura et al. 2011) and a Maximum-Likelihood analysis (1000 bootstrap) using a TIM3+I+G model as determined by JModel test (Darriba et al. 2012) in PhyML v3.0 (Guindon et al. 2010).

Once the total number of unique OTUs was defined, a matrix with all the OTUs, their abundance and sampling site was constructed in order to perform diversity, ordination and correlation analyses.

Diversity analyses. To evaluate sampling effort and compare diversity among sampling sites, we determined coverage values, constructed rarefaction curves and estimated alpha diversity per pond with Chao 1, Shannon and Simpson diversity indices for all samples. Coverage values were determined using the Good estimator (Good 1953) following the equation $C = (1 - n/N) \times 100$, where C is the percentage of coverage of the library, n the number of singletons and N the total number of clones analyzed (Ragon et al. 2012). Beta diversity was calculated using Sorensen binary index (using dissimilarity 1-Sorensen) (Krebs C J 1999). Rarefaction curves and diversity indices were calculated in mothur v.1.31.2 (Schloss et al. 2009).

Ordination analyses. To visualize patterns of environmental variables distribution and evaluate their correspondence with OTU distribution, we followed multivariate and ordination approaches (Ramette 2007). Principal Component Analysis (PCA) was performed to visualize environmental variance of the sampling sites (rda function carried out in Vegan package (Oksanen et al. 2013)

of R, version 3.0.1 (2013-05-16) (R-Core-Team 2013)), and Canonical Correspondence Analysis (CCA) was performed to model species response to environmental parameters (Ter Braak 1987) with OTUs abundance and standardized environmental data (cca function implemented in Ade4 package of R (Dray & Dufour 2007)).

To standardize environmental variables, raw data were transformed to z-scores using decostand function (standardized method) implemented in R (Vegan package (Oksanen et al. 2013)). The non-redundant and explanatory environmental variables (except geographic location) were identified by using three statistical criteria for the standardized environmental variables: PCA's loadings (higher than 0.70), paired correlation between variables (Pearson) and Colldiag's index (> 50 , Suppl. Table 1). In addition, we also considered ecological relevance. The first component of the PCA analysis accounted for 69.2% of the total variation, mainly involving DOC (Dissolved Organic Carbon), Sal (salinity), IC (Inorganic Carbon), TDS (Total Dissolved Solids) and conductivity. Of these five variables, we found that DOC, Sal and IC are correlated (Suppl. Fig. 1), and we decided to reduce these three variables to only IC as the representative explanatory variable. Likewise, we discarded TDS due to its correlation with conductivity, since they are derived measures. Consequently, we only selected IC and conductivity from the first PCA component. The remaining four variables were retained in order to explore nutrient variables (NO_2 , TP (Total Phosphorus) or because their ecological relevance in determining aquatic bacterial communities (pH and TC (Total Carbon)) (Logue & Lindström 2008). All these

calculations were performed using R packages as following: the correlation between the variables was explored using cor function of Stats R package, and the graphics (order.single function) with Gclus (Hurley 2012) using panelutils.r script (Gillet 2007). Condition indexes and variance decomposition proportions (Colldiag function) were generated in Perturb v.2.05 (Hendrickx 2012).

Correlation analyses. To statistically determine the relative importance of geographic and contemporary environmental conditions we conducted correlational analyses using OTUs incidence data, environmental and geographical parameters. These analyses were conducted with functions implemented in Vegan package (Oksanen et al. 2013) in R, unless specified otherwise.

Three kinds of data matrices were constructed, a biological matrix composed by 1 - Sorensen index (estimated with Vegdist function), an environmental matrix using selected environmental variables (standardized z-scores) was constructed calculating Euclidian distance (dist function, Stats R package (R-Core-Team 2013)), and a spatial matrix that contained distances between ponds calculated from coordinates data.

To test the effects of geographic distance vs. environmental dissimilarity on the assemblages composition (beta-diversity), we conducted pairwise Mantel tests among the three matrices using Mantel function (Pearson's correlation, and 9999 permutations) (Mantel 1967).

To identify the environmental variables that best correlated with community dissimilarities (biological matrix) we used Bioenv function (using Pearson correlation and 1-Sorensen (Bray binary) arguments).

Nucleotide sequence accession numbers

Sequences were deposited in GenBank with accession numbers KJ998817 to KJ999144.

Results

OTU composition: moderate turnover and phylum dominance

Diversity estimated at OTU and phylum levels was distributed differently among sampling sites. We identified a total of 48 OTUs at 97% sequence similarity across 10 phyla (Figure 2a). From the 48 different OTUs, 17 were shared (i.e., found in more than one pond, Figure 2b). Regarding OTU abundance, all ponds had few abundant OTUs (more than 15 sequences per OTU) and many were rare (one or two sequences per OTU, Fig. 2b). The abundant OTUs in one pond were not exclusive, and were generally found in all the other ponds, but at lower numbers (Fig 2b). At the phylum level, each pond showed differences in relative abundance (Fig. 2c). The most abundant phyla were Actinobacteria in the A pond, Betaproteobacteria in B, Cyanobacteria in C, and Verrucomicrobia in pond F. In order to describe differences in composition among ponds, we calculated a pairwise dissimilarity index (1-Sorensen, that ranges from 0 -identical composition of samples- to 1 -no similitude-) based on OTU composition,

represented in a UPGMA (Figure 3), and then we calculated the average for all the paired comparisons (0.548 ± 0.0973), which indicates moderate turnover between ponds.

Despite the fact that rarefaction curves did not reach an asymptote (Fig. 4a), we observed that the coverage value for each gene library exceeded 80% and Chao1 showed that the predicted number of OTUs per site is similar to the one observed (Fig. 4b). Therefore, we conducted diversity comparisons among sampling sites through the calculation of different indices and used rarefaction curves as a comparative method of species richness from samples of different sizes. For the majority of indices (Shannon and Simpson) and rarefaction curves (Fig. 4a), A and B ponds were the most diverse sites and F pond the less diverse (Fig. 4b).

Low environmental variation across aquatic ponds

Environmental parameters among ponds did not show major differences. For example, conductivity measures ranged from 1940 to 2883.2 μS , and pH ranged from 7.53 to 7.81 (Table 1). Due to the extreme oligotrophy of these sites, some nutrients were in such low concentration that could not be detected (Table 1). As shown in the PCA analysis including all environmental variables, there is some environmental variation across ponds (Fig. 5). The results of PCA indicate that the first component explains 69.19 % of the variation dividing ponds in two: the F pond in one side and A, B, and C ponds in the other, being the F pond slightly different from the others in conductivity, salinity, TDS, IC and DOC (values in Table 1).

Neutral factors are influencing Pozas Azules bacterial assemblies

To explore the association between OTUs composition and environmental parameters, we followed a canonical correspondence analysis (CCA) using OTU abundance and environmental parameters data for each pond. The original 15 environmental parameters (Table 1) were reduced to six non-redundant explanatory variables (except geographic location), based on the three statistical criteria described in methods. The final environmental variables used for the analyses were total carbon (TC), inorganic carbon (IC), total phosphorus (TP), conductivity (Cond), nitrite (NO₂) and pH. Many OTUs were specifically associated to only one pond (unique), and corresponded to specific environmental variables. For example, overlapping OTUs in the upper left of Figure 6 correspond with the highest levels of CT found in pond A. In contrast, there are OTUs distributed across all axes in the graphic, which indicates that they do not correspond with any specific environmental gradient or pond; these OTUs were present in more than one pond.

To determine the relative importance of geographic (neutral) and contemporary environmental conditions (niche-based) in microbial diversity distribution, we conducted correlational analyses using environmental and geographical parameters, and OTUs incidence data. Environmental and geographic distance matrices were independently tested for correlation with the biological distance matrix (differences among sites represented by the dissimilarity index 1-Sorensen), using Mantel tests. The biological distance matrix did not show significant association with the environmental matrix ($p=0.1671$, $r=0.5876$). In

contrast, the biological distance matrix was significantly correlated with geographic distance ($p=0.04117$, $r=0.7979$). To further investigate the environmental variables that best correlated with community differences (i.e., the biological matrix), we performed a Bioenv test that identified TC, NO₂, pH and conductivity as the best-fitted parameters; however, a Mantel correlation for these parameters was not statistically significant ($p=0.0862$, $r=0.1635$).

Overall, the presented evidence shows statistical correlation between bacterial composition and geographic distance.

Discussion

Diversity distribution studies have helped to understand the influence of geographic (neutral) and environmental (niche-based) factors in biogeographic patterns. For many years, microbial diversity studies in a biogeographical context were inaccessible, hindering the understanding of the ecological and evolutionary processes shaping diversity at the microbial scale. Technological advances developed in the past two decades have allowed the exploration of microbial diversity in natural environments and to test hypotheses regarding factors determining its distribution.

To properly investigate the causes of current diversity distribution by the analysis of biogeographical patterns, sampling design is critical and requires samples from locations that are geographically delimited, and have low environmental variation to avoid obscuring distance (neutral) effects (Bissett et al. 2010, Logares et al. 2013).

In this study, we followed a culture independent approach with a sampling strategy that minimized environmental variation by looking at bacterial diversity in four ponds within an aquatic system known as Pozas Azules in the CCB, México. Our findings show for the first time in the CCB the influence of neutral or historical factors in prokaryotic diversity distribution.

OTU composition: moderate turnover and dominance by one particular phylum

The phylogenetic analysis of microbial diversity showed one numerically dominant phylum in each pond that could nonetheless be found in all the other ponds in lower numbers, which reflects the reported complexity of bacterial community assembly mechanisms: “*no two, naturally assembled bacterial communities appear to be the same*” (Curtis & Sloan 2004).

The average beta diversity between ponds indicates that near 50% of OTUs were shared among them. These shared *taxa* (many of them also abundant), can be explained because they are widely dispersed and known as generalist bacteria (Langenheder & Székely 2011). The same phyla were previously reported in aquatic systems of the CCC valley (Escalante et al. 2008, Pajares et al. 2012, 2013), as well as in other freshwater systems (e.g. Zhao et al. 2012). Specifically, Actinobacteria have been described as widely distributed in freshwater lakes (Glöckner et al. 2000, Warnecke et al. 2004), because of dispersal capacity and spore formation (Newton et al. 2011). Another generalist group, shared among the studied ponds, is Cyanobacteria, of which Chlorococcales (Pajares et al. 2012) is the most abundant order in freshwater lakes (Newton et al. 2011). Alphaproteobacteria (Rhodobacterales) has also been

reported with high abundance world wide (Fu et al. 2013), and Betaproteobacteria as predominant in mesocosms experiments by Langenheder and Székely (Langenheder & Székely 2011). The most abundant Verrucomicrobia OTU in our study is *Luteolibacter*, a non-motile bacteria, reported with a wide distribution (Yoon et al. 2012).

In contrast to the dominant and shared phyla, other groups like Bacteroidetes, although shared, were not abundant, despite being registered as ubiquitous in aquatic environments (O'Sullivan et al. 2006). We also found OTUs that were exclusive to one sampling site, nonetheless most of these OTUs are also usually considered common generalist *taxa* of water bodies (e.g. Rhodobacteriaceae, OTU B_10) (Allers et al. 2007, De Corte et al. 2013).

Overall, either shared or not, the generalist *taxa* were found in different abundances across ponds. Langenheder and Székely (2011) suggested that generalist groups usually represent neutrally assembled *taxa*, i.e., neutral processes may be predominant if most *taxa* correspond to generalist groups.

The local factors in Pozas Azules system: low environmental variation

Environmental conditions of the sampled ponds do not show major variation among them. The similarity of environmental parameters among ponds is key to statistically disentangle the influence of neutral processes (geographic distance) from environmental variables in diversity distribution (Bissett et al. 2010, Logares et al. 2013).

The 15 measured environmental parameters were reduced to six non-redundant explanatory variables. Of these, the four variables selected by Bioenv test (TC,

Cond, NO₂ and pH), although not statistically correlated with community dissimilarities among ponds, are variables reported as important elements shaping bacterial lake communities (Logue & Lindström 2008, Yang et al. 2013). However, given our sampling strategy where environmental variation was selected to be minimal among sites, no statistical correlation of these variables with assemblage dissimilarity was expected. Despite low environmental variability, we did observe a deviation of one pond (F) from the other three (Figure PCA), which accounts for the possibility that with finer measurements at smaller spatial scales, environmental variation can be detected and in turn explain some observations such as differences in abundance of some phyla among ponds.

Neutral factors are influencing Pozas Azules bacterial assembles

Our sampling strategy allowed us to independently evaluate the correlation between community dissimilarity among ponds and geographic distance or environmental variables.

Regarding environmental variables, Mantel correlation test was not statistically significant, and in the CCA analysis most OTUs did not show associations with any specific environmental condition. These non-conclusive correlations and associations were expected, given that environmental variability was chosen to be minimal. With respect to geographic effects on community composition, our results showed association of the aquatic bacterial composition with geographic distance among the ponds.

Despite several studies in the CCB focused in describing bacterial diversity (Souza et al. 2006, Escalante et al. 2008, Cerritos et al. 2011, Rodríguez-Verdugo et al. 2012, Bonilla-Rosso et al. 2012, López-Lozano et al. 2012, 2013, Peimbert et al. 2012, Pajares et al. 2012, 2013, Rebollar et al. 2012), this is the first time that environmentally similar sites have been analyzed to specifically determine the contribution of neutral factors (history, geographic distance) vs. niche-based factors (physicochemical conditions) in the observed distribution of microbial diversity. Two previous studies described aspects of the microbial diversity from Pozas Azules (Pajares et al. 2012, 2013), but their focus was in analyzing the effect of environmental changes in the composition and diversity of bacterial communities in a mesocosms long-term experiment. Pajares and colleagues (Pajares et al. 2012, 2013) sampling was conducted two years later than the study described in this paper, including one of the ponds here analyzed, and it is worth to note several coincidences on the reported OTUs ($\approx 50\%$ of this study OTUs are reported in Pajares bacterial communities). The fact that we found coincidences in OTU composition with Pajares, suggests that there could exist a common pool of successful bacterial *taxa* that are also constant over time, supporting our idea that these *taxa* are generalist, which could result in the neutrally assembled dynamics that we are observing (Langenheder & Székely 2011, Logares et al. 2013). These results also suggest that community dynamics in these ponds is associated with life history traits, as reported by Nemergut et al. (2013) which are dependent on dispersion capacity, and other bacterial ecological attributes (generalist or specialist). In this sense, rare and slow

growing *taxa* are reported with no geographic association in aquatic communities (Langenheder & Székely 2011), such as preliminary data for Planctomycetes in our samples (data not-shown), a slow growing phylum (Lage & Bondoso 2012), that, as stated, is departing from the assumptions of the neutral model (Logares et al. 2013).

The geographical pattern found in this study is consistent with Martiny's 3rd hypothesis that proposes the existence of limits to dispersal of bacteria, and states that this dispersal history is reflected in the correlation of beta diversity coupled with geographic distances. Behind this pattern, the dynamics of the community is mainly driven by rates of loss (extinction or emigration) or gain (immigration or speciation) of *taxa* in the community, as recognized by Hubbel's neutral theory (Hubbell 2001).

The geographical association was not statistically significant when using beta diversity indexes that contemplate abundance (Morisita-Horn or chord distance, data not shown). This fact supports the idea that using presence- absence data is a form to homogenize life histories in a community, cancelling out the effect of some specific group. Particular groups (hence, particular life histories) are better represented when using abundance-type beta diversity indexes, but considering relative abundances of groups makes the detection of geographic effects difficult, since neutral factors are only found when life history differences are excluded. As He and Hu stated "*neutral theory (...) is a "mean field" theory that describes what communities are expected to be like if all species obeyed the mean*" (He & Hu 2005). Moreover, the Pozas Azules system allowed us to detect signals of

neutral processes, since it has been already reported that unconnected freshwater bodies where dispersal is limited are systems that allow to detect similar patterns such as distance-decay relationship (also shaped by neutral processes) (Hanson et al. 2012).

Conclusions and perspectives

We found evidence that neutral processes are of primarily influence in the composition of the aquatic bacterial communities in the Pozas Azules system. We were able to show the correlation between beta diversity (1-Sorensen) and geographic distance through the use of a sampling design that minimizes the effect of environmental variability. Many of the OTUs detected in Pozas Azules were similar to generalist *taxa*, which usually represent neutrally assembled *taxa* as previously reported (Langenheder & Székely 2011).

To further advance in the investigation of the causes of diversity distribution of prokaryotes, it would be necessary to simultaneously control for environmental variation (using for example both contrasting and similarly environmental sampling sites) to test the relative influence of both neutral and niche-based, factors (Martiny et al. 2006). It would be relevant to also incorporate the study of specialist or rare *taxa* within the total communities (i.e. using specific primers or phylochip designed approaches, and also a highthroughput sequencing approach that would increase the number of *taxa* sampled (e.g. Logares et al. 2013, Escalante et al. 2014). Finally, life-history aspects such as relative abundance, dispersion capacities, and ecological range (i.e. generalists vs.

specialist) are important traits to consider in future studies, in order to better understand the processes behind biogeographical patterns.

Acknowledgments

We are grateful to Felipe García Oliva (CIEco, Biogeoquímica de suelos) for nutrient analysis of the samples. We thank Juan Carlos Ramírez Gloria (PRONATURA), Eria Rebollar, Morena Avitia, Esmeralda López-Lozano, René Cerritos, Enrique Scheinvar and Germán Bonilla for field assistance, and we specially thank PRONATURA Noreste for the access to the Pozas Azules ranch. We also thank Juan Carlos Ibarra Flores from Área de Protección de Flora y Fauna Cuatro Ciénegas (APFFC) for letting us use their Pozas Azules aerial photograph. We are grateful to Eria Rebollar for reviewing the manuscript and to Laura Figueroa and Santiago Ramírez Barahona for their collaboration in the realization of the figures. We also thank Dr. Erika Aguirre for technical support and English review of the manuscript. The project was supported by Grants 0237A1 Secretaría de Educación Pública- Consejo Nacional de Ciencia y Tecnología (SEP-CONACyT) and World Wildlife Found (WWF) - Alianza Carlos Slim. This paper is part of the doctoral dissertation of the first author, who thanks the Posgrado en Ciencias Biomédicas (Universidad Nacional Autónoma de México) and Consejo Nacional de Ciencia y Tecnología (CONACYT), grant no. 113997, for financial support.

Conflict of Interests

The authors declare that they have no conflict of interests regarding the publication of this paper.

References

- Allers, E., L. Gómez-Consarnau, J. Pinhassi, J. M. Gasol, K. Simek & J. Pernthaler, 2007. Response of Alteromonadaceae and Rhodobacteriaceae to glucose and phosphorus manipulation in marine mesocosms. *Environmental microbiology*, 9: 2417–29.
- Bissett, A., A. E. Richardson, G. Baker, S. Wakelin & P. H. Thrall, 2010. Life history determines biogeographical patterns of soil bacterial communities over multiple spatial scales. *Molecular Ecology*, 19: 4315–4327.
- Bonilla-Rosso, G., M. Peimbert, L. D. Alcaraz, I. Hernández, L. E. Eguiarte, G. Olmedo-Alvarez & V. Souza, 2012. Comparative metagenomics of two microbial mats at Cuatro Ciénegas Basin II: community structure and composition in oligotrophic environments. *Astrobiology*, 12: 659–73.
- Cerritos, R., L. E. Eguiarte, M. Avitia, J. Siefert, M. Travisano, A. Rodríguez-Verdugo & V. Souza, 2011. Diversity of culturable thermo-resistant aquatic bacteria along an environmental gradient in Cuatro Ciénegas, Coahuila, México. *Antonie van Leeuwenhoek*, 99: 303–18.
- Curtis, T. P. & W. T. Sloan, 2004. Prokaryotic diversity and its limits: microbial community structure in nature and implications for microbial ecology. *Current opinion in microbiology*, 7: 221–6.
- Darriba, D., G. L. Taboada, R. Doallo & D. Posada, 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature methods*, 9: 772.
- De Corte, D., E. Sintès, T. Yokokawa & G. J. Herndl, 2013. Comparison between MICRO–CARD–FISH and 16S rRNA gene clone libraries to assess the active versus total bacterial community in the coastal Arctic. *Environmental microbiology reports*, 5: 272–81.
- DeSantis, T. Z., P. Hugenholtz, K. Keller, E. L. Brodie, N. Larsen, Y. M. Piceno, R. Phan & G. L. Andersen, 2006. NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic acids research*, 34: W394–9.

- Dodson, S., 1992. Predicting crustacean zooplankton species richness. *Limnology and Oceanography*, 37: 848–856.
- Dray, S. & A. Dufour, 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of statistical software*, 22: 1–20.
- Edgar, R. C., B. J. Haas, J. C. Clemente, C. Quince & R. Knight, 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27: 2194–2200.
- Elser, J. J., J. H. Schampel, F. Garcia-Pichel, B. D. Wade, V. Souza, L. Eguiarte, A. Escalante & J. D. Farmer, 2005. Effects of phosphorus enrichment and grazing snails on modern stromatolitic microbial communities. *Freshwater Biology*, 50: 1808–1825.
- Escalante, A. E., L. E. Eguiarte, L. Espinosa-Asuar, L. J. Forney, A. M. Noguez & V. Souza Saldivar, 2008. Diversity of aquatic prokaryotic communities in the Cuatro Ciénegas basin. *Fems Microbiology Ecology*, 65: 50–60.
- Escalante, A. E., S. Jardón-Barbolla, Lev Ramírez-Barahona & L. E. Eguiarte, 2014. The study of biodiversity in the era of massive sequencing. *Revista Mexicana de Biodiversidad*.
- Forti, P., I. Giulivo, M. Mecchia & L. Piccini, 2004. Las aguas de Cuatro Ciénegas. Pages 76–87 in G. Badino, T. Bernabei, A. De Vivo, I. Guiulivo, & G. Savino (eds.) *Bajo el desierto. El Misterio de las aguas de Cuatro Ciénegas*, 1era edition. La Venta - Esplorazioni Geografiche, Italia.
- Fu, Y., K. F. Keats, R. B. Rivkin & A. S. Lang, 2013. Water mass and depth determine the distribution and diversity of Rhodobacterales in an Arctic marine system. *FEMS microbiology ecology*, 84: 564–76.
- Gillet, F., 2007. *Panelutils.R*.
- Glöckner, F. O., E. Zaichikov, N. Belkova, L. Denissova, J. Pernthaler, A. Pernthaler & R. Amann, 2000. Comparative 16S rRNA analysis of lake bacterioplankton reveals globally distributed phylogenetic clusters including an abundant group of actinobacteria. *Applied and environmental microbiology*, 66: 5053–65.
- Good, I. J., 1953. The population frequencies of species and the estimation of population parameters. *Biometrika*, 40: 237–264.
- Guindon, S., J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk & O. Gascuel, 2010. New algorithms and methods to estimate maximum-likelihood

- phylogenies: assessing the performance of PhyML 3.0. *Systematic biology*, 59: 307–21.
- Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series*, 41: 95–98.
- Hanson, C. A., J. A. Fuhrman, M. C. Horner-Devine & J. B. H. Martiny, 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature reviews. Microbiology*, 10: 497–506. Nature Publishing Group.
- Hatosy, S. M., J. B. H. Martiny, R. Sachdeva, J. Steele, J. A. Fuhrman & A. C. Martiny, 2013. Beta diversity of marine bacteria depends on temporal scale. *Ecology*, 94: 1898–1904.
- He, F. & X.-S. Hu, 2005. Hubbell's fundamental biodiversity parameter and the Simpson diversity index. *Ecology Letters*, 8: 386–390.
- Hendrickx, J., 2012. Perturb: tools for evaluating collinearity.
- Hubbell, S., 2001. The unified neutral theory of biodiversity and biogeography. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press., Princeton, NJ.
- Huber, T., G. Faulkner & P. Hugenholtz, 2004. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics (Oxford, England)*, 20: 2317–9.
- Hurley, C., 2012. Gclus: clustering graphics.
- Johannesson, K. H., A. Cortés & K. C. Kilroy, 2004. Reconnaissance isotopic and hydrochemical study of Cuatro Ciénegas groundwater, Coahuila, México. *Journal of South American Earth Sciences*, 17: 171–180.
- Krebs C J, 1999. *Ecological methodology*, 2nd edition. Wesley Longman, New York, NY.
- Lage, O. M. & J. Bondoso, 2012. Bringing Planctomycetes into pure culture. *Frontiers in microbiology*, 3: 405.
- Lane, D. J., 1991. Nucleic acid techniques in bacterial systematics. Pages 115–175 *in* E. Stackebrandt & M. Goodfellow (eds.) *Nucleic acid techniques in bacterial systematics*. John Wiley and Sons, New York, NY.

- Langenheder, S. & A. J. Székely, 2011. Species sorting and neutral processes are both important during the initial assembly of bacterial communities. *The ISME journal*, 5: 1086–94.
- Lindström, E. S., 2006. External control of bacterial community structure in lakes. *Limnology and Oceanography*, 51: 339–342.
- Lindström, E. S. & S. Langenheder, 2012. Local and regional factors influencing bacterial community assembly. *Environmental microbiology reports*, 4: 1–9.
- Logares, R., E. S. Lindström, S. Langenheder, J. B. Logue, H. Paterson, J. Laybourn-Parry, K. Rengefors, L. Tranvik & S. Bertilsson, 2013. Biogeography of bacterial communities exposed to progressive long-term environmental change. *The ISME journal*, 7: 937–48.
- Logue, J. B. & E. S. Lindström, 2008. Biogeography of Bacterioplankton in Inland Waters Biogeography of bacterioplankton in inland waters. *Freshwater Reviews*, 1: 99–114.
- López-Lozano, N. E., L. E. Eguiarte, G. Bonilla-Rosso, F. García-Oliva, C. Martínez-Piedragil, C. Rooks & V. Souza, 2012. Bacterial communities and the nitrogen cycle in the gypsum soils of Cuatro Ciénegas Basin, coahuila: a Mars analogue. *Astrobiology*, 12: 699–709.
- López-Lozano, N. E., K. B. Heidelberg, W. C. Nelson, F. García-Oliva, L. E. Eguiarte & V. Souza, 2013. Microbial secondary succession in soil microcosms of a desert oasis in the Cuatro Cienegas Basin, Mexico. *PeerJ*, 1: e47.
- Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27: 209–220.
- Martiny, J. B. H., B. J. M. Bohannan, J. H. Brown, R. K. Colwell, J. A. Fuhrman, J. L. Green, M. C. Horner-Devine, M. Kane, J. A. Krumins, C. R. Kuske, P. J. Morin, S. Naeem, L. Ovreås, A.-L. Reysenbach, V. H. Smith & J. T. Staley, 2006. Microbial biogeography: putting microorganisms on the map. *Nature reviews. Microbiology*, 4: 102–12.
- Nemergut, D. R., S. K. Schmidt, T. Fukami, S. P. O'Neill, T. M. Bilinski, L. F. Stanish, J. E. Knelman, J. L. Darcy, R. C. Lynch, P. Wickey & S. Ferrenberg, 2013. Patterns and processes of microbial community assembly. *Microbiology and molecular biology reviews* : MMBR, 77: 342–56.
- Newton, R. J., S. E. Jones, A. Eiler, K. D. McMahon & S. Bertilsson, 2011. A guide to the natural history of freshwater lake bacteria. *Microbiology and molecular biology reviews* : MMBR, 75: 14–49.

- Noguez, A. M., H. T. Arita, A. E. Escalante, L. J. Forney, F. García-oliva & V. Souza, 2005. Microbial macroecology : highly structured prokaryotic soil assemblages in a tropical deciduous forest. *Global Ecology and Biogeography*, 14: 241–248.
- O’Sullivan, L. A., J. Rinna, G. Humphreys, A. J. Weightman & J. C. Fry, 2006. Culturable phylogenetic diversity of the phylum “Bacteroidetes” from river epilithon and coastal water and description of novel members of the family Flavobacteriaceae: *Epilithonimonas tenax* gen. nov., sp. nov. and *Persicivirga xylanidelens* gen. nov., sp. *International journal of systematic and evolutionary microbiology*, 56: 169–80.
- Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O’Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens & H. Wagner, 2013. *Vegan: Community Ecology Package*.
- Pajares, S., G. Bonilla-Rosso, M. Travisano, L. E. Eguiarte & V. Souza, 2012. Mesocosms of aquatic bacterial communities from the Cuatro Ciénegas Basin (Mexico): a tool to test bacterial community response to environmental stress. *Microbial ecology*, 64: 346–58.
- Pajares, S., L. E. Eguiarte, G. Bonilla-Rosso & V. Souza, 2013. Drastic changes in aquatic bacterial populations from the Cuatro Ciénegas Basin (Mexico) in response to long-term environmental stress. *Antonie van Leeuwenhoek*, 104: 1159–1175.
- Peimbert, M., L. D. Alcaraz, G. Bonilla-Rosso, G. Olmedo-Alvarez, F. García-Oliva, L. Segovia, L. E. Eguiarte & V. Souza, 2012. Comparative metagenomics of two microbial mats at Cuatro Ciénegas Basin I: ancient lessons on how to cope with an environment under severe nutrient stress. *Astrobiology*, 12: 648–58.
- R-Core-Team, 2013. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ragon, M., M. C. Fontaine, D. Moreira & P. López-García, 2012. Different biogeographic patterns of prokaryotes and microbial eukaryotes in epilithic biofilms. *Molecular ecology*, 21: 3852–68.
- Ramette, A., 2007. Multivariate analyses in microbial ecology. *FEMS microbiology ecology*, 62: 142–60.
- Rebollar, E. A., M. Avitia, L. E. Eguiarte, A. González-González, L. Mora, G. Bonilla-Rosso & V. Souza, 2012. Water-sediment niche differentiation in ancient marine lineages of *Exiguobacterium* endemic to the Cuatro Ciénegas Basin. *Environmental microbiology*, 14: 2323–33.

- Rodríguez-Verdugo, A., V. Souza, L. E. Eguiarte & A. E. Escalante, 2012. Diversity across Seasons of Culturable *Pseudomonas* from a Desiccation Lagoon in Cuatro Ciénegas, Mexico. *International journal of microbiology*, 2012: Article ID 201389, 10 pages.
- Sambrook, J. & D. W. Russell, 2006. Purification of nucleic acids by extraction with phenol:chloroform. *CSH protocols*: doi: 10.1101/pdb.prot4455.
- Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. a Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn & C. F. Weber, 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology*, 75: 7537–41.
- Souza, V., L. E. Eguiarte, J. Siefert & J. J. Elser, 2008. Microbial endemism: does phosphorus limitation enhance speciation? *Nature reviews. Microbiology*, 6: 559–64.
- Souza, V., L. E. Eguiarte, M. Travisano, J. J. Elser, C. Rooks & J. L. Siefert, 2012. Travel, sex, and food: what's speciation got to do with it? *Astrobiology*, 12: 634–40.
- Souza, V., L. Espinosa-asuar, A. E. Escalante, L. E. Eguiarte, J. Farmer, L. Lloret, J. M. Rodríguez-martínez, X. Soberón, R. Dirzo, J. Elser, L. Forney, X. Sobero & J. J. Elser, 2006. An endangered oasis of aquatic microbial biodiversity in the Chihuahuan desert. *Proceedings of the National Academy of Sciences of the United States of America*, 103: 6565–70.
- Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei & S. Kumar, 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*, 28: 2731–9.
- Ter Braak, C. F., 1987. The analysis of vegetation-environment relationships by canonical correspondence analysis. *Vegetatio*, 69: 69–77. Kluwer Academic Publishers.
- Wang, J., J. Shen, Y. Wu, C. Tu, J. Soininen, J. C. Stegen, J. He, X. Liu, L. Zhang & E. Zhang, 2013. Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *The ISME journal*, 7: 1310–21. Nature Publishing Group.
- Wang, Q., G. M. Garrity, J. M. Tiedje & J. R. Cole, 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology*, 73: 5261–7.

- Warnecke, F., R. Amann & J. Pernthaler, 2004. Actinobacterial 16S rRNA genes from freshwater habitats cluster in four distinct lineages. *Environmental Microbiology*, 6: 242–253.
- Wolaver, B. D., L. J. Crossey, K. E. Karlstrom, J. L. Banner, M. B. Cardenas, C. G. Ojeda & J. M. Sharp, 2012. Identifying origins of and pathways for spring waters in a semiarid basin using He, Sr, and C isotopes: Cuatrociénegas Basin, Mexico. *Geosphere*, 9: 113–125.
- Wright, E. S., L. S. Yilmaz & D. R. Noguera, 2012. DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Applied and environmental microbiology*, 78: 717–25.
- Yang, J.-K., Z.-B. Cheng, J. Li & L.-H. Miao, 2013. Community composition of nirS-type denitrifier in a shallow eutrophic lake. *Microbial ecology*, 66: 796–805.
- Yoon, J., Y. Matsuo & H. Kasai, 2012. Full Paper *Portibacter lacus gen. nov., sp. nov.*, a new member of the family Saprospiraceae isolated from a saline lake, 4: 191–197.
- Zhao, D., R. Huang, J. Zeng, W. Yan, J. Wang, T. Ma, M. Wang & Q. L. Wu, 2012. Diversity analysis of bacterial community compositions in sediments of urban lakes by terminal restriction fragment length polymorphism (T-RFLP). *World journal of microbiology & biotechnology*, 28: 3159–70.
- Zheng, Y.-M., P. Cao, B. Fu, J. M. Hughes & J.-Z. He, 2013. Ecological drivers of biogeographic patterns of soil archaeal community. *PloS one*, 8: e63375.

Figures

Fig. 1 The Pozas Azules system. a) Map of México displaying the location of Cuatro Ciénegas Basin (CCB), and Pozas Azules system (squares) (modified from Rebollar et al. 2012). b) Aerial view of the C pond and nearest ponds. The circular form of the pools is visibly appreciated. Clear blue ponds (color not showed) that are associated with low conductivity values, such as C pond, were selected for this study (photograph used with permission of APFFCC). c) Spatial

distribution of the sampling sites at Pozas Azules ponds (ponds A, B, C and F), dashed lines show distance between them in kilometers (Google Earth image)

Fig. 2 Taxa description of the Pozas Azules bacterial communities. a) Neighbor joining tree of representative 16S rDNA gene sequences. The representative 48 OTUs of sequences are defined at 0.03 distance cutoff. Number after OTUs name correspond to sequence abundance (ML phylogeny gave very similar topology (data not shown). b) OTUs abundance bars. Each tone refers to different pond. c) Relative abundance (in percent) of bacterial phyla from the clone library data of the four ponds. Each bar refers to a different pond, and the number of clones obtained from each site (n) is indicated

Fig. 3 UPGMA dendrogram showing bacterial composition similarities among ponds. Branch lengths represent 1-Sorensen similarity coefficient, based on presence/absence of 16S rDNA sequences OTUs (detected at 97% cutoff). The numbers (height) are 1-Sorensen's values

Fig. 4. Diversity-related data of bacterial communities in Pozas Azules CCB. a) Rarefaction curves showing 95 %confidence interval of 16S rDNA gene sequences for each pond, displaying number of OTU detected (at 97% cutoff) versus number of sequences analyzed by pond. b) Table with different diversity data obtained with 16S rRNA sequences for each pond (OTUs at 97 % cutoff).

Figure 5. Biplot generated from a principal component analysis (PCA) of the standardized environmental variables. TC=total carbon, IC=inorganic carbon, NO₂=nitrite, TP=total phosphorous, Sal=salinity, DOC=dissolved organic carbon, TDS=total dissolved solids, Cond=conductivity and pH. The different ponds are shown in the graph (A, B, C and F). Vectors show scaled environmental variables. The first component of the PCA analysis accounted for 69.2% of the total variation, mainly involving DOC, Sal, IC, and conductivity

Fig. 6 Canonical correspondence analysis (CCA) modeling OTUs response to environmental parameters. Text correspond to OTUs names for the bacterial communities of each ponds (letters A, B, C and F), and vectors show scaled environmental variables. Six explanatory variables are shown (TC=total carbon, IC=inorganic carbon, Cond=conductivity, NO₂=nitrite, TP=total phosphorous and pH)

Supplementary figure 1. Pearson correlation coefficients for the standardized environmental variables. Diagonal boxes represent the environmental values registered for the four ponds. Down-left side boxes are the regression lines for the plotted data of each pair of variables; upper-right side boxes are the r values for the correlated variables (red for positive and blue for negative correlations). Asterisks represent statistically significant associated variables (* p= 0.05, ** p=0.01, *** p=0.001)

Tables

Table 1. Environmental data values obtained for each pond.

	A	B	C	F
Lat	N 26 49'24.4''	N 26 49'31.8''	N 26 49'34.7''	N 26 48'37.08''
Long	W 102 00'53.2''	W 102 01'5''	W 102 01'22.8''	W 102 00'47.64''
pH	7.76 (0)	7.53 (0)	7.712 (0.09)	7.81 (0.10)
Conductivity (μ S)	2501.6 (8.73)	2529 (3.67)	2883.2 (3.63)	1940 (0)
Salinity (mg/L)	8.70 (0.53)	8.51 (1.24)	7.66 (0.82)	3.18 (0)
TDS (ppt)	1.3 (0)	1.3 (0)	1.5 (0)	1 (0)
TC (mg/L)	120.265(0.54)	109.22 (14.03)	110.41 (19.36)	102.28 (3.84)
IC (mg/L)	120.265 (0.54)	117.51 (7.13)	111.28 (18.12)	41.87 (1.38)
DOC (mg/L)	0	0	0	60.4 (3.54)
TP (mg/L)	0.83 (0.090)	0.89 (0.018)	0.87 (0.051)	0.92 (0.025)
IP (mg/L)	0	0	0	0
DOP (mg/L)	0.83 (0.090)	0.89 (0.018)	0.87 (0.051)	0.92 (0.025)
TN (mg/L)	0	0	0	0
DIN (NH ₄) (mg/L)	0	0	0	0
DON (mg/L)	0	0	0	0
NO ₂ (mg/L)	0	0	0.0026 (0.0015)	0

Nutrient and other environmental data except coordinates are mean values, standard deviations are in parentheses. Lat=latitude, Long=longitude, TDS=total dissolved solids, TC=total carbon, IC=inorganic carbon, DOC=dissolved organic carbon, TP=total phosphorous, DOP=dissolved organic phosphorous, TN=total nitrogen, DIN=dissolved inorganic nitrogen in form of ammonium, DON=dissolved organic nitrogen.

Supplementary Table 1. Colldiag tests values obtained for the standardized environmental variables

Condition											
Index	Variance	Decomposition Proportions									
		Intercept	TDS	Cond	Sal	pH	TC	IC	DOC	TP	NO ₂
1	1	0.002	0.002	0.001	0.006	0.001	0	0.005	0.001	0	0.002
2	2.759	0.001	0	0	0.01	0.001	0	0.01	0.156	0	0.088
3	3.827	0	0	0	0.049	0	0	0.033	0.09	0	0.436
4	67.937	0.996	0.998	0.999	0.936	0.998	1	0.952	0.754	1	0.474

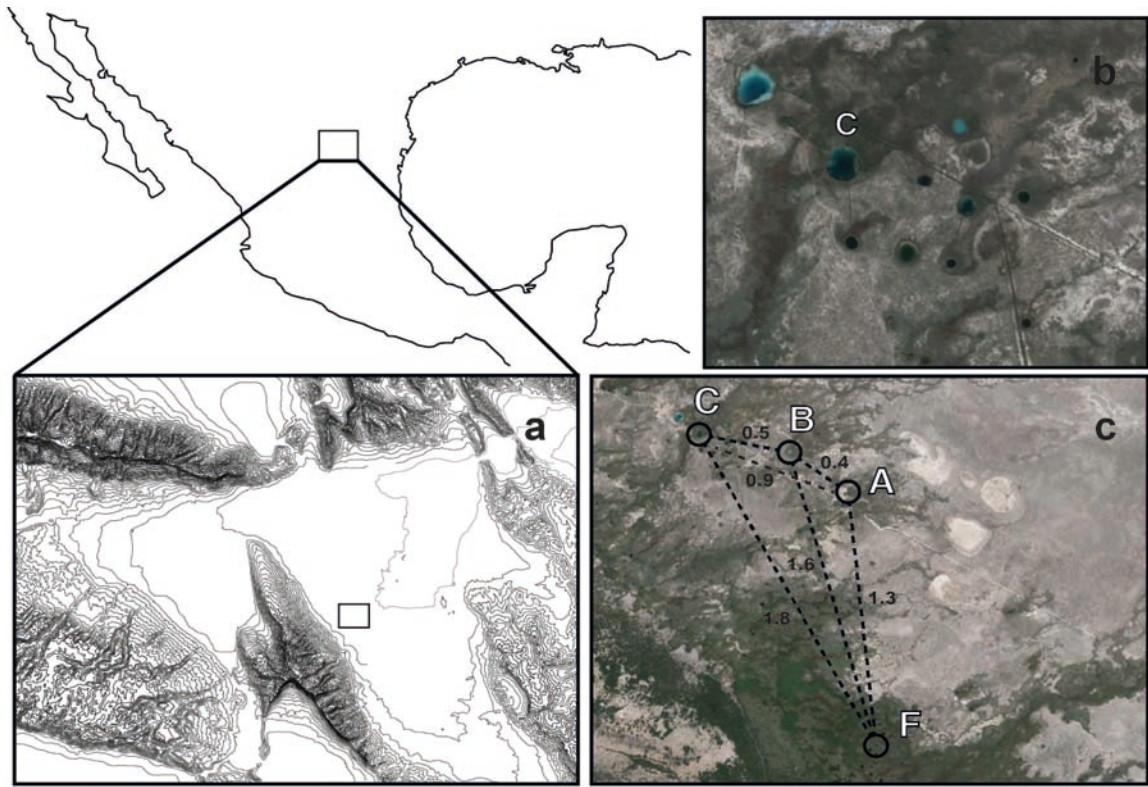
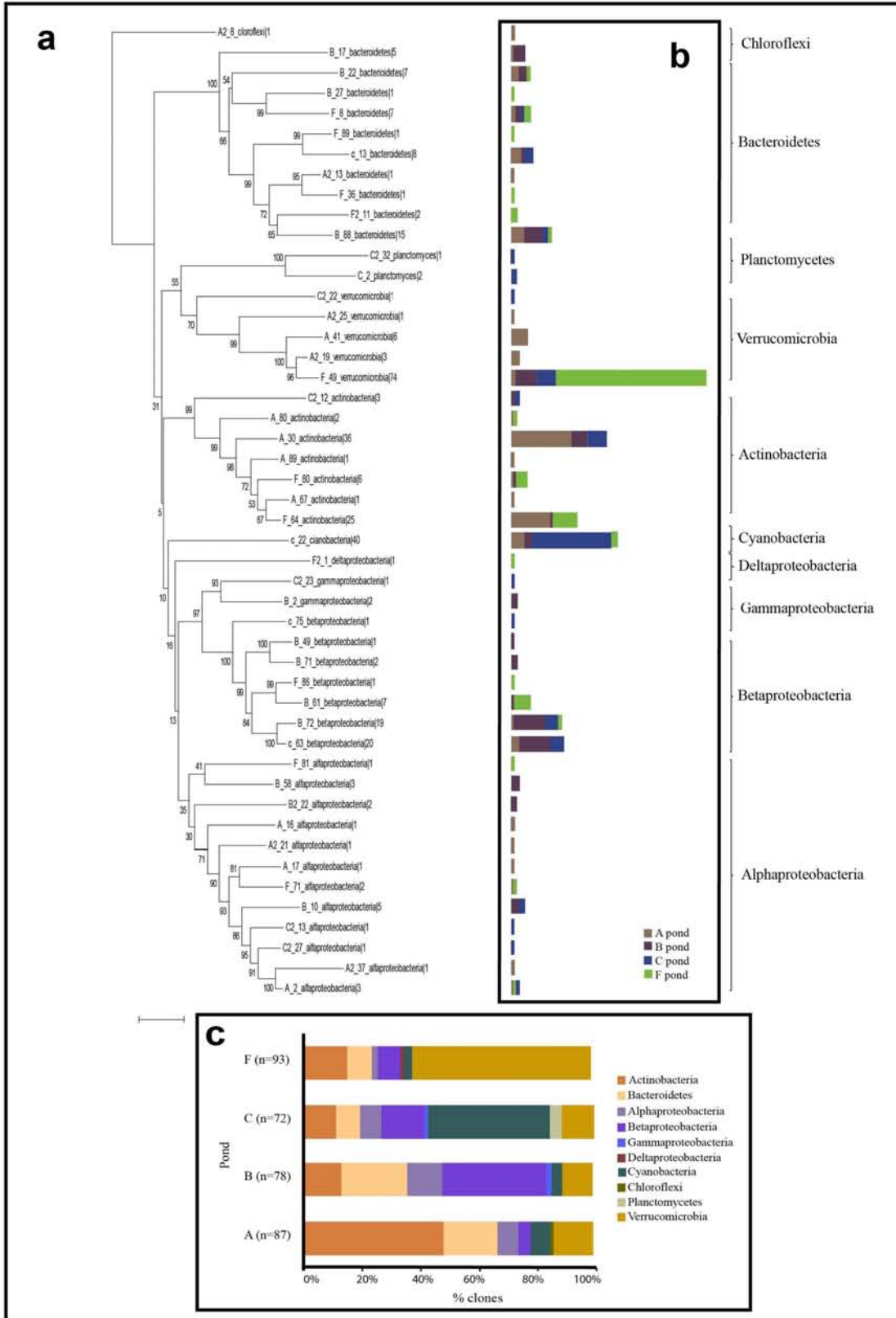


Figure 1



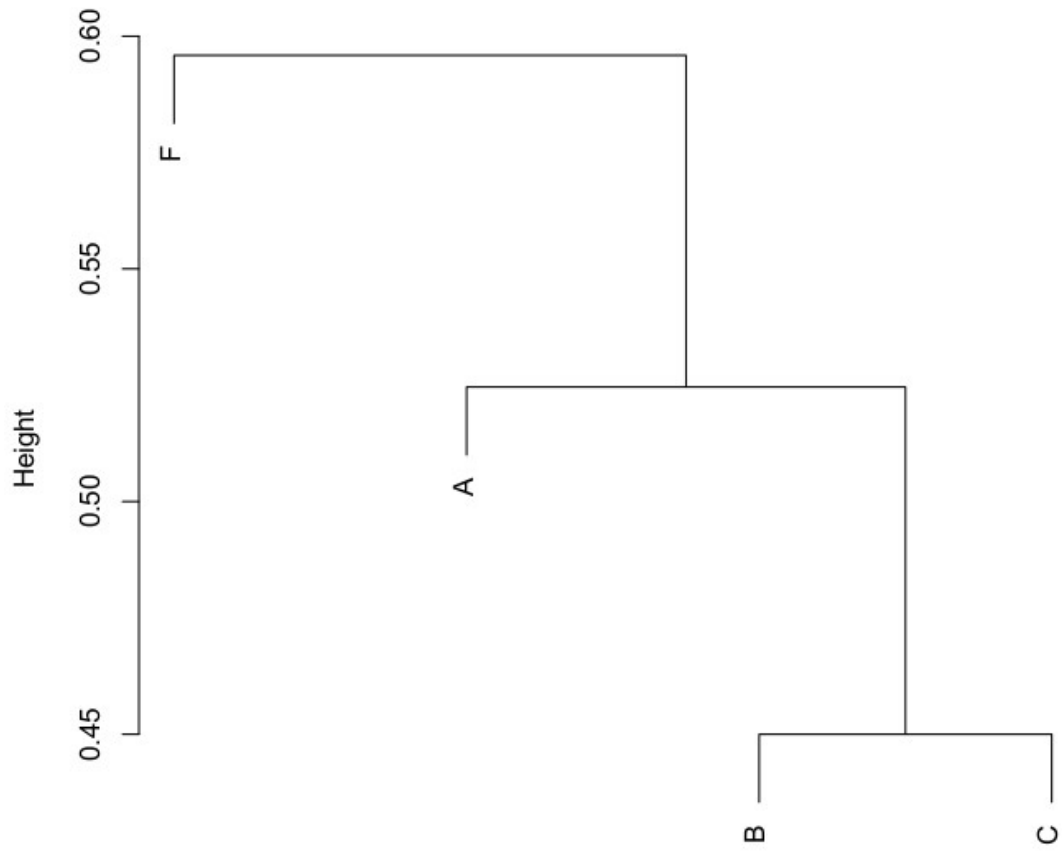


Figure 3

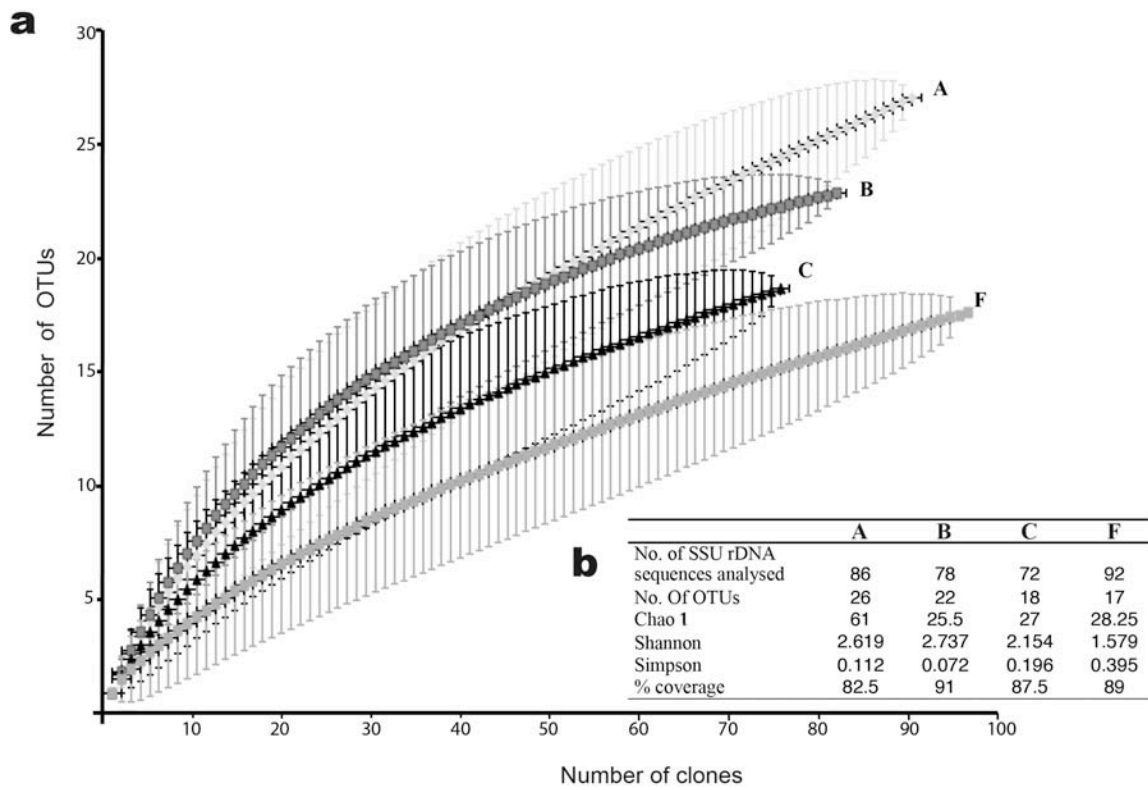


Figure 4

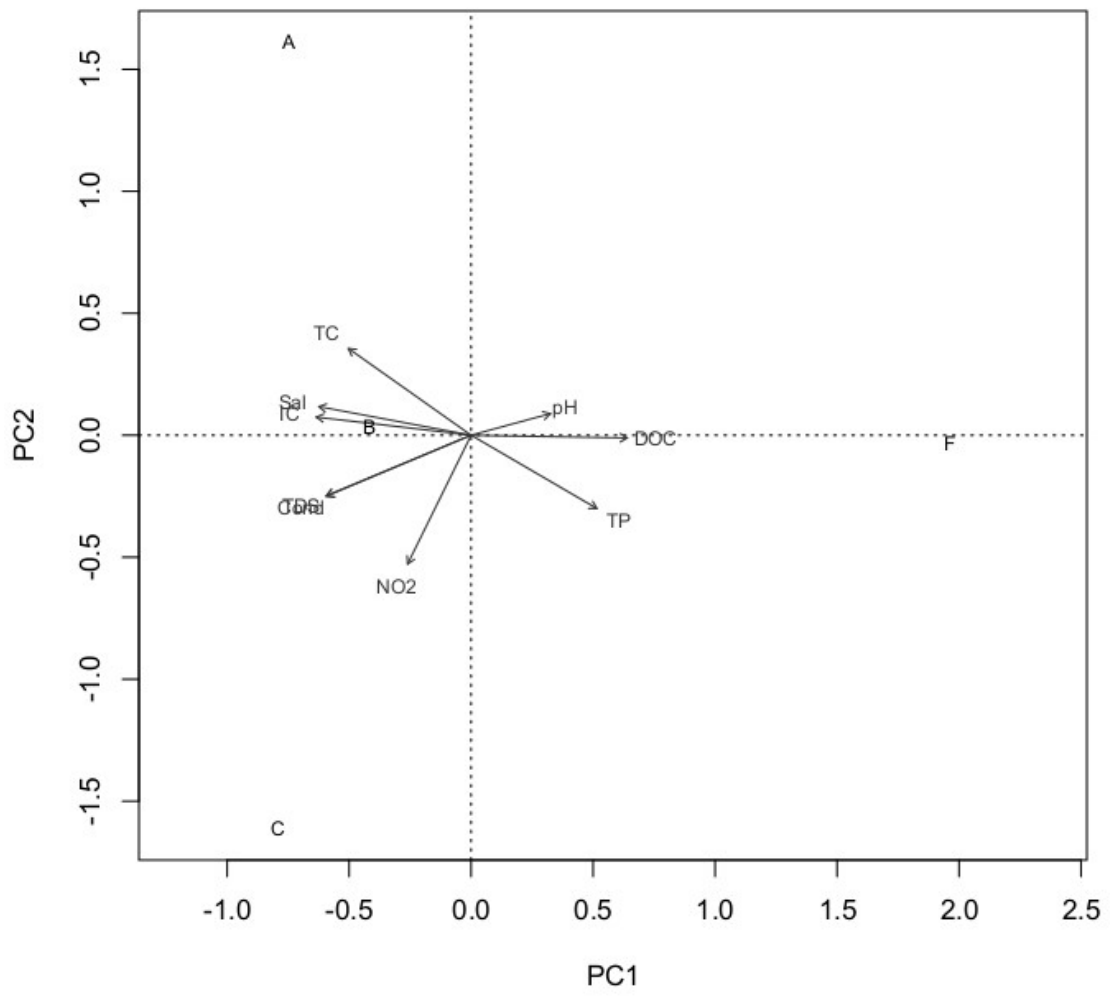


Figure 5

Species ordination

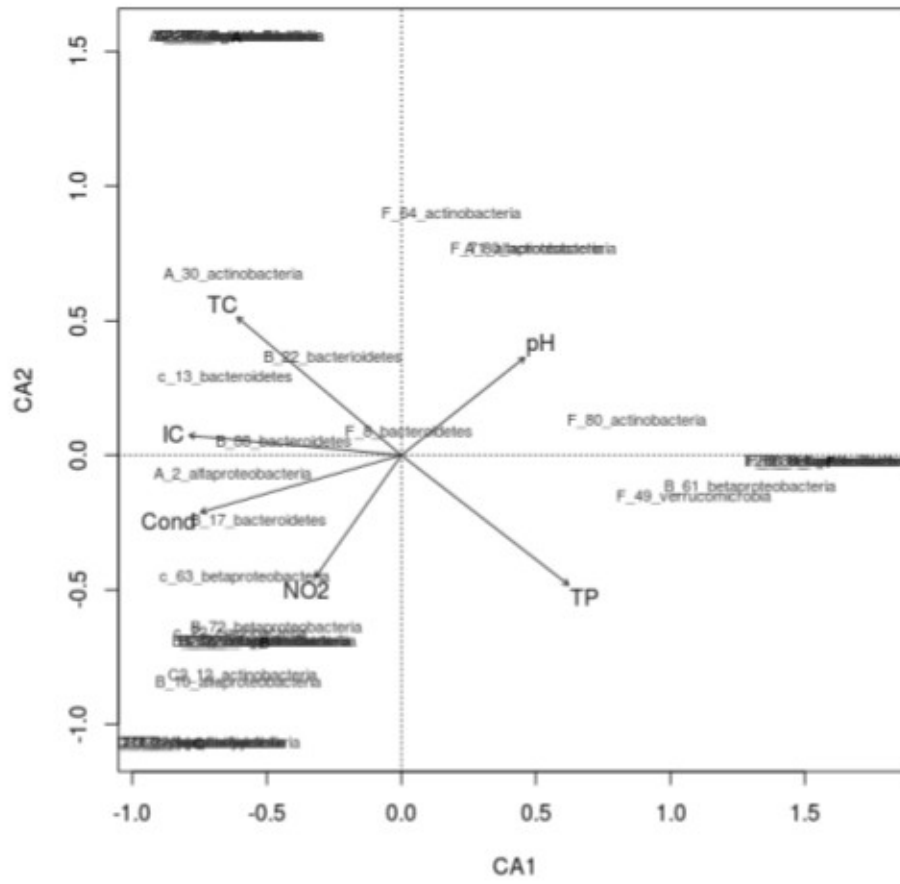
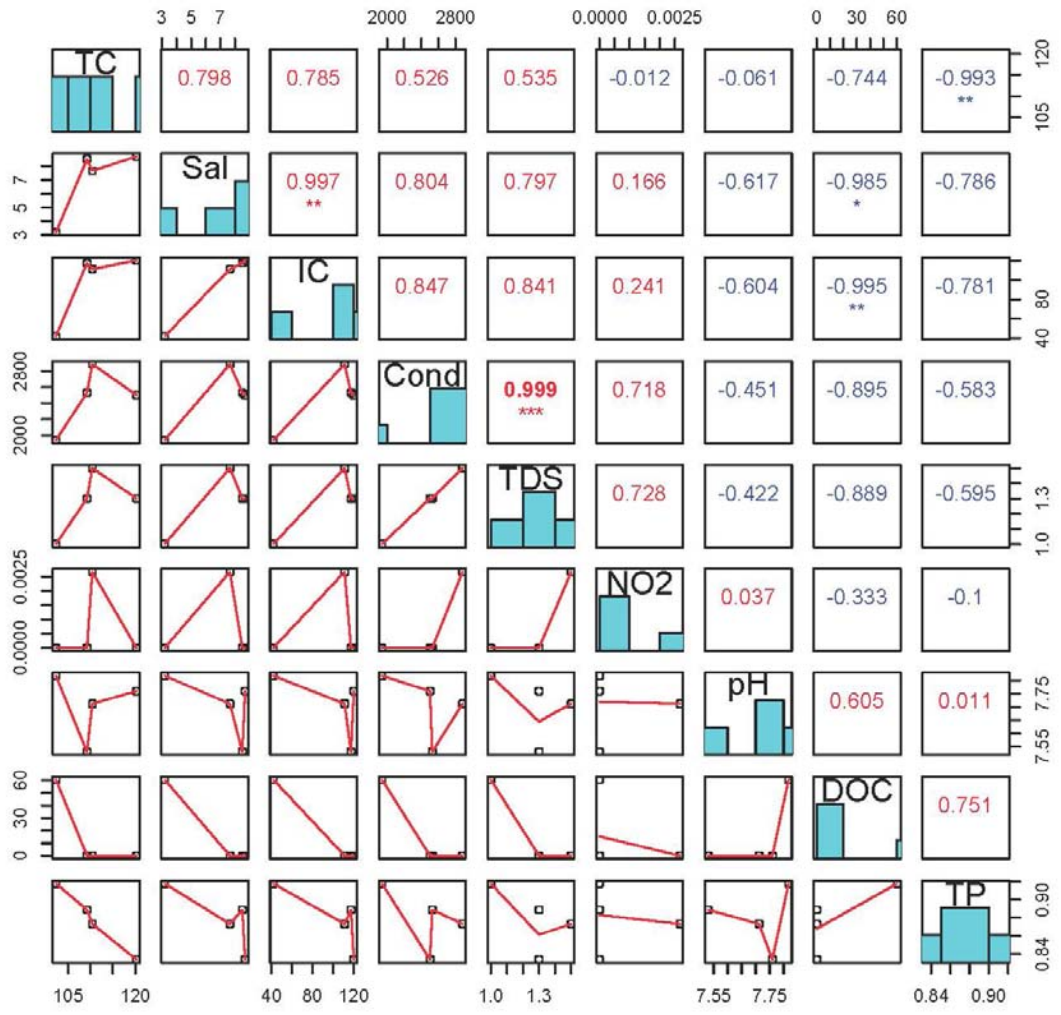


Figure 6

Pearson correlation matrix



Supplementary Figure 1

CAPÍTULO III

NOTAS SOBRE EL MUESTREO REGIONAL

Análisis de la distribución y diversidad de las bacterias acuáticas en el desierto Chihuahuense (Coahuila)

1 Objetivos

- Caracterizar la distribución y diversidad de comunidades acuáticas de bacterias en un muestreo a una escala regional (hasta 180 km) en el desierto Chihuahuense (Coahuila), utilizando la técnica de TRFLPs con el gen 16S ADNr.
- Correlacionar la composición de estas comunidades de bacterias con características ambientales (pH y conductividad) de los sitios de muestreo, así como con la distancia geográfica entre ellos.

2 Métodos

En abril del 2004 se muestrearon 27 sitios en distintos sistemas acuáticos. 10 muestras fueron tomadas en el valle de Cuatro Ciénegas, y 17 fuera de él, en el estado de Coahuila, a lo largo de un transecto aproximadamente circular de ca. 180 Km de radio (Figura 1 y Tabla 1).

Cada uno de los sitios muestreados fue geo-referenciado utilizando un GPS, y con una sonda Hydrolab YSI600R (YSI Inc., Yellow Springs, OH, USA), con la cual se hicieron mediciones de pH y conductividad *in situ*.

Se tomaron muestras de agua superficial (< 12cm de profundidad) en tres puntos distintos de cada sitio en cantidades iguales. Estas tres muestras se combinaron en botellas de plástico estériles de un galón, para formar una muestra compuesta para cada sitio.



Figura 1. Muestreo en un transecto de aprox. 180 km de diámetro. Los puntos señalan la posición de los 27 sitios de muestreo, y el número corresponde a la clave de la muestra tomada (ver Tabla 1 para especificaciones). Caracteres en blanco corresponden a sitios dentro del Valle de Cuatro Ciénegas, y en negro los que fueron muestreados fuera del valle.

Tabla 1. Especificaciones sobre los sitios de muestreo para comunidades microbianas acuáticas.

Muestra	Lugar	Tipo de muestra
1	CCC, manantial Churince	Manantial
2	CCC, sistema Pozas Azules, Poza Azul II	Poza
3	Municipio San Buenaventura	Pozo
4	Municipio San Buenaventura	Río
5	Valle del Hundido	Pozo
6	Valle del Sobaco	Noria 15m
7	Talía, pueblo cerca de Laguna de Mayrán	Agua del pueblo
8	Laguna de Mayrán	Laguna
9	Laguna de Mayrán	Laguna
10	Valle Ocampo	Arroyo
11	Valle Calaveras	Acequia
12	CCC	Acequia
13	CCC	Agua de llave
14	CCC	Canal con peces
15	CCC	Pozo a pivote
16	Cerca del pueblo El Rincón	Cascada entre montañas
17	Sierra de Australia	Charco tipo presón
18	Sierra de Australia	Charco tipo presón
19	Valle Ocampo	Río
20	Valle Ocampo	Agua de riego
21	Valle Calaveras	Pozo
22	CCC	Agua de pozo para uso doméstico
23	CCC	Depósito de agua de manantial
24	El Rincón	Pozo, agua del pueblo
25	CCC, Laguna Grande Churince	Laguna
26	Municipio San Buenaventura	Pozo
27	CCC, sistema Pozas Azules, Poza Azul I	Poza

Las muestras de agua fueron filtradas en fresco utilizando membranas de policarbonato 0.45 μm (Durapore, Merck Millipore ©). Los filtros se congelaron inmediatamente a -80° en un contenedor de nitrógeno líquido.

En el laboratorio se extrajo el ADN de las bacterias en los filtros (Apéndice 1). Para cada muestra se amplificó el gen 16S ADNr por PCR utilizando oligonucleótidos marcados con fluorescencia (Apéndice 2).

Los fragmentos amplificados fueron digeridos con la enzima HaeIII para generar patrones de restricción en cada muestra, los cuales fueron procesados con la técnica de T-RFLPs (Apéndice 3), lo que nos permite conocer los perfiles de las comunidades bacterianas a partir de los OTUs obtenidos. Un OTU (Unidad Taxonómica Operacional en español) se refiere a cada uno de los fragmentos de restricción del gen 16S ADNr que se generaron para cada una de las comunidades analizadas (ejemplo en Figura 2).

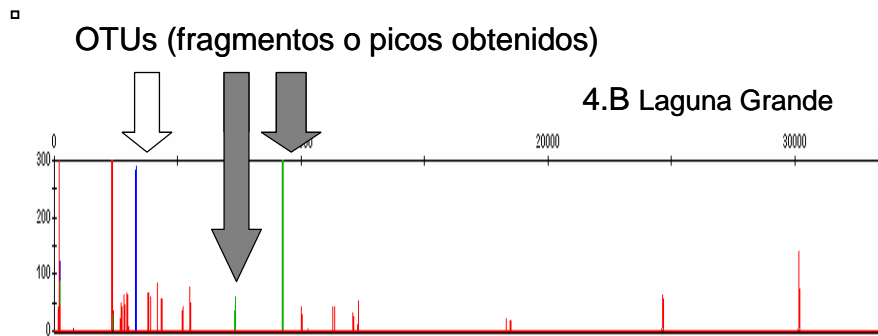


Figura 2. Perfil de TRFLPs para el sitio 25 (Laguna Grande en Churince, CCC): cada pico es equivalente a una banda de ADN en un gel. Las flechas en gris señalan los picos que fueron utilizados en este trabajo (color verde) y corresponden al oligonucleótido 27F, en el inicio del gen 16S ADNr. La flecha en blanco señala el fragmento obtenido con el oligonucleótido 1492, al final del gen. El resto de los picos (en rojo) corresponden al marcador de peso molecular.

Los fragmentos de T-RFLPs que se produjeron con el oligonucleótido 27F, fueron analizados en GenMapper y registrados en una tabla de abundancia, tomando en cuenta el tamaño (lo cual refleja el taxón) y la altura de cada fragmento de restricción (que refleja la abundancia relativa) (Fierer & Jackson, 2006). Estos datos fueron usados para calcular la diversidad beta en R (version 3.0.1 (2013-05-16) (R-Core-Team, 2013), utilizando el índice de disimilitud 1-Sorensen, que fue generado con la función `vegdist` del paquete `Vegan` (Oksanen *et al.*, 2013).

Los datos de pH y conductividad fueron transformados a z-score utilizando la función `decostand` (con el método “*standardized*”) también con `Vegan`, y con estos valores se construyó la matriz de similitud ambiental, calculando la distancia euclidiana entre las variables transformadas (utilizando la función `dist`). A partir de las dos matrices (disimilitud 1-Sorensen y la matriz de similitud ambiental) se construyeron los dos clusters de UPGMA también en R (función `hclust`). La matriz de distancia geográfica (en Km) se calculó a partir de los valores de latitud y longitud entre cada par de sitios.

Se realizaron pruebas de Mantel entre la matriz de disimilitud 1-Sorensen contra las otras dos matrices, la de similitud ambiental y la de distancia geográfica, usando la función `Mantel` (correlación de Pearson y 9,999 permutaciones), también en R.

3 Resultados

3.1 Descripción general

Para los 27 sitios muestreados se registró un total de 100 OTUs diferentes. Se obtuvieron de 2 a 24 OTUs para cada muestra, con un promedio de 11.4 OTUs por sitio. El 39% de OTUs estuvieron presentes una sola vez, y el resto pertenece a la proporción de OTUs que fueron compartidos por 2 y hasta en 23 sitios (Figura 3).

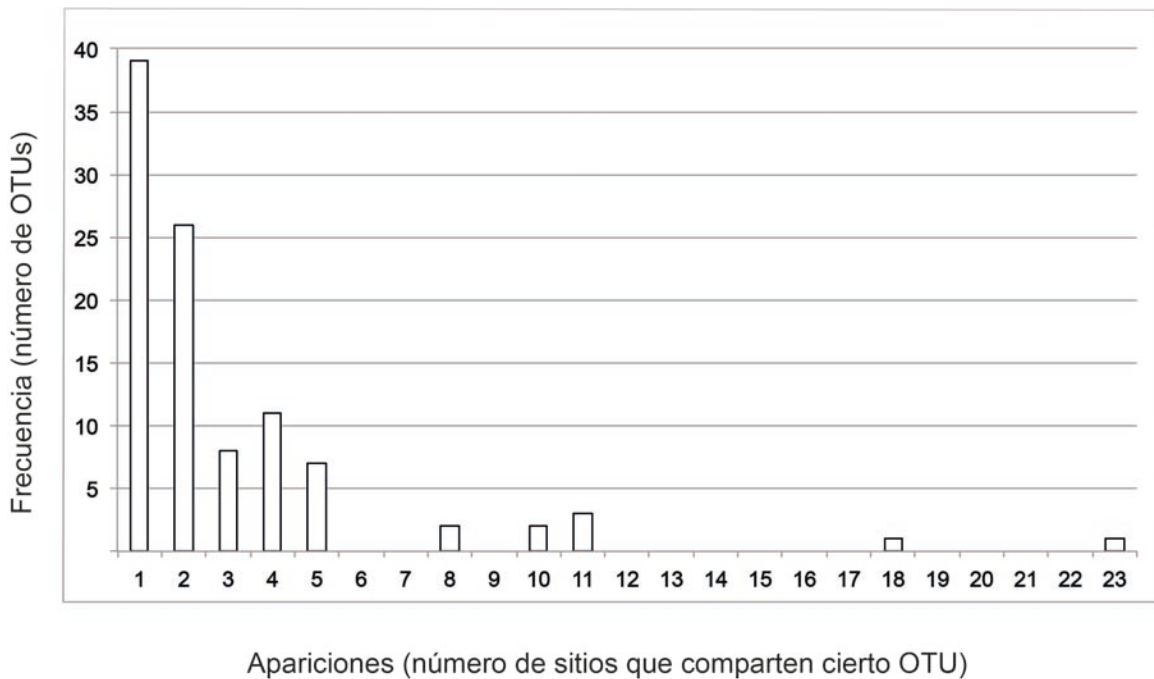


Figura 3. Frecuencias de aparición de OTUs. La suma de todas las frecuencias en esta gráfica corresponde al número total de OTUs obtenidos (100), representados en el eje y (número de OTUs). La mayoría de los OTUs o fragmentos estuvieron presentes en pocos sitios (de uno a cinco, representados en el eje x), y los fragmentos con una amplia distribución no son muy frecuentes, ya que sólo un número pequeño de OTUs fue registrado en muchos sitios (menos de cinco OTUs compartidos en 8 y hasta 23 sitios).

3.2 Análisis de *clusters*

El análisis de *clusters* usando el algoritmo UPGMA generado con la diversidad beta (1-Sorensen), representa claramente que la mayoría de los sitios de muestreo tienen una composición muy distinta entre sí, separados por distancias largas casi todos ellos (Figura 4). Sólo 7 de las 351 comparaciones resultaron menores a 0.5, y de esas 7, el valor más bajo obtenido fue de 0.3 .

Los sitios de muestreo 2 y 23, que corresponden a Poza Azul I y II en Pozas Azules, Cuatro Ciénegas, son los únicos sitios que son cercanos en composición de OTUS (i.e., tienen una mayor similitud de sus comunidades bacterianas) y geográficamente (ver Figura 1). Para el resto no hay una asociación fuerte con la distancia entre muestras y la similitud de la comunidad. Tampoco se observa alguna asociación entre las muestras que provienen dentro o fuera del valle de Cuatro Ciénegas, o entre los valles que podrían estar conectados geológicamente con CCC. En resumen, la composición en cada uno de los sitios es particular y pocos sitios son similares entre sí respecto a la composición de OTUs.

En cuanto al *cluster* que representa la similitud ambiental (Figura 5), estas mismas dos muestras tomadas en Pozas Azules que fueron similares en composición *beta* (2 y 27) también resultaron similares ambientalmente. Además hubo otras que ambientalmente son equivalentes en Cuatro Ciénegas (12 vs. 14, tomadas de agua que fluye en un canal y de una acequia, y 15 vs. 23 (pozo y depósito de agua).

Sobre las muestras fuera de CCC, resultaron ambientalmente más cercanas las muestras 17 vs. 18 (charcos en la Sierra de Australia) así como la 5 vs. 7 (un pozo en el Hundido y agua en Talía, pueblo cerca de Mayrán) y por último la 10 vs. 20 (agua de riego y un río en Ocampo). No se observa una asociación particular entre el tipo de muestras, tampoco entre sitios dentro y fuera de Cuatro Ciénegas, ni entre sitios cercanos geográficamente.

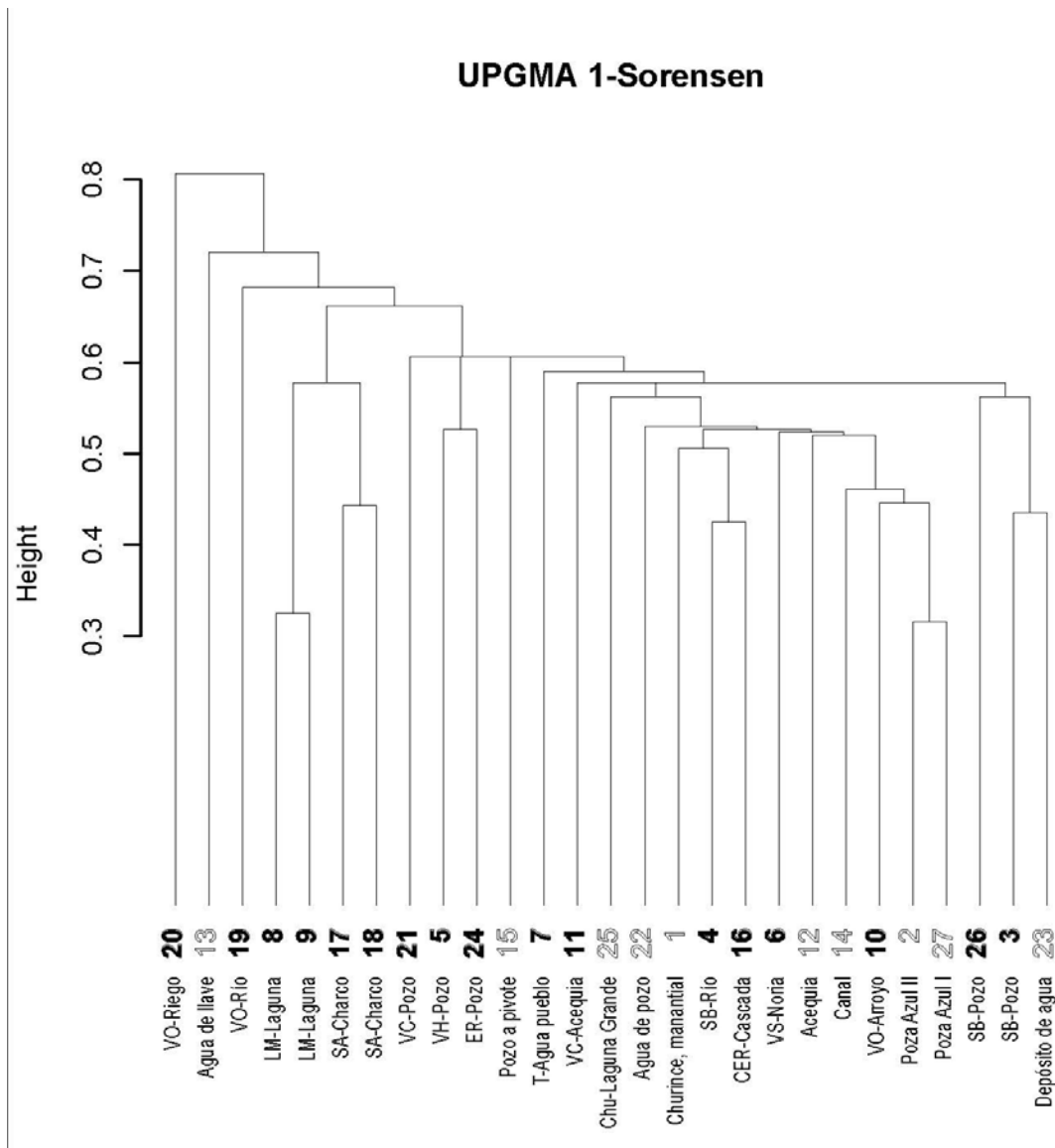


Figura 4. Análisis de *cluster* usando el algoritmo UPGMA construido con la matriz de disimilitud 1-Sorensen. Los números en negritas son sitios de muestreo fuera del valle de Cuatro Ciénegas (VO=Valle Ocampo, LM=Laguna de Mayrán, SA=Sierra de Australia, VC=Valle Calaveras, VH=Valle del Hundido, ER=El Rincón, T=Talía, SB=San Buenaventura, CER=Cerca de El Rincón, VS=Valle del Sobaco), y las muestras tomadas dentro del valle están representadas por los números más claros (Chu=Churince). La mayor parte de los sitios tiene distancias grandes entre ellos (cuyo valor está representado por la altura (*height*), lo que significa que la composición de los OTUs encontrados para cada sitio es muy diferente.

UPGMA ambiental

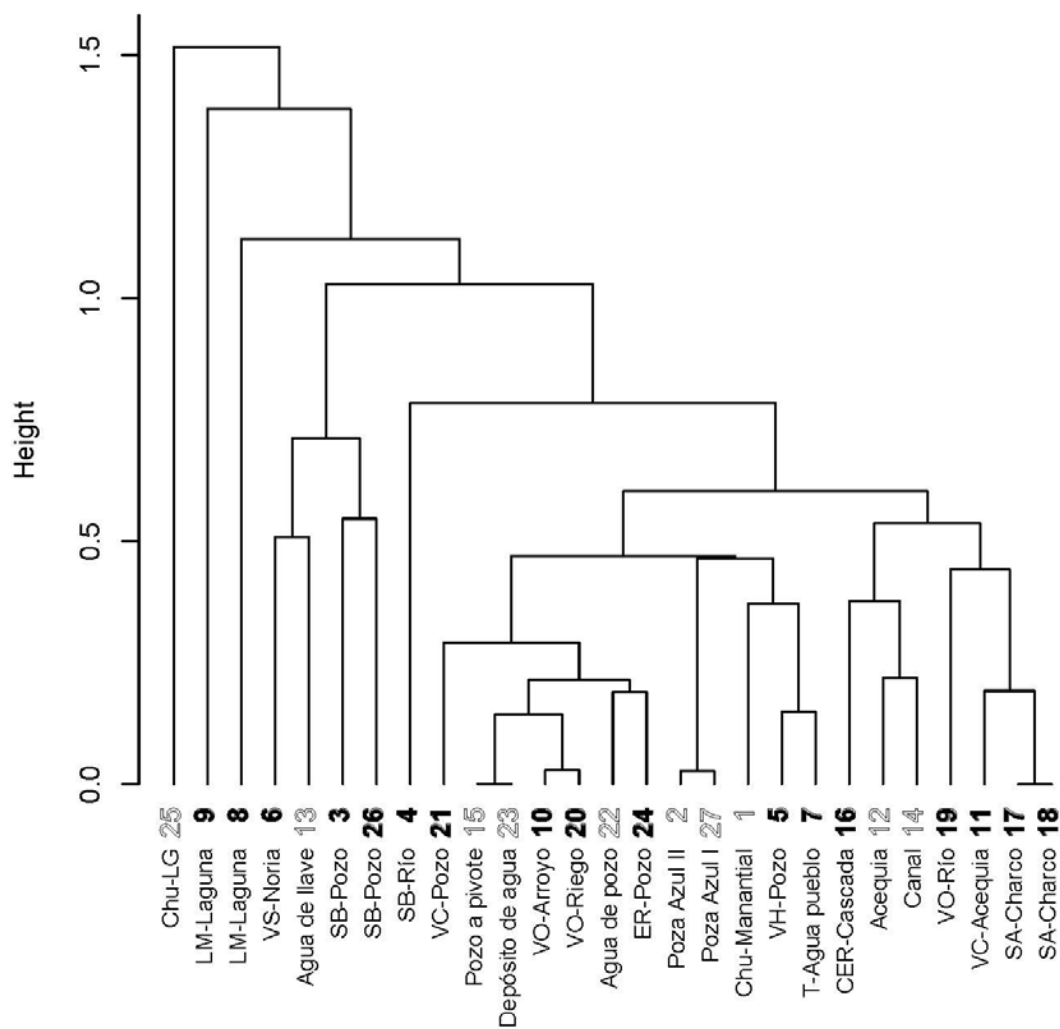


Figura 5. Análisis de *cluster* usando el algoritmo UPGMA construido a partir de la matriz ambiental (pH y conductividad). Al igual que en la Figura 4, los números en negritas son sitios de muestreo fuera del valle de Cuatro Ciénegas (VO=Valle Ocampo, LM=Laguna de Mayrán, SA=Sierra de Australia, VC=Valle Calaveras, VH=Valle del Hundido, ER=El Rincón, T=Talía, SB=San Buenaventura, CER=Cerca de El Rincón, VS=Valle del Sobaco), y las muestras tomadas dentro del valle están representadas por los números más claros (Chu=Churince, LG=Laguna Grande).

3.3 Patrones biogeográficos. Comparación entre diversidad *beta* (1-Sorensen) vs. variables ambientales y distancia geográfica

La prueba de Mantel entre la diversidad *beta* (1-Sorensen) contra variables ambientales (pH y conductividad) y distancia geográfica resultó significativa para

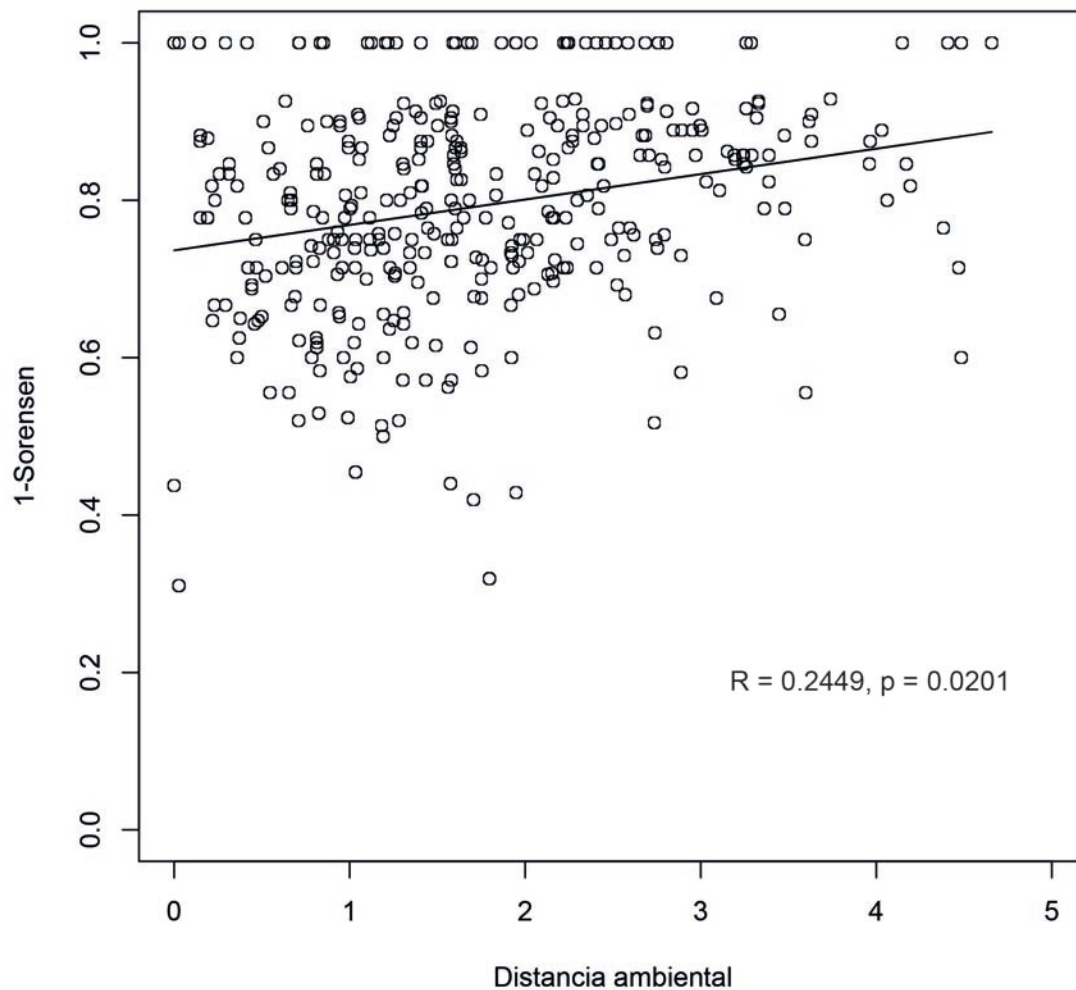


Figura 6. Diversidad beta (1-Sorensen) vs. distancia ambiental. Cada uno de los puntos graficados representa la comparación entre un par de sitios (para 27 sitios son 351 comparaciones). La posición de cada punto corresponde a la diferencia ambiental que hay entre los dos sitios y la disimilitud entre ellos (1-Sorensen). Se calculó también la regresión lineal correspondiente ($R^2 = 0.05729$, $p = 3.436 \times 10^{-6}$).

la matriz ambiental (Figura 6), lo que indica que los sitios ambientalmente similares, comparten mayor número de OTUs. No se encontró asociación con la distancia geográfica (Figura 7), lo cual indica que OTUs similares no se encuentran necesariamente cerca en el espacio.

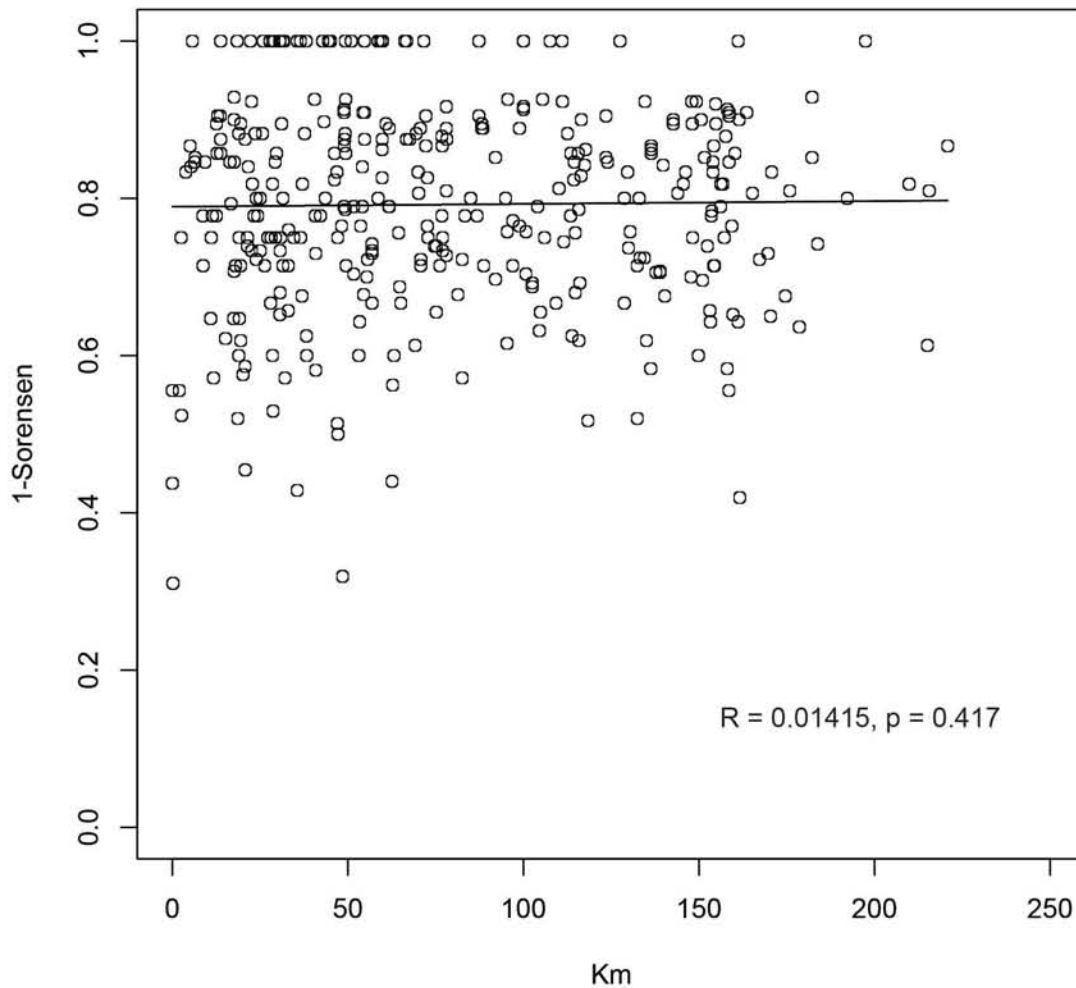


Figura 7. Diversidad beta (1-Sorensen) vs. distancia geográfica. Análogo a la Figura 6, cada uno de los puntos (rectángulos y círculos) corresponde a un par de sitios, cuya posición en este caso está dada por la distancia geográfica entre los dos, así como por la disimilitud entre ellos (1-Sorensen). Se grafica también la regresión lineal ($R^2 = -0.002665$, $p = 0.7917$).

4 Discusión

“La historia evolutiva y fragmentada de los desiertos en el mundo ha sido la fuerza que ha moldeado su rareza biológica, su adaptación a condiciones locales, y su especialización a los ambientes aislados. Después de millones de años en aislamiento, las fuerzas de evolución y fragmentación han llevado a formas únicas de vida en cada desierto, extrañas plantas y animales extraordinarios. Los desiertos del mundo son islas biológicas y culturales, tierras de aventura y fantasía, hábitats sorprendentes, en las que encontramos muy frecuentemente con enmarañadas formas de crecimiento y territorios de una inmensa belleza natural”
(Ezcurra, 2006) .

Cuatro Ciénegas es un valle anclado en el desierto Chihuahuense, y tal como lo describe Ezcurra (2006) refiriéndose a los desiertos, “de una inmensa belleza natural y con hábitats sorprendentes”. Ha sido investigado por nuestro grupo desde hace unos quince años, y en ese entonces ya existían estudios sobre los macroorganismos que ahí habitan. Estos estudios describen a un gran número de especies de animales y plantas endémicas, pero en aquellos años no existían datos sobre la diversidad microbiana del valle.

A lo largo de este tiempo, nuestro grupo de investigación se ha dedicado a analizar esta diversidad, con lo que se ha generado un número cada vez mayor de publicaciones utilizando enfoques diferentes, desde aquéllas que analizan poblaciones o comunidades de bacterias acuáticas (Avitia *et al.*, 2014, Cerritos *et al.*, 2011; Escalante *et al.*, 2008; Noguez *et al.*, 2008; Rebollar *et al.*, 2012; Rodríguez-Verdugo, *et al.*, 2012; Souza *et al.*, 2006) o describen tapetes microbianos (Bonilla-Rosso *et al.*, 2012; Peimbert *et al.*, 2012), hasta microcosmos en agua (Pajares *et al.*, 2013, 2012) o mesocosmos de suelo (López-Lozano *et al.*, 2013, 2012).

El presente trabajo, además de analizar comunidades de bacterias acuáticas dentro del valle de CCC, también considera otras comunidades acuáticas del desierto Chihuahuense que están ubicadas fuera del valle. Este diseño es una de las dos aproximaciones usadas para analizar la distribución de la diversidad de bacterias acuáticas de la presente tesis, en este caso utilizando una escala regional (180 km), y con sitios de muestreo que tienen grandes diferencias de pH y conductividad entre ellos. Se empleó una técnica

independiente de cultivo, que permite de una manera rápida y precisa comparar *grosso modo* varias localidades al mismo tiempo (Noguez *et al.*, 2005; Reed & Martiny, 2013), conocida como T-RFLPs (para una revisión sobre esta y otras técnicas moleculares, ver Espinosa-Asuar *et al.*, 2014).

Describir la diversidad bacteriana desde un enfoque ecológico permite aportar información sobre una de las grandes polémicas que existen en la literatura desde hace algunos años, que cuestiona si los microorganismos también están sujetos y responden a procesos ecológicos y evolutivos tal como lo hacen los macroorganismos (*e.g.* límites de dispersión) (O'Malley, 2008; Sul *et al.*, 2013), y se pregunta cuáles son las causas de sus patrones de distribución (Hanson *et al.*, 2012; Lindström & Langenheder, 2012; Martiny *et al.*, 2006; Nemergut *et al.*, 2013).

Los resultados obtenidos en este trabajo muestran que la distribución de la diversidad de las comunidades estudiadas dentro y fuera de CCC tuvo algunas semejanzas con patrones de distribución de macroorganismos. La primera similitud fue sobre el patrón que tiene la distribución de abundancias (Figura 3), con muchas especies (o en nuestro caso OTUs) presentes en pocos sitios, y pocas especies presentes en un gran número de sitios (McGeoch & Gaston, 2002; Pedrós-Alió, 2012), tal como se reporta para macroorganismos. Otra evidencia importante se obtuvo en el índice de disimilitud entre comunidades (diversidad *beta*), que resultó con un valor elevado (mayor a 0.5 para la mayoría de las comparaciones), lo que refleja una composición particular para cada uno de los sitios muestreados (Figuras 4 y 5). Estas dos evidencias reflejan cierto grado de aislamiento en las comunidades bacterianas muestreadas.

Como parte de la polémica mencionada, algunos grupos de trabajo (*e.g.* Finlay, 2002), basándose en postulados que han prevalecido desde hace casi un siglo en la ecología microbiana (Baas Becking, 1934), han sugerido que el tamaño pequeño de los microorganismos, así como sus grandes tasas de dispersión, producen una distribución que no es similar a la de los

macroorganismos, porque los microbios estarán “en todos lados”, y por lo tanto se esperarían comunidades que pueden tener una gran similitud entre ellas.

Las dos primeras evidencias del presente trabajo, apuntan en otro sentido. Sí se encontraron microorganismos (OTUs) que están “en todos lados”, pero al igual que los macroorganismos, la mayoría de los OTUs obtenidos presentan distribuciones restringidas, lo cual genera diferencias elevadas en la composición de los sitios muestreados, por lo que la disimilitud entre la mayor parte de los sitios fue mayor al 0.5 en casi todos los casos (Figura 4). En otros estudios similares, realizados en comunidades acuáticas de bacterias, también se ha reportado una composición diferente entre sitios. Como ejemplos podemos citar estudios en lagos belgas (Van der Gucht *et al.*, 2005), españoles (Reche *et al.*, 2005), en distintos sistemas acuáticos dentro del valle de Cuatro Ciénegas (Escalante *et al.*, 2008; Espinosa-Asuar *et al.*, n.d.), o en comunidades marinas de bacterioplancton (Ghiglione *et al.*, 2012), entre otros.

Por otro lado, si analizamos detenidamente las dos gráficas (Figuras 6 y 7) que correlacionan la diversidad *beta* (1-Sorensen) vs. la distancia geográfica o la diferencia ambiental, puede observarse un patrón interesante. Los puntos (comparaciones pareadas) que provienen de la comparación de sitios que resultaron con grandes diferencias ambientales o con una gran distancia entre ellos (aquéllos puntos con tendencia hacia la derecha de la gráfica), corresponden siempre a comunidades de bacterias cuya diferencia en composición (diversidad *beta*) también tuvo valores muy altos (y por lo tanto situados hacia arriba de las gráficas), esto es, no son comunidades similares. De esta forma no existen puntos que caigan en el área derecha y hacia abajo de ninguna de las dos gráficas, lo que significa que en este muestreo no hubo comunidades de bacterias que tuvieran grandes diferencias ambientales o distantes entre ellas, que además fueran similares en composición (con bajos valores de diversidad *beta*). Esto nos sugiere que existe una distribución de bacterias que no es al azar. Dicho de otra forma, una gran diferencia ambiental o una gran distancia geográfica entre sitios siempre estuvo asociada a comunidades que son muy distintas entre sí.

Respecto a las pequeñas diferencias ambientales o de distancia, tanto en la Figura 6 como en la 7 puede observarse que estos puntos (situados hacia la izquierda de las dos gráficas) están situados en un gradiente que va desde la parte alta hacia la parte baja de las gráficas.

En el caso de la distancia geográfica (Figura 7), esto se explica porque dos comunidades que son geográficamente cercanas pueden además ser ambientalmente distintas, y por lo tanto podrían tener comunidades muy distintas entre sí porque están adaptadas a dos ambientes contrastantes. Por el contrario, dos comunidades vecinas también podrían ser ambientalmente equivalentes y probablemente tendrán una composición (diversidad beta) similar. De esta forma, las comunidades de bacterias que son geográficamente cercanas pueden ubicarse entre estos dos extremos de posibilidades dependiendo del gradiente ambiental en el que se ubiquen, creándose este gradiente de puntos. En relación con este hecho, es importante recordar que ya ha sido reportado que grandes diferencias ambientales pueden enmascarar los efectos de la distancia geográfica (Bissett *et al.*, 2010; Logares *et al.*, 2013) tal como podría reflejarlo el gradiente de puntos que se observa en esta gráfica.

Para el caso de la gráfica ambiental ocurre algo similar, ya que comunidades ambientalmente equivalentes podrían estar geográficamente muy alejadas, por lo cual pueden tener una diferencia mayor en su composición, que aquéllas comunidades ambientalmente equivalentes, que estén geográficamente más cercanas. En este caso (Figura 7) el gradiente entre los puntos (esto es, la distribución hacia arriba y abajo de la gráfica) no resultó tan amplio como en el caso de la gráfica de distancia geográfica. Grandes diferencias ambientales pueden enmascarar los efectos de la distancia geográfica, pero no ocurrió al contrario, con lo cual la matriz ambiental resultó tener una correlación estadísticamente significativa con la matriz de composición (diversidad beta) de las comunidades bacterianas estudiadas en este trabajo.

Con esta misma metodología (T-RFLPs) también se han publicado patrones de diversidad o composición microbiana que están asociados a factores ambientales (Fierer & Jackson, 2006; Noguez *et al.*, 2008; Reed &

Martiny, 2013; Zhao *et al.*, 2012). Las dos variables analizadas, pH y conductividad, son factores ambientales reportados (entre otros) para ecosistemas acuáticos, como elementos que son determinantes en la estructura de la composición de comunidades de bacterias (Jones *et al.*, 2012) tal como se ha demostrado en este estudio.

Por otro lado, en algunos artículos que analizan la distribución de diversidad bacteriana utilizando la prueba de Mantel, es posible encontrar un patrón similar al descrito previamente (e.g. Bissett *et al.*, 2010; Logares *et al.*, 2012, Lear *et al.*, 2014). Sería interesante analizar este tipo de distribución desde una aproximación estadística más formal. Una opción es modelar los datos reportados en este tipo de estudios, para comprobar, por ejemplo, si estas comunidades con grandes diferencias ambientales y espaciales siempre están asociadas a una composición muy distinta, y si esta asociación es significativamente distinta a lo que se esperaría por azar.

Por último, es necesario mencionar que muchas de las técnicas moleculares conocidas como *fingerprinting*, así como también las librerías de clonas, y las nuevas tecnologías de secuenciación de nueva generación (NGS), cuando se hacen a baja cobertura, detectan únicamente a los componentes abundantes de una comunidad. En el caso de técnicas de *fingerprint*, detectan abundancias mayores al 1% de la comunidad total (Casamayor *et al.*, 2000, Lee, 2014), pero no son capaces de evaluar la contribución de los *taxa* raros (Espinosa-Asuar *et al.*, 2014; Jones *et al.*, 2012). De esta manera, es importante aclarar que las conclusiones de este y otros trabajos con metodologías similares, representan a los *taxa* abundantes de las comunidades. Existen pocos reportes que se enfocan al estudio de *taxa* raros, por lo cual estos patrones están hoy en día en proceso de describirse (Langenheder & Székely, 2011; Logares *et al.*, 2013).

5 Conclusiones

En resumen, no encontramos efecto de procesos neutros en las comunidades muestreadas. Por otra parte no se observaron patrones sencillos entre los sitios que se encuentran dentro y fuera de Cuatro Ciénegas. La composición obtenida para cada uno de los sitios es particular y pocos sitios son similares entre sí respecto a la composición de OTUs. Por último encontramos que la estructura de la comunidad de los sitios muestreados está influenciada por los factores ambientales medidos en el trabajo (pH y conductividad).

6. Bibliografía

- Acinas S. G., Sarma-Rupavtarm R., Klepac-Ceraj V., Polz M.F., 2005. PCR-Induced Sequence Artifacts and Bias: Insights from Comparison of Two 16S rRNA Clone Libraries Constructed from the Same Sample. *Appl. Env. Microbiol.* 71, 8966-8969.
- Avitia M, Escalante A.E., Rebollar E.A., Moreno-Letelier A., Eguiarte L.E. Souza V., 2014. Common population expansions in coexisting bacterial lineages revealed by genetic evidence. *PeerJ* in press.
- Baas Becking, L.G.M., 1934. *Geobiologie of inleiding tot de milieukunde*. Hague, Netherlands W.P. Van Stock. Zoon (in Dutch).
- Bissett, A., Richardson, A.E., Baker, G., Wakelin, S., Thrall, P.H., 2010. Life history determines biogeographical patterns of soil bacterial communities over multiple spatial scales. *Mol. Ecol.* 19, 4315–4327.
- Bonilla-Rosso, G., Peimbert, M., Alcaraz, L.D., Hernández, I., Eguiarte, L.E., Olmedo-Alvarez, G., Souza, V., 2012. Comparative metagenomics of two microbial mats at Cuatro Ciénegas Basin II: community structure and composition in oligotrophic environments. *Astrobiology* 12, 659–73.
- Casamayor E.O., Schafer H., Baneras L., Pedros-Alio C., Muyzer G. 2000. Identification of and spatiotemporal differences between microbial

assemblages from two neighboring sulfurous lakes: Comparison by microscopy and denaturing gradient gel electrophoresis. *Applied and environmental microbiology* 66: 499-508.

- Cerritos, R., Eguiarte, L.E., Avitia, M., Siefert, J., Travisano, M., Rodríguez-Verdugo, A., Souza, V., 2011. Diversity of culturable thermo-resistant aquatic bacteria along an environmental gradient in Cuatro Ciénegas, Coahuila, México. *Antonie Van Leeuwenhoek* 99, 303–18.
- Escalante, A.E., Eguiarte, L.E., Espinosa-Asuar, L., Forney, L.J., Noguez, A.M., Souza Saldivar, V., 2008. Diversity of aquatic prokaryotic communities in the Cuatro Cienegas basin. *Fems Microbiology Ecol.* 65, 50–60.
- Espinosa-Asuar, L., Escalante, A.E., Falcón, L.I., Bonilla-Rosso, G., Ramírez-Barahona, S., Eguiarte, L.E., Souza, V., 2014. : Comparación de tres métodos moleculares para el análisis de procariontes ambientales en el mar del canal de Yucatán. *Hidrobiologica* 24.
- Espinosa-Asuar, L., Escalante, A.E., Gasca-Pineda, J., Blaz, J., Peña, L., Eguiarte, L.E., Souza, V., n.d. Aquatic bacterial assemblage structure in Pozas Azules, Cuatro Ciénegas Basin in México: neutral vs. niche-based factors. Unpublished.
- Ezcurra, E., 2006. Natural history and evolution of the world deserts, in: Ezcurra, E. (Ed.), *Global Deserts Outlook*. San Diego Natural History Museum, San Diego, California.
- Fierer, N., Jackson, R.B., 2006. The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* 103, 626–31.
- Finlay, B.J., 2002. Global dispersal of free-living microbial eukaryote species. *Science* 296, 1061–3.
- Ghiglione, J.-F., Galand, P.E., Pommier, T., Pedrós-Alió, C., Maas, E.W., Bakker, K., Bertilson, S., Kirchmanj, D.L., Lovejoy, C., Yager, P.L., Murray, A.E., 2012. Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* 109, 17633–8.
- Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., Martiny, J.B.H., 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.* 10, 497–506.
- Jones, S.E., Cadkin, T. a, Newton, R.J., McMahon, K.D., 2012. Spatial and temporal scales of aquatic bacterial beta diversity. *Front. Microbiol.* 3, 318.

- Langenheder, S., Székely, A.J., 2011. Species sorting and neutral processes are both important during the initial assembly of bacterial communities. *ISME J.* 5, 1086–94.
- Lee J. 2014. Niche and neutral processes in aquatic bacterial communities: Are all things equal? PhD thesis. Lincoln University.
- Lindström, E.S., Langenheder, S., 2012. Local and regional factors influencing bacterial community assembly. *Environ. Microbiol. Rep.* 4, 1–9.
- Logares, R., Lindström, E.S., Langenheder, S., Logue, J.B., Paterson, H., Laybourn-Parry, J., Rengefors, K., Tranvik, L., Bertilsson, S., 2013. Biogeography of bacterial communities exposed to progressive long-term environmental change. *ISME J.* 7, 937–48.
- López-Lozano, N.E., Eguiarte, L.E., Bonilla-Rosso, G., García-Oliva, F., Martínez-Piedragil, C., Rooks, C., Souza, V., 2012. Bacterial communities and the nitrogen cycle in the gypsum soils of Cuatro Ciénegas Basin, coahuila: a Mars analogue. *Astrobiology* 12, 699–709.
- López-Lozano, N.E., Heidelberg, K.B., Nelson, W.C., García-Oliva, F., Eguiarte, L.E., Souza, V., 2013. Microbial secondary succession in soil microcosms of a desert oasis in the Cuatro Cienegas Basin, Mexico. *PeerJ* 1, e47.
- Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., Morin, P.J., Naeem, S., Ovreås, L., Reysenbach, A.-L., Smith, V.H., Staley, J.T., 2006. Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.* 4, 102–12.
- McGeoch, M. a, Gaston, K.J., 2002. Occupancy frequency distributions: patterns, artefacts and mechanisms. *Biol. Rev. Camb. Philos. Soc.* 77, 311–31.
- Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F., Knelman, J.E., Darcy, J.L., Lynch, R.C., Wickey, P., Ferrenberg, S., 2013. Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* 77, 342–56.
- Noguez, A.M., Arita, H.T., Escalante, A.E., Forney, L.J., García-oliva, F., Souza, V., 2005. Microbial macroecology: highly structured prokaryotic soil assemblages in a tropical deciduous forest. *Glob. Ecol. Biogeogr.* 14, 241–248.
- Noguez, A.M., Escalante, A.E., Forney, L.J., Nava-Mendoza, M., Rosas, I., Souza, V., García-Oliva, F., 2008. Soil aggregates in a tropical deciduous

- forest: effects on C and N dynamics, and microbial communities as determined by t-RFLPs. *Biogeochemistry* 89, 209–220.
- O'Malley, M.A., 2008. "Everything is everywhere: but the environment selects": ubiquitous distribution and ecological determinism in microbial biogeography. *Stud. Hist. Philos. Biol. Biomed. Sci.* 39, 314–25.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2013. *Vegan: Community Ecology Package*.
- Pajares, S., Bonilla-Rosso, G., Travisano, M., Eguiarte, L.E., Souza, V., 2012. Mesocosms of aquatic bacterial communities from the Cuatro Ciénegas Basin (Mexico): a tool to test bacterial community response to environmental stress. *Microb. Ecol.* 64, 346–58.
- Pajares, S., Eguiarte, L.E., Bonilla-Rosso, G., Souza, V., 2013. Drastic changes in aquatic bacterial populations from the Cuatro Ciénegas Basin (Mexico) in response to long-term environmental stress. *Antonie Van Leeuwenhoek* 104, 1159–1175.
- Pedrós-Alió, C., 2012. The Rare Bacterial Biosphere. *Ann. Rev. Mar. Sci.* 4, 449–466.
- Peimbert, M., Alcaraz, L.D., Bonilla-Rosso, G., Olmedo-Alvarez, G., García-Oliva, F., Segovia, L., Eguiarte, L.E., Souza, V., 2012. Comparative metagenomics of two microbial mats at Cuatro Ciénegas Basin I: ancient lessons on how to cope with an environment under severe nutrient stress. *Astrobiology* 12, 648–58.
- R-Core-Team, 2013. *R: A language and environment for statistical computing*.
- Rebollar, E.A., Avitia, M., Eguiarte, L.E., González-González, A., Mora, L., Bonilla-Rosso, G., Souza, V., 2012. Water-sediment niche differentiation in ancient marine lineages of *Exiguobacterium* endemic to the Cuatro Ciénegas Basin. *Environ. Microbiol.* 14, 2323–33.
- Reche, I.E., Pulido-Vilena, R., Morales-Baquero, Casamayor, E.O., 2005. Does ecosystem size determine aquatic bacterial richness? *Comment. Ecology* 86, 1715–1722.
- Reed, H.E., Martiny, J.B.H., 2013. Microbial composition affects the functioning of estuarine sediments. *ISME J.* 7, 868–79.
- Rodríguez-Verdugo, A., Souza, V., Eguiarte, L.E., Escalante, A.E., 2012. Diversity across Seasons of Culturable *Pseudomonas* from a Desiccation

Lagoon in Cuatro Ciénegas, Mexico. *Int. J. Microbiol.* 2012, Article ID 201389, 10 pages.

Souza, V., Espinosa-Asuar, L., Escalante, A.E., Eguiarte, L.E., Farmer, J., Lloret, L., Rodríguez-Martínez, J.M., Soberón, X., Dirzo, R., Elser, J., Forney, L., Sobero, X., Elser, J.J., 2006. An endangered oasis of aquatic microbial biodiversity in the Chihuahuan desert. *Proc. Natl. Acad. Sci. U. S. A.* 103, 6565–70.

Sul, W.J., Oliver, T.A., Ducklow, H.W., Amaral-Zettler, L.A., Sogin, M.L., 2013. Marine bacteria exhibit a bipolar distribution. *Proc. Natl. Acad. Sci. U. S. A.* 110, 2342–7.

Van der Gucht, K., Vandekerckhove, T., Vloemans, N., Cousin, S., Muylaert, K., Sabbe, K., Gillis, M., Declerk, S., De Meester, L., Vyverman, W., 2005. Characterization of bacterial communities in four freshwater lakes differing in nutrient load and food web structure. *FEMS Microbiol. Ecol.* 53, 205–20.

Zhao, D., Huang, R., Zeng, J., Yan, W., Wang, J., Ma, T., Wang, M., Wu, Q.L., 2012. Diversity analysis of bacterial community compositions in sediments of urban lakes by terminal restriction fragment length polymorphism (T-RFLP). *World J. Microbiol. Biotechnol.* 28, 3159–70.

APÉNDICE 1

Protocolo de extracción de ADN con “*Ultraclean water DNA Kit*” Marca MoBio. Cat. No. 148 00-25

1. Colocar 1 ó 2 filtros en el “*bead tube*” y agregar 4 ml de “*bead solution*”, vortex 1 mn. Es posible colocar más de 2 filtros en un mismo tubo, en particular para muestras de agua que utilizaron muchos filtros. Si es necesario, poner 0.5ml más de solución.
2. Agregar 0.5ml de solución WD1 y 5 µl de RNAsa, vortex 30 secs. Colocar los tubos horizontalmente en el agitador y mezclar a toda velocidad por 15 mn.
3. Centrifugar los tubos a 5000 rpm por 3 mn. Eliminar los filtros con cuidado antes de centrifugar.
4. Transferir el sobrenadante a un tubo limpio de 15 ml y agregar 0.6 ml de solución WD2. Vortex 5 secs, incubar a 4° por 30 mn.
5. Centrifugar a 5 000 rpm por 6 mn. Transferir el sobrenadante con la pipeta a un tubo limpio y agregarle 8 ml de solución WD3. Vortex 5 sec.
6. Poner la mezcla en un “tubo-filtro” y centrifugar a 4000 rpm por 3 mn. Eliminar el filtrado.
7. Agregar 3 ml de solución WD4 y centrifugar a 4000rpm por 3 mn. Descartar el filtrado.
8. Repetir paso 7 con WD4 para asegurar que el ADN quede muy limpio.
9. Centrifugar nuevamente a 5000rpm por 5 mn.
10. Secar el filtro durante 20 mn a temperatura ambiente.
11. Colocar el filtro en un nuevo tubo y agregar 3 ml de solución WD5.
12. Dejar en reposo por 12 hrs. a 4°C
13. Centrifugar a 5000rpm por 3 mn. Ahora el ADN está eluído en el filtrado.

A los 3 ml obtenidos:

14. Transferir el ADN a un falcon de 15 ml.
15. Añadir 2.5 ml de fenol. Mezclar por inversión.
16. Centrifugar a 3,700 rpm por 15 mn en canastilla de columpio
17. Recuperar sobrenadante en un falcon limpio.
18. Añadir fenol:cloroformo: isoamílico 25:24:1

19. Centrifugar a 3,700 rpm por 15 mn en canastilla de columpio
20. Recuperar sobrenadante en un falcon limpio.
21. Añadir acetato de sodio a concentración final de 0.3M
22. Añadir 4 ml de glicoblue y 2 vols de etanol
23. Mezclar por inversión.
24. Dejar precipitar 15 mn a -20°
25. Centrifugar en ángulo fijo a 7,500 rpm por 30 mn a 4°
26. Quitar el alcohol y limpiar el precipitado con 2 ml de EtOH al 70%
27. Dejar secar el precipitado
28. Resuspender en 30 μ l de agua ultra pura
29. Dividir el ADN en dos alícuotas de de 15 μ l cada una. Si el ADN está en uso guardar una de ellas a 4° y la otra a -20° .

APÉNDICE 2

PCR del gen 16S ADN ribosomal, amplificado con fluorescencia para TRFLPs

Oligonucleótidos

27F AGA GTT TGA TCM TGG CTC AG marcado con VIC.

1492R TAC GGY TAC CTT GTT ACG ACT T marcado con FAM.

Concentración inicial de cada reactivo	Concentración final	Cantidad para 100µl (en µl)
Buffer 10X	1X	10
MgCl 30mM	2mM	6.6
dNTPs 10mM	0.8mM	8
Oligonucleótidos 10µM	0.3µM	3
Agua	---	68.2
ADN 10 a 100ng	0.2 a 2 ng/µl	1
Taq 5U/µl	1U en 50 µl	0.2

Ciclos de reacción

94° por 3 mn. 20 ciclos de 94° por 1 mn 52° por 1 mn y 72° por 1mn 20 seg. Extensión final por 3 mn.

APÉNDICE 3

Protocolo TRFLPs

1. PCR con oligos marcados 100ul c/u por triplicado, 20 ciclos máximo (sugerido por Acinas *et al.*, 2005)
2. Extraer banda en gel con el kit *Qiaquick gel extraction kit*, marca Quiagen. Cat. No. 28706)
3. Resuspender en 30 µl
4. Medir la concentración de producto en biofotómetro (Eppendorf ©) Utilizar 40 ng para cada digestión. En el secuenciador cargar 20 ng; y usar los otros 20 ng como "backup", guardados secos y congelados para que no se echen a perder).
5. Digestión con enzimas de restricción:

Digestión en 20 µl

40 ng del producto, en 18ul de agua.
2µl de Buffer 10X correspondiente
5 U de enzima (HaeIII en este caso)

Dejar a 37° por 12 hrs., y un ciclo de 65° final para desactivar la reacción, utilizar un termociclador para mayor exactitud en la temperatura.

6. Desalar restricción (dividir la rxn en 2 partes, y llevar cada una a 50ul)
 - a) 1/10V 125 mM EDTA a cada muestra (5ul para 50 µl)
 - b) Añadir 1/10 3MNaOAc (PH5.2) a cada uno. Mezclar con pipeta (5µl para 50 µl)
 - c) Añadir 2.5V 100%EtOH (125ul para 50 µl)
 - d) Vortex, incubar TA 15 mn
 - e) Centrifugar 2500g 30mn (placa) o 14 000 g 25 mn (tubo)
 - f) Inmediatamente voltear placa o quitar el etanol si es tubo
 - g) Centrifugar 185g 1mn para secar (placa)
 - h) Limpiar 35ul EtOH 70% a c/pozo (placa) o 100ul si es tubo. Centrifugar igual (e, f y/o g).
 - i) Centrifugar placa invertida 185g x 1mn (placa)
 - j) Resuspender en 9.5 ul formamida y 0.5ul ROX. O guardar botón congelado a -20oC.

7. Análisis de datos:

El producto se analizó en un secuenciador Abi Prism ® 3100-Avant (Applied Biosystems), en las siguientes condiciones:

Condiciones de corrida para el capilar de 50 cm: 12 V, 14 000 para el tiempo de corrida y 15 sec. de tiempo de inyección (puede variarse a 30).

Se realizaron corridas por triplicado para cada muestra.

Utilizando el programa GenScanner se analizaron los picos mayores a 20 Unidades de fluorescencia, sólo si estaban presentes en las tres réplicas.

DISCUSIÓN GENERAL

1 Sobre patrones y métodos moleculares

Los resultados que se describen en el Capítulo I de esta tesis, con el análisis de una muestra de mar de Yucatán, permiten reflexionar sobre los alcances de las distintas herramientas moleculares utilizadas. En este contexto, en cuanto a patrones de diversidad y distribución de comunidades bacterianas se refiere, es importante resaltar el sesgo de no detección de los *taxa* raros.

Para comprender la influencia que este sesgo puede tener en los patrones de comunidades analizadas en esta tesis, es necesario mencionar que Zinger y colaboradores desde 2012 desarrollaron la idea de que los patrones de abundancia y diversidad de comunidades bacterianas no resultan sustancialmente distintos si se incluye o no a *taxa* raros. Esta hipótesis ya había sido propuesta previamente (Agogue *et al.*, 2011; Gobet *et al.*, 2010; Pommier *et al.*, 2010). En un artículo reciente Zinger *et al.*, (2014) esta propuesta se confirma, utilizando bases de datos muy amplias con 438 muestras, en un detallado análisis de la diversidad *beta* en comunidades microbianas marinas. Tomando en cuenta estos antecedentes, es posible afirmar que por un lado los patrones que se obtienen con el uso de tecnologías sencillas, como las que se usaron en esta tesis (clonación y TRFLPs) están asociados a *taxa* abundantes en las comunidades estudiadas, y esta abundancia es lo que permite su detección con estas técnicas. Por otro lado, y como idea muy importante, este tipo de patrones son válidos aún en estudios con muestreos relativamente limitados, y por lo tanto su discusión dentro de la ecología microbiana tiene validez, al igual que aquéllos estudios que han utilizado nuevas tecnologías, más costosas y poco accesibles para laboratorios de bajo presupuesto. Esto último además implica un problema informático que no es trivial, en lo que se refiere al manejo e interpretación correcta de toda la información que se genera.

Así, el uso de estas tecnologías bajo costo, e incluso más sencillas experimentalmente (como lo es el caso de los TRFLPs), vuelve posible que laboratorios con poco presupuesto, puedan caracterizar robusta y válidamente su propia (y poco explorada hasta ahora, como es el caso de países

latinoamericanos) diversidad microbiana.

2 De patrones y factores

Existe ya una gran cantidad de información publicada que permite afirmar que es posible describir a los patrones de distribución de microorganismos con los mismos principios ecológicos y estadísticos que se utilizan para describir a los de macroorganismos (Martiny *et al.*, 2006; Hanson *et al.*, 2012; Nemergut *et al.*, 2013; Zinger *et al.*, 2014).

Entre los ejemplos de patrones para microorganismos reconocidos en la literatura (para una buena revisión ver Hanson *et al.*, 2012 y Nemergut *et al.*, 2013), se encuentra el decaimiento por distancia (Astorga *et al.*, 2012), otro es la estructura de las comunidades, asociadas en su mayoría a condiciones ambientales (92% de las publicaciones encuentran asociación ambiental, de acuerdo a Hanson *et al.* (2012), pero también a procesos neutros (cerca del 69%). A nivel filogenético, un patrón importante que se ha reportado es que los atributos simples (por ejemplo el uso de sustratos de carbono) están dispersos en la filogenia y los conservados y más complejos (fotosíntesis o metanogénesis) se encuentran en *clusters* que provienen de ramas muy largas en la filogenia (Martiny *et al.*, 2013). Respecto a los patrones reportados para comunidades acuáticas en particular, la mayoría de los reportes también asocian la estructura de las comunidades a factores ambientales (96% vs. \approx 50% a efecto de distancia) (Hanson *et al.*, 2012). En coincidencia con esta idea, se ha reportado que la adaptación local es un factor predominante en este hábitat, ya que en los metagenomas de comunidades de lagos predominan genes adaptados a cada ambiente particular (Eiler *et al.*, 2014). Es interesante señalar que el tamaño efectivo del genoma ha resultado menor en mar y mayor en lagos (Eiler *et al.*, 2014), y se han publicado evidencias de cuellos de botella y clonalidad en estos hábitats (Zaremba-Niedzwiedzka *et al.*, 2013).

Así como existen patrones claros y reconocidos en la literatura, también es importante mencionar que los mecanismos asociados a la distribución de los organismos en comunidades ecológicas en general son muchos y muy

complejos (Vellend, 2010; Nemergut *et al.*, 2013). Desde hace más de un siglo los ecólogos se debaten entre cuáles son las formas posibles en las que las comunidades pueden estructurarse. El debate clásico (Clements vs. Gleason ca. 1926) implicaba reconocer o negar la existencia de conceptos importantes, como son las interacciones ecológicas y las comunidades (Clements), o el papel del azar y de las condiciones ambientales (Gleason). Actualmente, más que negar la existencia de una posibilidad u otra, se reconoce la necesidad de incorporar a todas las opciones conceptuales (Levins & Lewontin 1985). En este mismo sentido, la cuestión que se ha vuelto popular en los últimos 20 años para el caso de comunidades de microorganismos, ha sido debatir si están estructuradas por factores abióticos o factores neutros, y el fondo de esta cuestión radica en el reconocimiento o negación de la existencia de límites de dispersión en bacterias. Si se retoma el planteamiento de Levins y Lewontin (1985), reconociendo ambas posibilidades sin necesidad de negar una u otra, esta cuestión quedaría zanjada; lo cual es confirmado por las evidencias que apuntan hacia una síntesis de ambas perspectivas (Leibold & McPeck, 2006). Se ha reportado por ejemplo, para comunidades bacterianas de lagos suecos, que mientras más contrastantes sean las características ambientales, éstas serán las que determinen la composición de la comunidad (Östman *et al.*, 2010, 2012), y por el contrario, en ambientes similares, el efecto neutro podría ser el determinante (Bissett *et al.*, 2010; Logares *et al.*, 2013).

Además del análisis de la estructuración de las comunidades bacterianas por factores neutros o factores abióticos, como lo hace este trabajo y muchos de los artículos que actualmente se publican, las interacciones bióticas son un factor poco contemplado en la literatura, sobre todo en aquéllos reportes que utilizan técnicas para analizar el ADN genómico microbiano. Existe una limitante metodológica importante, relacionada con la imposibilidad de determinar a qué individuo de la comunidad total le pertenece cada una de las muchas secuencias que se generan en este tipo de análisis. Por lo tanto, la información obtenida sobre los fenotipos presentes en la comunidad asociados a algún taxón en particular no resulta muy precisa. Una solución sería obtener genomas

completos para cada taxón presente en una comunidad, pero el tamaño de los fragmentos obtenidos a partir de ADN ambiental aún no permite el armado de cada genoma, con lo cual sería posible la caracterización exacta para cada taxón. Algunos artículos han inferido el fenotipo de manera indirecta, mediante la comparación de los fragmentos del gen 16S ADNr (lo cual permite una caracterización taxonómica) con genomas completos ya publicados (e.g. Goberna *et al.*, 2014, Barberán *et al.*, 2014), lo que resulta en inferencias aproximadas sobre el fenotipo de cada individuo; con estos datos han generado resultados interesantes sobre la estructura de las comunidades bacterianas. Otra manera que resulta más precisa para el estudio de las interacciones bacterianas ha sido a partir de bacterias en cultivo, con las restricciones que implica una sub-representación de la comunidad, ya que sólo el 1% de las bacterias pueden ser cultivadas.

2.1 Sobre patrones triangulares

En cuanto a los patrones en forma de triángulo que se observan en las pruebas de Mantel reportadas en la literatura (e.g. Bissett *et al.*, 2010; Lear *et al.*, 2014; Logares *et al.*, 2013) y que también se obtuvieron en este trabajo (Capítulo III), generados al asociar la diversidad *beta* con las distancia ambiental o geográfica de las comunidades, éstos podrían ser un reflejo de cómo ambos factores –tanto neutros como abióticos– pueden estar influenciando a la vez la estructura de las comunidades, en coincidencia con la propuesta de Leibold & McPeck, (2006) hacia una síntesis de ambas perspectivas.

Para un mayor entendimiento de este patrón triangular, se plantea un muestreo hipotético en tres montañas (Figura 1, arriba), asumiendo que sus comunidades están estructuradas tanto por factores ambientales (abióticos en este caso), como por factores geográficos, tal como lo proponen Leibold & McPeck (2006). En la figura se comparan distintas combinaciones de comunidades (lejanas o cercanas, similares o diferentes ambientalmente), y se muestra la posición que tendrían los puntos correspondientes a esas comparaciones (ver Figura 1, abajo). Hay puntos que van formando tendencias

claras, por ejemplo los de las gráficas c y f. En el primer caso (c) las comunidades son ambientalmente equivalentes y muestran un patrón asociado a distancia geográfica, En el segundo caso (f) son comunidades que provienen de ambientes diferentes, y la asociación es ambiental. En otros ejemplos el patrón se pierde o puede no ser tan evidente, debido a la interrelación tanto de los factores ambientales como los de distancia (e.g. puntos de las gráficas d y e), por lo cual caen dentro del triángulo, llenándolo sin orden aparente. Si continuáramos modelando comparaciones como las que se marcan en la parte inferior de la figura, para cada una de las combinaciones pareadas de las comunidades hipotéticas, el patrón sería análogo al que se generó en los ejemplos que se proporcionan. Por otra parte, la zona que no se ocupa (el triángulo invertido, ver gráficas g y h), representa a comunidades similares (zona de arriba) con diferencias ambientales o geográficas (zona hacia la derecha), una combinación que es muy poco probable que ocurra. De esta forma, la existencia de estos patrones triangulares en comunidades naturales podría ser una evidencia de que la combinación de ambos factores (ambientales y distancia geográfica) produce una distribución de puntos restringidos al área triangular mencionada.

Si se estudian únicamente comunidades ambientalmente equivalentes (e.g. puntos en la gráfica c), los patrones de distancia geográfica podrían ser más claros (eliminando a las comunidades “ruido” como las de la gráfica e), en coincidencia con Bissett *et al.*, (2010). Por último, es importante mencionar que el ambiente contrastante es una característica muy frecuente en la naturaleza. Los muestreos en esa condición resultarán, de acuerdo a Östman *et al.* (2010, 2012), en la detección –frecuente, tal como se ha reportado por Hanson *et al.*, (2012)– de patrones ambientales.

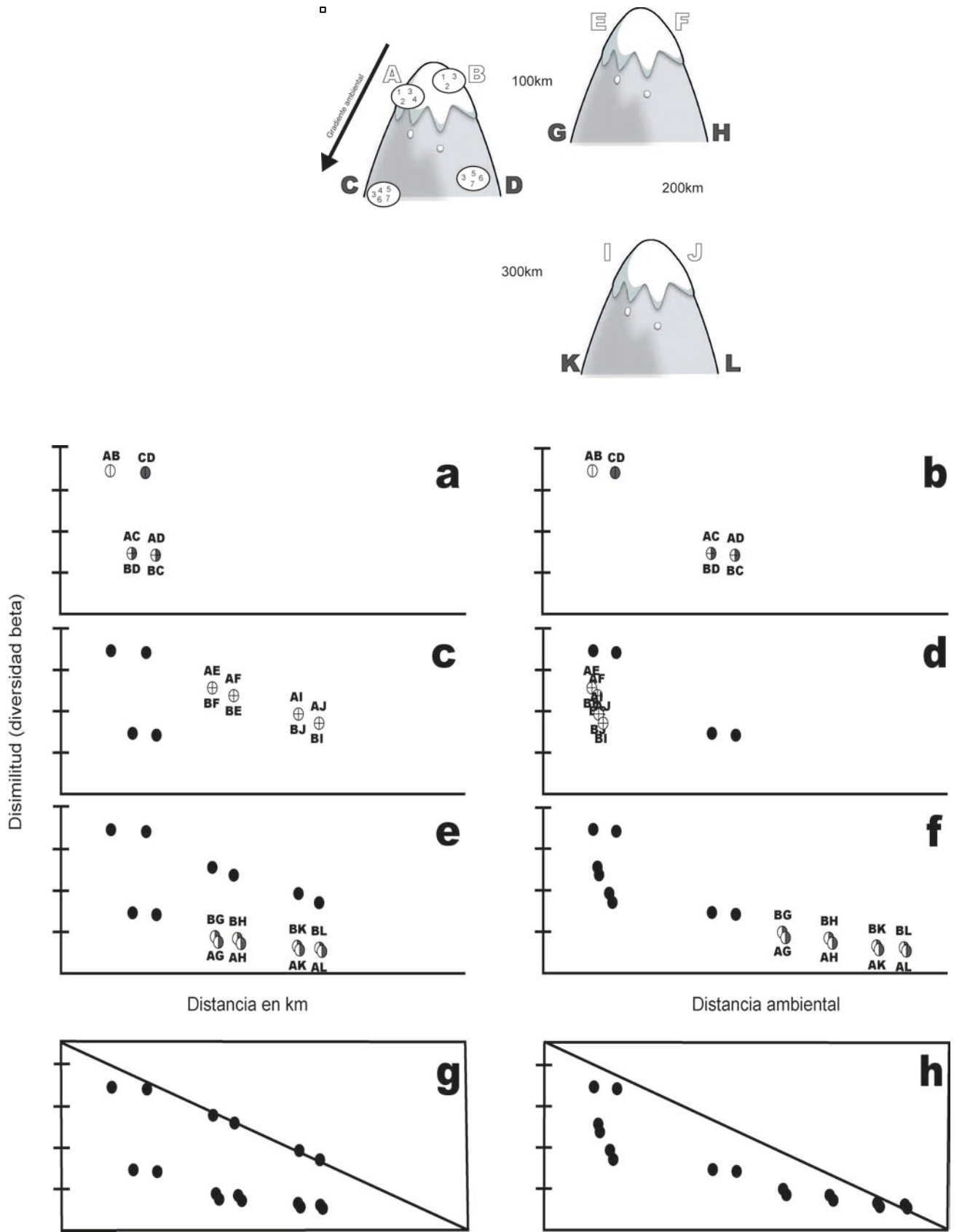


Figura 1

Fig. 1. Muestreo hipotético en tres montañas (arriba), en cada una se examinan dos ambientes distintos (en blanco o en gris), y la distancia entre las montañas es diferente. Se comparan distintos tipos de comunidades, mostrando la posición que tendrían los puntos correspondientes a esas comparaciones (abajo). En esta parte inferior, todas las gráficas del lado izquierdo hacen referencia a distancia geográfica, y las que se encuentran del lado derecho, a distancia ambiental (abiótica).

(a) y (b) representan una simulación para las comunidades de una misma montaña. Los puntos marcados como AB y CD están posicionados de acuerdo a la comparación de las comunidades A vs. B (equivalentes ambientalmente y representadas con un punto blanco dividido a la mitad) y C vs. D (también ambientalmente equivalentes y representadas en gris), de acuerdo a la distancia geográfica (a) o distancia abiótica (b). Estos puntos están situados arriba de la gráfica ya que las comunidades son similares. Los puntos marcados como AC/BD y BD/BC corresponden a la comparación de comunidades ambientalmente distintas (blanco vs. gris), por lo cual tienen menor valor de similitud y se encuentran más abajo. La posición de los puntos correspondientes a las comparaciones AC y BD sería equivalente, por eso se encuentran en un mismo sitio, y para estos casos un mismo punto se divide en cuatro, señalizando que son dos comparaciones que comparten la misma posición en la gráfica.

(c) y (d) corresponden a una simulación para comunidades de dos montañas diferentes, pero ambientalmente equivalentes. Los puntos en negro son los de la simulación previa (en este caso (a) y (b)), y se agregan nuevas comparaciones de comunidades similares (blancas). Los puntos están divididos en cuatro porque son dos comparaciones distintas, que ocupan la misma posición.

(e) y (f) simulan comunidades de dos montañas distintas, y con diferentes condiciones ambientales. La explicación es parecida a la anterior, pero para comunidades distintas (blancas vs. grises).

(g) y (h) marcan el área triangular que se va ocupando en cada simulación (ambiental o geográfica).

Consultar el texto para la discusión del patrón.

3 Patrones y escalas

La importancia relativa de los distintos procesos es diferente de acuerdo la escala de estudio (Vellend, 2010), y en el caso de los microorganismos, aún no se comprende claramente el papel que tiene esta última, ya sea espacial o temporal, en el ensamble de sus comunidades (Nemergut *et al.*, 2013). Una propuesta ha sido que a escala local la dispersión tiene gran importancia, por lo cual la dispersión limitada de las bacterias determina la estructura de sus comunidades, y el efecto geográfico se hace evidente, a diferencia de escalas mayores en las que la selección (factores ambientales) es el factor determinante (Martiny *et al.*, 2011).

En este contexto se diseñaron los dos muestreos que se describen en los capítulos II y III, uno de ellos a escala local (menos de 2 km) y otro a escala regional (distancia mayor de 180 km entre dos sitios) y las características ambientales para cada diseño fueron diferentes. En el primero, los sitios de muestreo eran equivalentes entre ellos, mientras que en el segundo se eligieron sitios contrastantes, y los patrones que se obtuvieron en estos dos muestreos resultaron diferentes. La estructura de las comunidades en el muestreo con ambientes contrastantes estuvo asociada a factores ambientales, y la distancia geográfica tuvo correlación significativa con la distribución *beta* de las comunidades muestreadas en hábitats equivalentes.

Estos dos patrones se ajustan con algunas predicciones ya reportadas: características ambientales contrastantes en las comunidades bacterianas tendrán un fuerte efecto en la composición de las comunidades (Östman *et al.*, 2010, 2012), y en ambientes similares, el efecto neutro podría ser el determinante (Bissett *et al.*, 2010; Logares *et al.*, 2013). Sobre la escala a la que se realizaron ambos muestreos no es posible dar conclusiones claras. En coincidencia con Martiny *et al.*, (2011), el efecto geográfico fue detectado a escala local, y a una escala mayor se detectó un efecto ambiental. Pero es importante mencionar que otras revisiones sobre reportes publicados para comunidades de bacterias no muestran un patrón asociado a una escala en particular (e.g. Hanson *et al.*, 2012). Para poder tener conclusiones más claras,

sería necesario un muestreo a nivel regional en condiciones ambientales equivalentes, y viceversa, un muestreo a nivel local en condiciones ambientales contrastantes.

4 De los cuatro procesos

La propuesta de Vellend (2010), retomada por Nemergut *et al.* (2013) sobre una síntesis conceptual que enfatice los cuatro procesos (dispersión, deriva, diversificación y selección) para explicar los patrones observados en las comunidades, resulta interesante como una aproximación más sencilla para el entendimiento de dichos patrones.

Por ejemplo, la dispersión es un proceso que cobra importancia a escalas pequeñas, tal como lo reportan Martiny *et al.* (2011), ya que una baja dispersión puede generar comunidades estructuradas con procesos neutros. Adicionalmente, una dispersión limitada también puede verse como un proceso característico de un taxón específico (independientemente de la escala), y en coincidencia con esta idea, existen reportes que encuentran asociación a la distancia geográfica en *taxa* con baja dispersión (Bissett *et al.*, 2010; Astorga *et al.*, 2012). Para este caso en particular, el patrón estará estructurado por una baja dispersión, siendo este el proceso que genera la estructura neutra en las comunidades, independientemente de si está relacionado con alguna escala.

De esta forma, poniendo en énfasis en el proceso (en este caso una baja dispersión asociado a una estructura neutra) se vuelve posible generalizar un patrón y dar una explicación sobre el proceso que lo estructura . En este mismo sentido, también podría proponerse que el proceso de selección, acoplado a condiciones ambientales contrastantes (independientemente de la escala), generará una estructura determinística en las comunidades, en coincidencia con lo planteado por Östman *et al.* (2010, 2012). Por último, y en concordancia con esta propuesta de describir patrones de acuerdo a los procesos asociados, aquellas comunidades que tengan ambientes no contrastantes, estarán estructuradas de forma neutra siempre y cuando el proceso de dispersión limitada esté operando en ellas.

Con estas ideas en mente, es posible proponer que los dos patrones observados en esta tesis, independientemente de la escala de estudio, se encuentran asociados a dos procesos: selección, para el muestreo de ambientes contrastantes, y dispersión limitada en las comunidades con hábitats semejantes.

5 Sobre los *taxa*

Es necesario enfatizar que los patrones obtenidos en este trabajo representan a los *taxa* más abundantes de cada comunidad, ya que ninguno de los métodos utilizados detectan *taxa* poco abundantes (Curtis & Sloan, 2005; Sogin *et al.*, 2006).

Otra idea importante es que no necesariamente los *taxa* más abundantes en estos muestreos son *taxa* generalistas, aunque algunos autores apoyan lo anterior (e.g. Székely & Langenheder, 2014). En el caso de las comunidades de Pozas Azules, los OTUs que sí fue posible clasificar son similares a *taxa* que tienen una amplia distribución, y que se adaptan a diferentes condiciones ambientales, por lo que probablemente, este tipo de OTUs, pertenecen a *taxa* generalistas. Pero otros OTUs abundantes (los cuales no fueron clasificados por ser secuencias sin parecido a ninguna otra en las bases de datos mundiales) podrían pertenecer a algún tipo de microorganismos que estén fuertemente adaptados a alguna de las características de las pozas (especialistas), siendo que estas últimas son ambientes equivalentes. Esta es sólo una propuesta, que podrá ser corroborada cuando existan publicaciones sobre secuencias de bacterias similares, lo que permitirá dilucidar la historia de vida con mayor precisión. Mientras tanto, con esta posibilidad en mente, no hay que descartar la idea de que incluso los *taxa* especialistas pueden estar sujetos también a procesos neutros además de los procesos de adaptación que los caracterizan. Y de acuerdo a lo planteado en el apartado anterior (sobre enfatizar los procesos detrás de los patrones), los *taxa* especialistas que estarían sujetos a procesos neutros, serán aquéllos que tengan como característica una baja dispersión.

Por lo tanto, se vuelve necesario evitar generalizaciones: desde asegurar

por ejemplo que los *taxa* abundantes son siempre generalistas, hasta asegurar que cierto tipo de *taxa* (generalista o especialista, abundante o raro) está asociado con factores neutros o factores ambientales, porque puede llevar a contradicciones como las que ya existen en la literatura. Por ejemplo la mencionada en la Introducción sobre los *taxa* raros asociados factores neutros por tener mala dispersión (Lepère *et al.*, 2013) o no asociados a factores neutros, debido a su lento crecimiento (Logares *et al.*, 2013).

Por lo tanto, la descripción precisa de características particulares es importante para esclarecer los patrones. Por ejemplo, es más correcto afirmar que los *taxa* de crecimiento rápido (que pueden ser tanto generalistas como especialistas) tienden a ensamblarse en forma neutra (Langenheder & Székely, 2011) y también lo hacen aquéllos *taxa* con poca capacidad de dispersión (Bissett *et al.*, 2010). De esta forma, mientras más meticulosa y clara sea la información obtenida, los patrones podrán asociarse con mayor certeza a los procesos correspondientes.

6 Sobre patrones en Cuatro Ciénegas

Se vuelve necesario preguntarse el porqué de la asociación con la distancia geográfica en Pozas Azules (Capítulo II), en un sitio con grandes adaptaciones tal como se ha descrito para ambientes lacustres, caracterizado por poblaciones clonales y cuellos de botella (Avitia *et al.*, 2014) y con tantas evidencias de procesos de selección (Cerritos *et al.*, 2011; Rebollar *et al.*, 2012; López-Lozano *et al.*, 2012).

El efecto de distancia es detectado con mayor frecuencia en comunidades de bacterias que colonizan hábitats más amplios, como sucede por ejemplo en el mar, con casi un 80% de los reportes, comparados con cerca del 50% de los artículos dedicados a estudiar comunidades de lagos (Hanson *et al.*, 2012). Las adaptaciones locales que caracterizan a los sistemas acuáticos continentales (más del 90% de los reportes, de acuerdo a Hanson *et al.* (2012), reflejan la enorme variación ambiental que existe entre estos sistemas, lo cual tiene como consecuencia que la asociación a factores neutros sea difícil de detectar

(Ostman *et al.*, 2010), y que predominen factores locales en la caracterización de sus comunidades bacterianas. Este es el caso del estudio a gran escala que se hizo en este trabajo, y también de otros estudios microbianos en Cuatro Ciénegas con características ambientales contrastantes (Cerritos *et al.*, 2011; Rebollar *et al.*, 2012; López-Lozano *et al.*, 2012).

Dentro de esta reflexión sobre el efecto de distancia encontrado en Pozas Azules, es necesario considerar el contexto de la integración de ambos factores (neutros y ambientales) en la estructura de las comunidades bacterianas, propuesto por Leibold & McPeck (2006). Así, además del proceso tan evidente de selección que hay en las comunidades de Cuatro Ciénegas, también la distancia (y por lo tanto la dispersión) está operando en la estructura de los ensamblajes, tal como se evidencia en las comunidades de Pozas Azules. Por lo tanto, y dado que incluso los organismos muy especializados en un hábitat o ambiente particular también poseen cierta capacidad de dispersión, es posible afirmar que el efecto de la distancia también influirá en organismos que están bajo procesos de selección, lo cual se vuelve evidente al utilizar ambientes equivalentes. Sería interesante explorar a estas comunidades de bacterias “cieneguenses”, utilizando diseños experimentales enfocados en los dos tipos de ambientes –tanto equivalentes como contrastantes— con lo cual será posible evidenciar si ambos factores (neutros y locales) están operando en las comunidades, de acuerdo a Leibold & McPeck (2006).

Los datos de las comunidades de Pozas Azules, en comparación con otras comunidades bacterianas en el valle, concuerdan en características de la diversidad y distribución de las comunidades previamente reportadas. La diversidad *alfa* (índice de Shannon desde 1.579 hasta 2.737) se encuentra dentro de los límites publicados para comunidades acuáticas (rangos de 1.3 (Escalante *et al.*, 2008) y 1.5 (Pajares *et al.*, 2012, 2013) hasta 3.13-3.8 respectivamente). Otra coincidencia ha sido la característica de tener un taxón abundante (particular para cada sitio), y muchos *taxa* poco abundantes (Escalante *et al.*, 2008; Pajares *et al.*, 2012), tal como se reporta para distintas comunidades bacterianas (Sogin *et al.*, 2006). Se ha afirmado en otros estudios,

que este tipo de estructura puede presentarse en forma cíclica en los ensamblajes bacterianos (ver Fuhrman *et al.*, 2006).

Es pertinente preguntarse el porqué de un taxón abundante y característico en comunidades como las de Pozas Azules, ambientalmente similares y en las que se ha detectado asociación a la distancia, ya que podría esperarse una estructura determinada por el fenómeno de invarianza regional, con una abundancia de *taxa* similar para todos los sistemas, dado que son sitios ambientalmente equivalentes. Es posible explicar esta predominancia particular, ya que el orden de colonización en las comunidades puede resultar en comunidades divergentes, incluso si las condiciones ambientales son iguales y el pool de especies es el mismo (Gleason, 1927), apoyado por evidencias en la literatura que reportan ambientes similares con grandes diferencias de composición (Woodcock *et al.*, 2007).

Por último, el complejo sistema hidrogeológico del valle de Cuatro Ciénegas invita a pensar en la conexión subterránea que podría existir entre los cuerpos de agua estudiados. Pruebas preliminares utilizando la técnica de T-RFLPs en el sistema de Pozas Azules, mostraron que las pozas conectadas por canales superficiales visibles tenían perfiles de comunidades casi idénticos tanto en presencia de *taxa* como en abundancia. Esto no sucedió para las cuatro pozas muestreadas, lo cual nos permitió suponer que no existe conexión subterránea entre ellas, pero es una posibilidad que no puede descartarse con certeza.

CONCLUSIONES

1. Actualmente es posible estudiar comunidades microbianas con tecnologías sencillas y económicas, las cuales proporcionan información importante así como patrones ecológica y estadísticamente válidos.
2. El efecto de distancia geográfica (neutro) fue detectado a escala local. Este es un patrón que no había sido reportado previamente en otros trabajos sobre bacterias en Cuatro Ciénegas. A una escala mayor se detectó un efecto ambiental (nicho).
3. Los patrones triangulares en las pruebas de Mantel podían estar representando la síntesis de factores ambientales y de distancia geográfica (Leibold y McPeck 2006) en una comunidad.
4. El estudio de comunidades ambientalmente equivalentes permite obtener patrones de distancia geográfica más claros. Por otra parte, un muestreo que contemple condiciones ambientales contrastantes reflejará predominantemente patrones ambientales (Bissett *et al.*, 2010; Logares *et al.*, 2013; Östman *et al.*, 2010, 2012).
5. Los dos patrones observados en esta tesis se encuentran asociados a dos procesos: selección, para el muestreo de ambientes contrastantes, y dispersión limitada en las comunidades con hábitats semejantes.
6. La descripción precisa de características particulares de los *taxa* estudiados es importante para esclarecer los patrones (Nemergut *et al.*, 2013).
7. Los datos de las comunidades de Pozas Azules, en comparación con otros reportes similares en el valle de Cuatro Ciénegas, concuerdan en diversidad y distribución de otras comunidades previamente reportadas.
8. Las comunidades microbianas estudiadas en este trabajo presentaron patrones de distribución similares a las comunidades de macroorganismos: efecto de distancia y distribución restringida de OTUs.

PERSPECTIVAS

1. Es necesario explorar el efecto de la escala con distintas combinaciones de muestreo: tanto a gran escala como a una escala local, utilizando sitios contrastantes y sitios equivalentes ambientalmente.
2. Podría realizarse una buena caracterización previa de los sitios, que comprenda desde características físicas hasta detallados análisis químicos con el mayor número posible de variables, incluyendo diferentes opciones para nutrientes y para tipos de sales en el agua, entre otros.
3. Determinar la historia de vida de algunos *taxa* en particular. Por ejemplo, los que estén presentes en la mayoría de las muestras, o algunos OTUs asociados con determinada variable ambiental. Algunas posibilidades para hacerlo son:
 - Definir grupos funcionales a partir de una comparación de las secuencias que codifican para el gen 16S ADNr con los genomas completos que están disponibles en bases de datos públicas.
 - Caracterizar abundancia, así como el área de distribución de cada taxón (dispersión), con técnicas de *next-generation* (NGS).
 - Secuenciación exhaustiva para algunas de las muestras, lo que permitirá oligonucleótidos específicos dirigidos (incluso hacia *taxa* poco abundantes).
 - Filochip diseñado a partir de los datos obtenidos de NGS. De esta forma se generarán datos más precisos, de manera rápida y confiable, sobre la distribución y abundancia de los *taxa* analizados.
 - Sería importante avanzar en la genómica de célula única (*single cell genomics*, preseleccionando las células con fluorescencia), para explorar los genomas de algunos *taxa*. Con este tipo de análisis se obtendrán datos metabólicos y se podrán hacer inferencias sobre los procesos implicados en la estructura de las comunidades, incluso a nivel poblacional. Alternativamente, se pueden hacer metagenomas de las

comunidades completas, con los problemas informáticos que representa toda la información que se genera.

4. Llevar a cabo modelos estadísticos formales, usando datos de otros trabajos publicados, para determinar si el patrón triangular en la prueba de Mantel representa el efecto de ambos factores (ambientales y de distancia geográfica) en una comunidad. Será necesario considerar las variables y características que se reportan para cada muestreo.

5. Considerar las características e historia geológica de los sitios muestreados y compararlas con los patrones de comunidades obtenidos, con el fin de comprender con mayor claridad los procesos históricos que moldean a las comunidades.

BIBLIOGRAFÍA CORRESPONDIENTE A LA INTRODUCCIÓN Y DISCUSIÓN GENERAL

1. Agogue H, Lamy D, Neal P, Sogin M & Herndl G (2011) Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. *Mol. Ecol.* **20**: 258–274
2. Aguirre-von-Wobeser E, Soberón-Chávez G, Eguiarte LE, Ponce-Soto GY, Vázquez-Rosas-Landa M, Souza V. (2014) Two-role model of an interaction network of free-living γ -proteobacteria from an oligotrophic environment. *Environ. Microbiol.* **16**:1366-1377
3. Alcaraz LD, Moreno-Hagelsieb G, Eguiarte LE, Souza V, Herrera-Estrella L & Olmedo G (2010) Understanding the evolutionary relationships and major traits of *Bacillus* through comparative genomics. *BMC Genomics* **11**: 332
4. Alcaraz LD, Olmedo G, Bonilla G, Cerritos R, Hernández G, Cruz A, Ramírez E, Putonti C, Jiménez B, Martínez E, López V, Arvizu JL, Ayala F, Razo F, Caballero J, Siefert J, Eguiarte L, Vielle J-P, Martínez O, Souza V, et al (2008) The genome of *Bacillus coahuilensis* reveals adaptations essential for survival in the relic of an ancient marine environment. *Proc. Natl. Acad. Sci. U. S. A.* **105**: 5803–8
5. Andersson M, Berga M, Lindström ES & Langenheder S (2014) The spatial structure of bacterial communities is influenced by historical environmental conditions. *Ecology* **95**: 1134–1140
6. Astorga A, Oksanen J, Luoto M, Soininen J, Virtanen R & Muotka T (2012) Distance decay of similarity in freshwater communities: do macro- and microorganisms follow the same rules? *Glob. Ecol. Biogeogr.* **21**: 365–375
7. Avitia M, Escalante AE, Rebollar EA, Moreno-Letelier A, Eguiarte LE & Souza V (2014) Common population expansions in coexisting bacterial lineages revealed by genetic evidence. *PeerJ in press*:
8. Baas Becking LGM (1934) Geobiologie of inleiding tot de milieukunde. *Hague, Netherlands W.P. Van Stock. Zoon (in Dutch)*.
9. Barberán A, Fernández-Guerra A, Bohannan BJM & Casamayor EO (2012) Exploration of community traits as ecological markers in microbial metagenomes. *Mol. Ecol.* **21**: 1909–1917
10. Barberán A, Ramirez KS, Leff JW, Bradford M a, Wall DH & Fierer N (2014) Why are some microbes more ubiquitous than others? Predicting the habitat breadth of soil bacteria. *Ecol. Lett.* **17**: 794–802
11. Begon M, Harper JL & Townsend CR (2006) Ecology: Individuals, Populations and Communities. In *Blackwell science* p 738
12. Bell T, Ager D, Song JI, Newman JA, Thompson IP, Lilley AK & van der Gast CJ. (2005) Larger islands house more bacterial taxa. *Science.* **308**:1884
13. Bissett A, Richardson AE, Baker G, Wakelin S & Thrall PH (2010) Life history determines biogeographical patterns of soil bacterial communities over multiple spatial scales. *Mol. Ecol.* **19**: 4315–4327
14. Bonilla-Rosso G, Peimbert M, Alcaraz LD, Hernández I, Eguiarte LE, Olmedo-Alvarez G & Souza V (2012) Comparative metagenomics of two

- microbial mats at Cuatro Ciénegas Basin II: community structure and composition in oligotrophic environments. *Astrobiology* **12**: 659–73
15. Boto L (2010) Horizontal gene transfer in evolution: facts and challenges. *Proc. Biol. Sci.* **277**:819-27
 16. Breitbart M, Hoare A, Nitti A, Siefert J, Haynes M, Dinsdale E, Edwards R, Souza V, Rohwer F & Hollander D (2009) Metagenomic and stable isotopic analyses of modern freshwater microbialites in Cuatro Ciénegas, Mexico. *Environ. Microbiol.* **11**: 16–34
 17. Brown M V, Ostrowski M, Grzymalski JJ & Lauro FM (2014) A trait based perspective on the biogeography of common and abundant marine bacterioplankton clades. *Mar. Genomics* **15**: 17–28
 18. Burke C, Steinberg P, Rusch D, Kjelleberg S & Thomas T (2011) Bacterial community assembly based on functional genes rather than species. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 14288–93
 19. Calderón-Patrón JM, Moreno CE & Zuria I (2012) La diversidad beta: medio siglo de avances. *Rev. Mex. Biodivers.* **83**: 879–891
 20. Cerritos R, Eguiarte LE, Avitia M, Siefert J, Travisano M, Rodríguez-Verdugo A & Souza V (2011) Diversity of culturable thermo-resistant aquatic bacteria along an environmental gradient in Cuatro Ciénegas, Coahuila, México. *Antonie Van Leeuwenhoek* **99**: 303–18
 21. Clements F (1916) Plant Succession: An Analysis of the Development of Vegetation. *Carnegie Inst. Washington, Washington, DC.* **242**: 1–512
 22. Clements FE (1936) Nature and structure of the climax. *J. Ecol.* **24**: 252–284
 23. Cottenie K (2005) Integrating environmental and spatial processes in ecological community dynamics. *Ecol. Lett.* **8**: 1175–82
 24. Curtis TP & Sloan WT (2005) Microbiology. Exploring microbial diversity-- a vast below. *Science* **309**: 1331–3
 25. Edwards KF, Litchman E & Klausmeier CA (2013) Functional traits explain phytoplankton responses to environmental gradients across lakes of the United States. *Ecology* **94**: 1626–1635
 26. Eguiarte LE, Aguirre-liguori JA, Jardón-barbolla L, Aguirre-planter E & Souza V (2013) Genómica de poblaciones: nada en evolución va a tener sentido si no es a la luz de la genómica, y nada en genómica tendrá sentido si no es a la luz de la evolución. *TIP Rev. Esp. Cienc. Quim. Biol* **16**: 5–19
 27. Eiler A, Zaremba-Niedzwiedzka K, Martínez-García M, McMahon KD, Stepanauskas R, Andersson SGE & Bertilsson S (2014) Productivity and salinity structuring of the microplankton revealed by comparative freshwater metagenomics. *Environ. Microbiol.* **16**: 2682–98
 28. Elser JJ, Schampel JH, Garcia-Pichel F, Wade BD, Souza V, Eguiarte L, Escalante A & Farmer JD (2005) Effects of phosphorus enrichment and grazing snails on modern stromatolitic microbial communities. *Freshw. Biol.* **50**: 1808–1825
 29. Escalante AE (2008) Ecología evolutiva de procariontes en Cuatro Ciénegas Coahuila, México.

30. Escalante AE, Eguiarte LE, Espinosa-Asuar L, Forney LJ, Noguez AM & Souza Saldivar V (2008) Diversity of aquatic prokaryotic communities in the Cuatro Ciénegas basin. *Fems Microbiology Ecol.* **65**: 50–60
31. Evans SB (2005) Using Chemical Data to Define Flow Systems in Cuatro Ciénegas Coahuila, Mexico. Doctoral dissertation, University of Texas, Austin.
32. Ferrenberg S, O'Neill SP, Knelman JE, Todd B, Duggan S, Bradley D, Robinson T, Schmidt SK, Townsend AR, Williams MW, Cleveland CC, Melbourne BA, Jiang L & Nemergut DR (2013) Changes in assembly processes in soil bacterial communities following a wildfire disturbance. *ISME J.* **7**: 1102–1111
33. Finlay BJ (2002) Global dispersal of free-living microbial eukaryote species. *Science* **296**: 1061–3
34. Forti, P, Giulivo I, Mecchia M & Piccini L (2004) Las aguas de Cuatro Ciénegas. 76–87 en *Bajo el desierto. El Misterio de las aguas de Cuatro Ciénegas* Badino G, Bernabei T, De Vivo A, Giulivo I & Savino G (eds) 1era edición. La Venta - Esplorazioni Geografiche, Italia.
35. Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown M V & Naeem S (2006) Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 13104–9
36. Galand PE, Casamayor EO, Kirchman DL & Lovejoy C (2009) Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 22427–22432
37. Ghiglione J-F, Galand PE, Pommier T, Pedrós-Alió C, Maas EW, Bakker K, Bertilson S, Kirchmanj DL, Lovejoy C, Yager PL & Murray AE (2012) Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* **109**: 17633–8
38. Gilbert B & Lechowicz MJ (2004) Neutrality, niches, and dispersal in a temperate forest understory. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 7651–7656
39. Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, Bibbs L, Eads J, Richardson TH, Noordewier M, Rappé MS, Short JM, Carrington JC & Mathur EJ. (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science.* **309**:1242-1245.
40. Gleason HA (1927) Further Views on the Succession-Concept. *Ecology* **8**: 299–326
41. Goberna M, Navarro-Cano JA, Valiente-Banuet A, García C & Verdú M (2014) Abiotic stress tolerance and competition-related traits underlie phylogenetic clustering in soil bacterial communities. *Ecol. Lett.* **17**:1191-1201.
42. Gobet A, Quince C & Ramette A (2010) Multivariate Cutoff Level Analysis (MultiCoLA) of large community data sets. *Nucleic Acids Res.* **38**: e155
43. Gravel D, Canham CD, Beaudet M & Messier C (2006) Reconciling niche and neutrality: The continuum hypothesis. *Ecol. Lett.* **9**: 399–409

44. Gravel D, Poisot T & Desjardins-Proulx P (2014) Using neutral theory to reveal the contribution of meta-community processes to assembly in complex landscapes. *J. Limnol.* **73**: 61–73
45. Green JL, Bohannan BJM & Whitaker RJ (2008) Microbial biogeography: from taxonomy to traits. *Science* **320**: 1039–43
46. Grote J, Thrash JC, Huggett MJ, Landry ZC, Carini P, Giovannoni SJ, & Rappé MS (2012) Streamlining and core genome conservation among highly divergent members of the SAR11 clade. *Mbio* **3**: e00252-12.
47. Hanson CA, Fuhrman JA, Horner-Devine MC & Martiny JBH (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.* **10**: 497–506
48. Harpole WS (2012) Neutral Theory of Species Diversity. *Nat. Educ. Knowl.* **3**: 60
49. Hatosy SM, Martiny JBH, Sachdeva R, Steele J, Fuhrman JA & Martiny AC (2013) Beta diversity of marine bacteria depends on temporal scale. *Ecology* **94**: 1898–1904
50. Horner-Devine MC, Carney KM & Bohannan BJ. 2004 An ecological perspective on bacterial biodiversity. *Proc. Biol. Sci.* **271**:113-122.
51. Hortal J, Nabout JC, Calatayud J, Carneiro FM, Padial A, Santos AMC, Siqueira T, Bokma F, Bini LM & Ventura M (2014) Perspectives on the use of lakes and ponds as model systems for macroecological research. *J. Limnol.* **73**: 46–60
52. Hubbell S (2001) The unified neutral theory of biodiversity and biogeography Princeton, NJ: Princeton University Press
53. Johannesson KH, Cortés A & Kilroy KC (2004). Reconnaissance isotopic and hydrochemical study of Cuatro Ciénegas groundwater, Coahuila, México. *J. South. Am. Earth Sciences* **17**: 171–180
54. Jones SE, Cadkin T a, Newton RJ & McMahon KD (2012) Spatial and temporal scales of aquatic bacterial beta diversity. *Front. Microbiol.* **3**: 318
55. Langenheder S & Székely AJ (2011) Species sorting and neutral processes are both important during the initial assembly of bacterial communities. *ISME J.* **5**: 1086–94
56. Lear G, Bellamy J, Case BS, Lee JE & Buckley HL (2014) Fine-scale spatial patterns in bacterial community composition and function within freshwater ponds. *ISME J.* **8**: 1715–26
57. Lee J (2014) Niche and neutral processes in aquatic bacterial communities: Are all things equal?
58. Leibold MA & McPeck MA (2006) Coexistence of the niche and neutral perspectives in community ecology. *Ecology* **87**: 1399–1410
59. Lepère C, Domaizon I, Taïb N, Mangot J-F, Bronner G, Boucher D & Debroas D (2013) Geographic distance and ecosystem size determine the distribution of smallest protists in lacustrine ecosystems. *FEMS Microbiol. Ecol.* **85**: 85–94
60. Levins R & Lewontin RC (1985) Dialectics and reductionism in ecology. 133-160 en *The dialectical biologist*. 1era edición Harvard University Press, USA

61. Lindström ES & Langenheder S (2012) Local and regional factors influencing bacterial community assembly. *Environ. Microbiol. Rep.* **4**: 1–9
62. Logares R, Lindström ES, Langenheder S, Logue JB, Paterson H, Laybourn-Parry J, Rengefors K, Tranvik L & Bertilsson S (2013) Biogeography of bacterial communities exposed to progressive long-term environmental change. *ISME J.* **7**: 937–48
63. Logue JB, Mouquet N, Peter H & Hillebrand H (2011) Empirical approaches to metacommunities: a review and comparison with theory. *Trends Ecol. Evol.* **26**: 482–91
64. Lomolino M V., Riddle BR & Brown JH (2006) The science of biogeography. In *Biogeography* pp 3–12.
65. López-Lozano NE, Eguiarte LE, Bonilla-Rosso G, García-Oliva F, Martínez-Piedragil C, Rooks C & Souza V (2012) Bacterial communities and the nitrogen cycle in the gypsum soils of Cuatro Ciénegas Basin, coahuila: a Mars analogue. *Astrobiology* **12**: 699–709
66. López-Lozano NE, Heidelberg KB, Nelson WC, García-Oliva F, Eguiarte LE & Souza V (2013) Microbial secondary succession in soil microcosms of a desert oasis in the Cuatro Cienegas Basin, Mexico. *PeerJ* **1**: e47
67. MacArthur RH & Wilson EO (1967) The theory of island biogeography Princeton University Press.
68. Martiny AC, Treseder K & Pusch G (2013) Phylogenetic conservatism of functional traits in microorganisms. *ISME J.* **7**: 830–8
69. Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR, Morin PJ, Naeem S, Ovreås L, Reysenbach A-L, Smith VH & Staley JT (2006) Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.* **4**: 102–12
70. Martiny JBH, Eisen JA, Penn K, Allison SD & Horner-Devine MC (2011) Drivers of bacterial beta-diversity depend on spatial scale. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 7850–4
71. Morlon H, Chuyong G, Condit R, Hubbell S, Kenfack D, Thomas D, Valencia R & Green JL (2008) A general framework for the distance-decay of similarity in ecological communities. *Ecol. Lett.* **11**: 904–917
72. Nekola J & White P (1999) The distance decay of similarity in biogeography and ecology. *J. Biogeogr.* **26**: 867–878
73. Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, Knelman JE, Darcy JL, Lynch RC, Wickey P & Ferrenberg S (2013) Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* **77**: 342–56
74. Ng ISY, Carr CM & Cottenie K (2008) Hierarchical zooplankton metacommunities: distinguishing between high and limiting dispersal mechanisms. *Hydrobiologia* **619**: 133–143
75. Noguez A, Arita T, Escalante AE, Forney LJ, García-Oliva F & Souza V (2005) Microbial macroecology: highly structured prokaryotic soil assemblages in a tropical deciduous forest. *Global Ecol. Biogeogr.* **14**: 241–248.

76. Östman Ö, Drakare S, Kritzberg E, Langenheder S, Logue J & Lindström E (2012) Importance of space and the local environment for linking local and regional abundances of microbes. *Aquat. Microb. Ecol.* **67**: 35–45
77. Ostman O, Drakare S, Kritzberg ES, Langenheder S, Logue JB & Lindström ES (2010) Regional invariance among microbial communities. *Ecol. Lett.* **13**: 118–27
78. Pajares S, Bonilla-Rosso G, Travisano M, Eguiarte LE & Souza V (2012) Mesocosms of aquatic bacterial communities from the Cuatro Ciénegas Basin (Mexico): a tool to test bacterial community response to environmental stress. *Microb. Ecol.* **64**: 346–58
79. Pajares S, Eguiarte LE, Bonilla-Rosso G & Souza V (2013) Drastic changes in aquatic bacterial populations from the Cuatro Ciénegas Basin (Mexico) in response to long-term environmental stress. *Antonie Van Leeuwenhoek* **104**: 1159–1175
80. Pandit SN, Kolasa J & Cottenie K (2009) Contrasts between habitat generalists and specialists: An empirical extension to the basic metacommunity framework. *Ecology* **90**: 2253–2262
81. Pedrós-Alió C (2012) The Rare Bacterial Biosphere. *Ann. Rev. Mar. Sci.* **4**: 449–466
82. Peimbert M, Alcaraz LD, Bonilla-Rosso G, Olmedo-Alvarez G, García-Oliva F, Segovia L, Eguiarte LE & Souza V (2012) Comparative metagenomics of two microbial mats at Cuatro Ciénegas Basin I: ancient lessons on how to cope with an environment under severe nutrient stress. *Astrobiology* **12**: 648–58
83. Pommier T, Neal P, Gasol J, Coll M, Acinas S & Pedrós-Alió C (2010) Spatial patterns of bacterial richness and evenness in the NW Mediterranean Sea explored by pyrosequencing of the 16S rRNA. *Aquat. Microb. Ecol.* **61**: 221–233
84. Prosser JI, Bohannan BJM, Curtis TP, Ellis RJ, Firestone MK, Freckleton RP, Green JL, Green LE, Killham K, Lennon JJ, Osborn AM, Solan M, van der Gast CJ & Young JPW (2007) The role of ecological theory in microbial ecology. *Nat. Rev. Microbiol.* **5**: 384–392
85. Rainey PB & Travisano M (1998) Adaptive radiation in a heterogeneous environment. *Nature* **32**: 69–72
86. Rebollar EA, Avitia M, Eguiarte LE, González-González A, Mora L, Bonilla-Rosso G & Souza V (2012) Water-sediment niche differentiation in ancient marine lineages of *Exiguobacterium* endemic to the Cuatro Ciénegas Basin. *Environ. Microbiol.* **14**: 2323–33
87. Ricklefs RE (2004) A comprehensive framework for global patterns in biodiversity. *Ecol. Lett.* **7**: 1–15
88. Rodríguez P & Vázquez-Domínguez E (2003) Escalas y diversidad de especies. In *Una perspectiva latinoamericana de la biogeografía* pp 109–114.

89. Rodríguez-Verdugo A, Souza V, Eguiarte LE & Escalante AE (2012) Diversity across Seasons of Culturable Pseudomonas from a Desiccation Lagoon in Cuatro Ciénegas, Mexico. *Int. J. Microbiol.* **2012**: Article ID 201389, 10 pages
90. Schmidt SK, Nemergut DR, Darcy JL & Lynch R (2014) Do bacterial and fungal communities assemble differently during primary succession? *Mol. Ecol.* **23**: 254–258.
91. Sloan WT, Lunn M, Woodcock S, Head IM, Nee S & Curtis TP (2006) Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ. Microbiol.* **8**: 732–740
92. Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR, Arrieta JM & Herndl GJ (2006) Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 12115–20
93. Soininen J (2014) A quantitative analysis of species sorting across organisms and ecosystems. *Ecology*: 140619091522005
94. Soininen J, McDonald R & Hillebrand H (2007) The distance decay of similarity in ecological communities. *Ecography (Cop.)*. **30**: 3–12
95. Souza V, Eguiarte LE, Siefert J & Elser JJ (2008) Microbial endemism: does phosphorus limitation enhance speciation? *Nat. Rev. Microbiol.* **6**: 559–64
96. Souza V, Espinosa-asuar L, Escalante AE, Eguiarte LE, Farmer J, Lloret L, Rodríguez-martínez JM, Soberón X, Dirzo R, Elser J, Forney L, Sobero X & Elser JJ (2006) An endangered oasis of aquatic microbial biodiversity in the Chihuahuan desert. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 6565–70
97. Souza V, Siefert JL, Escalante AE, Elser JJ & Eguiarte LE (2012) The Cuatro Ciénegas Basin in Coahuila, Mexico: an astrobiological Precambrian Park. *Astrobiology* **12**: 641-647.
98. Székely AJ & Langenheder S (2014) The importance of species sorting differs between habitat generalists and specialists in bacterial communities. *FEMS Microbiol. Ecol.* **87**: 102–12
99. Thompson R & Townsend C (2006) A truce with neutral theory: Local deterministic factors, species traits and dispersal limitation together determine patterns of diversity in stream invertebrates. *J. Anim. Ecol.* **75**: 476–484
100. Torsvik V & Øvreås L (2002) Microbial diversity and function in soil : from genes to ecosystems. *Curr. Opin. Microbiol.* **5**: 240–245
101. Unmack P & Minckley W (2008) The demise of desert springs. En *Aridland Springs in North America: Ecology and Conservation*, Lawrence ES & Meretsky JV (eds) The University of Arizona Press and the Arizona-Sonora Desert Museum, Tucson
102. Valdivia-Anistro J, Delgado G, Márquez P, Eguiarte LE, Olmedo G, Elser JJ & Souza V. (2015) Variability of rRNA operon copy number and growth rate dynamics of *Bacillus* isolated from an extremely oligotrophic aquatic ecosystem. Manuscrito en preparación.

103. Valiente-B A, Ramírez E, Verdú M & Montesinos A (2014) Plant community ecology. En *Frontiers in Ecology, Evolution and Complexity* Benítez M, Miramontes O & Valiente-Banuet A (eds) EditoraC3 (open access), México.
104. van der Heijden MG, Bardgett RD & van Straalen NM. (2008) The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol. Lett.* **11**: 296-310.
105. Vartoukian SR, Palmer RM & Wade WG. (2010) Strategies for culture of 'unculturable' bacteria. *FEMS Microbiol. Lett.* **309**: 1-7.
106. Vellend M (2010) Conceptual synthesis in community ecology. *Q. Rev. Biol.* **85**: 183–206
107. Wang J, Shen J, Wu Y, Tu C, Soininen J, Stegen JC, He J, Liu X, Zhang L & Zhang E (2013) Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *ISME J.* **7**: 1310–21
108. West SA, Diggle SP, Buckling A, Gardner A, & Griffin AS (2007) The Social Lives of Microbes *Annual Review of Ecology, Evolution, and Systematics* **38**: 53-77
109. Whittaker RH (1956) Vegetation of the Great Smoky Mountains. *Ecological Monographs* **26**:1-80. Whittaker RH (1960) Vegetation of the Sisiyou Mountains, Orgeon and California. *Ecol. Monogr.* **30**: 279–338
110. Williamson SJ & Yooseph S (2012) From bacterial to microbial ecosystems (metagenomics). In *Bacterial Molecular Networks*, van Helden J Toussaint A & Thieffry D (eds) pp 35–55. New York, NY: Springer New York
111. Wolaver BD, Sharp JM, Rodriguez JM & Flores JCI (2008) Delineation of regional arid karstic aquifers: an integrative data approach. *Ground Water* **46**: 396–413
112. Wolaver B.D. Crossey LJ, Karlstrom K.E, Banner JL, Cardenas MB, Gutiérrez Ojeda C & Sharp Jr JM (2013) Identifying origins of and pathways for spring waters in a semiarid basin using He, Sr, and C isotopes: Cuatrociénegas Basin, Mexico. *Geosphere* **9**:113-125.
113. Woodcock S, van der Gast CJ, Bell T, Lunn M, Curtis TP, Head IM & Sloan WT (2007) Neutral assembly of bacterial communities. *FEMS Microbiol. Ecol.* **62**: 171–80
114. Zaremba-Niedzwiedzka K, Viklund J, Zhao W, Ast J, Sczyrba A, Woyke T, McMahon K, Bertilsson S, Stepanauskas R & Andersson SGE (2013) Single-cell genomics reveal low recombination frequencies in freshwater bacteria of the SAR11 clade. *Genome Biol.* **14**: R130
115. Zarronaindia I, Smith D & Gilbert J (2013) Beyond the genome: community-level analysis of the microbial world. *Biol. Philos.* **28**: 261–282

116. Zheng Y-M, Cao P, Fu B, Hughes JM & He J-Z (2013) Ecological drivers of biogeographic patterns of soil archaeal community. *PLoS One* **8**: e63375
117. Zinger L, Boetius a & Ramette a (2014) Bacterial *taxa*-area and distance-decay relationships in marine environments. *Mol. Ecol.* **23**: 954–64
118. Zinger L, Gobet A & Pommier T (2012) Two decades of describing the unseen majority of aquatic microbial diversity. *Mol. Ecol.* **21**: 1878–96