



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE MÉXICO

DOCTORADO  
en  
CIENCIAS  
BIOMÉDICAS

INSTITUTO DE ECOLOGÍA

PROGRAMA DE DOCTORADO  
EN CIENCIAS BIOMÉDICAS

*Exiguobacterium* como modelo para entender la relación del ambiente  
con los procesos evolutivos en bacterias en el valle de Cuatro Ciénegas,  
Coahuila.

T E S I S

QUE PARA OBTENER EL GRADO DE:

DOCTORA EN CIENCIAS

P R E S E N T A:

ERIA ALAIDE REBOLLAR CAUDILLO

Tutora: DRA. VALERIA SOUZA SALDÍVAR



INSTITUTO  
DE ECOLOGÍA  
UNAM

MÉXICO, D. F.

AGOSTO 2012



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

<b>0. RESUMEN/ABSTRACT.....</b>	<b>1-2</b>
<b>1. INTRODUCCIÓN.....</b>	<b>3</b>
1.1 Diversidad y evolución de los procariontes en ambientes naturales	3
1.1.1 La gran diversidad de procariontes en el planeta	4-8
1.1.2 ¿Qué sabemos sobre la evolución de los procariontes?	8-12
1.2 Divergencia genética y especiación en procariontes	12-13
1.2.1 El aislamiento geográfico (especiación alopátrica).	13-16
1.2.2 La selección del ambiente (especiación ecológica).	16-20
1.2.3 El papel de la recombinación en la divergencia genética.	20-22
<b>2. ANTECEDENTES.....</b>	<b>23</b>
2.1 Diversidad y evolución bacteriana en Cuatro Ciénegas Coahuila.	23-28
2.2 El género <i>Exiguobacterium</i> .	28-31
<b>3. JUSTIFICACIÓN DEL PROYECTO.....</b>	<b>32-33</b>
<b>4. ARTÍCULO 1 .....</b>	<b>34</b>
“Diferenciación de nicho agua-sedimento en linajes marinos antiguos de <i>Exiguobacterium</i> endémicos del valle de Cuatro Ciénegas”.	
4.1 Resumen	34
4.2 Artículo en inglés	35-54
<b>5. DIVERSIDAD GENÉTICA Y DISRIBUCIÓN DE <i>Exiguobacterium</i> EN CCC.....</b>	<b>55</b>
5.1 Genética de poblaciones comparada: el caso de <i>Exiguobacterium</i> con respecto a otros géneros en CCC.	55-62
5.2 La distribución de los filogrupos de <i>Exiguobacterium</i> en el valle de CCC.	62-67

<b>6. ARTÍCULO 2.....</b>	<b>68</b>
“ <i>Exiguobacterium halophilus</i> sp. nov. y <i>Exiguobacterium coahuilensis</i> sp. nov., dos nuevas especies halófilas aisladas del valle de Cuatro Ciénegas Coahuila, Mexico.*	
6.1 Resumen	68-69
6.2 Artículo en inglés	70-87
<b>7. DISCUSIÓN Y CONCLUSIONES.....</b>	<b>88</b>
7.1 Los linajes de <i>Exiguobacterium</i> en CCC son cercanos a especies de origen marino...	88-90
7.2 Los linajes de <i>Exiguobacterium</i> en CCC indican historias evolutivas diferentes.....	90-93
7.3 Los linajes de <i>Exiguobacterium</i> en CCC muestran un patrón de diferenciación.....	93-96
ecológica asociada a los ambientes de agua y sedimento.	
<b>8. PERSPECTIVAS.....</b>	<b>97</b>
8.1 Caracterización genotípica y fenotípica de los linajes de <i>Exiguobacterium</i> en CCC.	97-98
8.2 Genómica comparada de <i>Exiguobacterium</i> .	99-102
<b>9. REFERENCIAS.....</b>	<b>103-113</b>

---

\* Este artículo se encuentra en preparación y los nombres de las especies son nombres propuestos sujetos a cambios.



## 0. RESUMEN

Los procariontes son los organismos más diversos y abundantes del planeta. Uno de los temas centrales de la microbiología evolutiva es la comprensión de la enorme diversidad genética, metabólica y ecológica de los procariontes. Cuatro Ciénegas Coahuila (CCC) se ubica en el desierto chihuahuense y posee múltiples sistemas acuáticos oligotróficos con condiciones ambientales únicas. La gran heterogeneidad de este sitio permite explorar el papel de los factores ambientales y geográficos en la diversificación y adaptación de los procariontes a diferentes ambientes. Así mismo, la gran diversidad de especies bacterianas en este sitio, hace posible profundizar en los mecanismos que promueven la divergencia y diferenciación de estas poblaciones. En este trabajo se estudió la evolución molecular de 183 aislados del género *Exiguobacterium* en CCC. Este género consta de bacterias halófilas y alcalófilas que se encuentran presentes en los sedimentos y el agua de los sistemas acuáticos de CCC. Se obtuvo la secuencia completa del gen *16srRNA* así como la secuencia parcial de cuatro genes codificantes. A partir de la filogenia concatenada de los cuatro genes codificantes se definieron tres filogrupos con diferentes valores de diversidad genética. Estos filogrupos no se encuentran asociados a un sistema acuático o a una salinidad específica. Sin embargo, la reconstrucción filogenética indicó la existencia de subgrupos compuestos de organismos asociados ya sea al sedimento o al agua. Este patrón de diferenciación se demostró mediante múltiples análisis estadísticos. Los resultados obtenidos confirman observaciones previas de la existencia de linajes bacterianos asociados a nichos ecológicos independientes; siendo este el primer trabajo que describe un evento de divergencia ecológica asociada a hábitats de sedimento y agua. Este trabajo contribuye de manera importante al entendimiento de los procesos ecológicos que influyen sobre la evolución de las poblaciones bacterianas naturales.

## ABSTRACT

Prokaryotes are both the most diverse and abundant organisms in the planet. One of the central subjects in evolutionary microbiology is to understand the reasons for this enormous genetic, metabolic and ecologic diversity. The Cuatro Ciénegas Basin (CCB) is part of the Chihuahuan desert and it comprises several oligotrophic aquatic systems with contrasting environmental conditions. The heterogeneity of this place makes it an excellent model to explore the role of ecological and geographical factors in the diversification and adaptation of prokaryotes to different environments. Moreover, the great diversity of bacterial species in this site, allows for a deeper understanding of the mechanisms involved in population divergence and differentiation. We studied the molecular evolution of 183 isolates of the *Exiguobacterium* genus from the CCB. This genus is comprised by halophilic and alkaliphilic bacteria and is extremely abundant in the sediments and water from the CCB aquatic systems. The complete sequence of the *16srRNA* gene and four partial housekeeping genes were obtained. Based on the concatenated phylogeny of the four coding genes we defined three main phylogroups with contrasting genetic diversity values. These phylogroups are not specific to an aquatic system or to a determined salinity. However the phylogenetic reconstruction indicates the existence of several small clusters composed by isolates from sediment or water only. In addition, multiple statistical analyses revealed a clear differentiation pattern associated to sediment and water habitats. The results shown here support previous observations of niche differentiation in other bacterial lineages, and represent the first work showing ecological divergence associated to water and sediment habitats within prokaryotic populations. This work contributes to the understanding of the ecological processes involved in the evolution of natural populations of bacteria.

# 1. INTRODUCCIÓN

## 1.1 Diversidad y evolución de los procariontes en ambientes naturales

La distinción entre los organismos procariontes y eucariontes fue descrita por primera vez en 1937 por Edouard Chaton (Sapp 2005). Casi sesenta años después, Carl Woese propuso que la vida se organizaba en tres dominios (Woese *et al.*, 1990), de los cuáles dos están conformados por organismos microscópicos y procariontes. Hoy en día, sabemos que los procariontes son los organismos más antiguos, abundantes y diversos del planeta. Los primeros acercamientos de los científicos a los organismos procariontes fueron gracias a las enfermedades humanas causadas por éstos. A la fecha, la gran mayoría de los genomas secuenciados corresponden a organismos de importancia médica o agrícola. Por lo anterior, los microorganismos patógenos (en particular las bacterias) son hasta la fecha los más estudiados, sin embargo, la mayor parte de la diversidad procarionte se encuentra en otros ambientes naturales (no asociados al humano)(Whitman *et al.*, 1998). Es en estos ambientes naturales, en donde se encuentran las comunidades más complejas de procariontes, así como las más diversas genética y metabólicamente (DeLong y Pace 2001).

El estudio de la diversidad genética de las poblaciones bacterianas actuales es la herramienta fundamental para entender cómo surgen los linajes bacterianos y qué mecanismos genéticos son importantes para el surgimiento de nuevas especies microbianas. El entender la evolución microbiana, no sólo es importante para explicarnos la gran diversidad de formas que existen en el planeta, sino también es fundamental para entender el surgimiento y propagación de linajes bacterianos patógenos y oportunistas.

En esta primera sección de la introducción se discutirán los principales hallazgos sobre la diversidad de las poblaciones procariontes en ambientes naturales así como los mecanismos evolutivos que moldean las poblaciones procariontes.

### 1.1.1 La gran diversidad de procariontes en el planeta

Los procariontes son los organismos más abundantes y diversos. Se estima que en el planeta existen alrededor de  $4 \times 10^{30}$  células procariontes distribuidas principalmente en los océanos, lagunas, ríos, suelos y sedimentos (Whitman *et al.*, 1998). Debido a su abundancia, los procariontes representan la mayor reserva de nutrientes en el planeta, incluso por arriba de la concentración de nutrientes presentes en las plantas (Whitman *et al.*, 1998). Debido a lo anterior y al gran repertorio metabólico que poseen, los procariontes son esenciales en el flujo de nutrientes (ciclos biogeoquímicos) en todos los ecosistemas del planeta, incluso en aquellos en donde los eucariontes no son capaces de desarrollarse.

Además de la gran abundancia y diversidad metabólica de estos organismos, los procariontes son genéticamente muy diversos. Lo anterior se debe probablemente a los altos tamaños poblacionales y los tiempos generacionales cortos (alta tasa de división celular) característicos de este grupo. Esto ha permitido sugerir que los procariontes tienen tasas bajas de extinción y altas tasas de especiación (Dykhuizen, 1998).

Debido a su tamaño y morfología, los procariontes sólo pueden ser estudiados mediante métodos de gran resolución que permitan analizar su metabolismo y sus rasgos genéticos. En las tres últimas décadas se han desarrollado una gran cantidad de técnicas moleculares para el estudio de la vida microscópica. Gracias a éstas, el estudio de la diversidad y evolución de los microorganismos se ha desarrollado ampliamente.

Por un lado, el desarrollo de técnicas independientes de cultivo ha permitido estimar la enorme diversidad de organismos procariontes presentes en el planeta (DeLong y Pace, 2001). Asimismo, estas técnicas han permitido detectar vida microscópica en

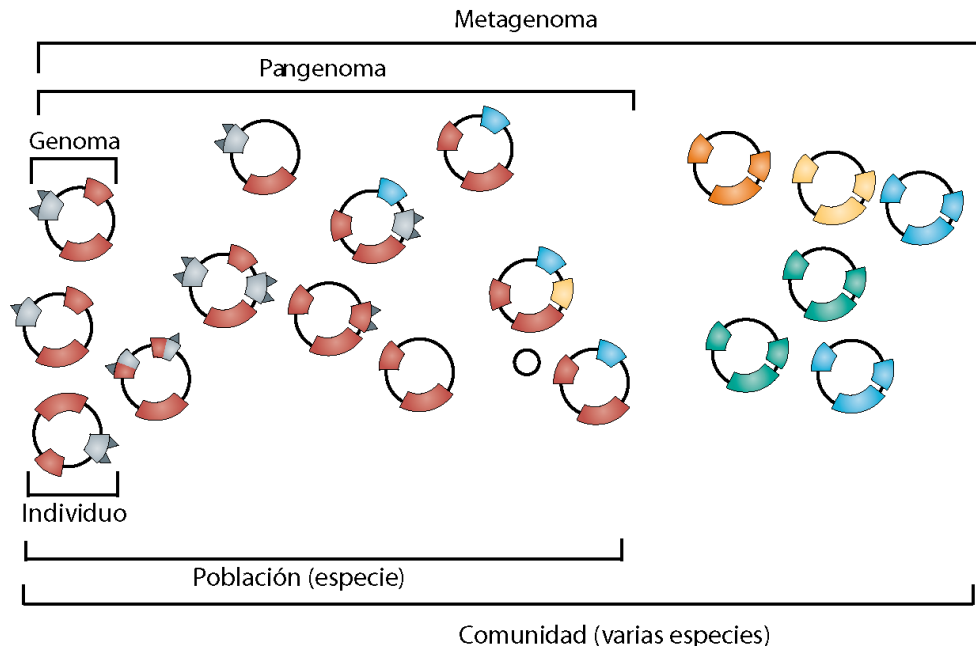
lugares que antes se consideraban inaccesibles para ésta, como es el caso de las ventilas hidrotermales, las regiones polares, las lagunas ácidas o hipersalinas y las zonas de actividad volcánica (Pikuta *et al.*, 2007; Chanal *et al.*, 2006; Barker y Boetius 2007). Por otro lado, gracias al análisis de metagenomas<sup>1</sup> de ambientes dominados por microorganismos se ha comenzado a entender la estructura de las comunidades bacterianas, las funciones de los microorganismos en una comunidad y el papel ecológico que cumplen en un ecosistema (Forney *et al.*, 2004; Whitaker y Banfield 2006; Xu, 2006). Antes del surgimiento de estas técnicas, la microbiología se enfocaba en el estudio de organismos cultivables que representan aproximadamente sólo el 1% de la diversidad microbiana (Oren, 2004).

Por otra parte, la secuenciación de DNA de organismos cultivables ha permitido determinar las relaciones evolutivas entre individuos mediante la comparación de secuencias, genes o genomas completos (Abby y Daubin 2007; Medini *et al.*, 2008).

El surgimiento de estrategias moleculares para analizar secuencias de DNA, ha permitido explorar la diversidad genética de los microorganismos a diferentes escalas (Figura 1). Por un lado, es posible describir a detalle el contenido genómico de un individuo. Sin embargo para la ecología y evolución molecular es más informativo describir la diversidad genómica de una especie mediante la secuenciación de una muestra de una población completa. Además, las técnicas independientes de cultivo permiten el análisis de ecosistemas microbianos completos mediante la secuenciación completa de los genes contenidos en una comunidad microbiana (Allen y Banfield 2005). Estos tres niveles de análisis son esenciales para el entendimiento de la evolución de los linajes procariontes así como de la dinámica de los ecosistemas microbianos.

---

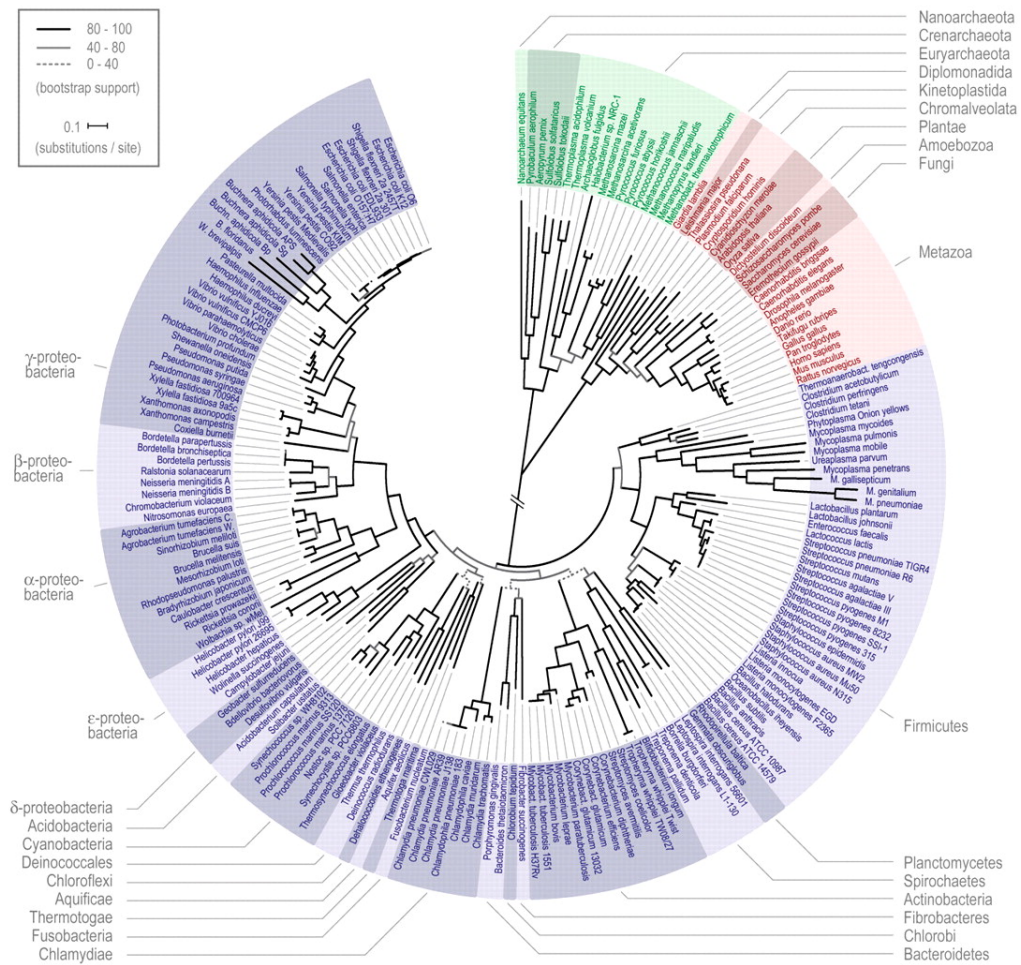
<sup>1</sup> Repertorio genético global de un nicho ambiental que está compuesto por microorganismos de vida libre o por microorganismos comensales de animal o planta hospedero (Medini *et al.*, 2008)



**Figura 1**

Análisis de la diversidad genética a distintos niveles: genoma, pangenoma y metagenoma. La información contenida en un genoma no es suficiente para describir la diversidad intra-especie. Por ende es necesario analizar múltiples genomas dentro de una especie (población). Para estudiar las relaciones inter-especie y la dinámica funcional de un ecosistema es necesario describir la comunidad completa.  
(tomada y modificada de Medini *et al.*, 2008).

Como evidencia de la gran diversidad microbiana en diversos ecosistemas, en la última década se han realizado estudios en comunidades microbianas que son extremadamente diversas genéticamente tanto a nivel de género y especie así como al interior de las poblaciones (Hewson y Furham 2004; Acinas *et al.*, 2004; Fourny *et al.*, 2004; Souza *et al.*, 2006). Hasta la fecha, la secuencia del gen 16s ribosomal (*16srRNA*) es el marcador universal para describir la diversidad del mundo procarionte (Woese, 1987). Sin embargo, recientemente se han utilizado secuencias de múltiples genes ubicuos que son menos conservados que el gen *16srRNA* y que por ende ofrecen una mayor resolución para determinar relaciones filogenéticas (Maiden 2006).



**Figura 2**

Filogenia concatenada de 31 secuencias de proteínas a partir de 191 genomas. En morado se muestra el dominio bacteriano, en verde el dominio arqueano y en rojo el dominio eucariote. (tomada de Ciccarelli *et al.*, 2006).

Mas allá del análisis de unos cuantos genes, la secuenciación de genomas completos ha permitido describir la diversidad genética en procariontes con mucho mayor resolución y así determinar que el dominio bacteriano es el genéticamente más diverso (Figura 2). Al comparar las secuencias de 31 proteínas conservadas a partir de los genomas de 191 organismos de los tres dominios, Ciccarelli y colaboradores muestran como la distancia genética entre el humano y el ratón es equivalente a la distancia entre dos especies de bacterias del mismo género (por ejemplo *Pseudomonas* o *Streptococcus*).

A partir de esta y otras observaciones surge la pregunta de ¿Porqué las bacterias son tan diversas? y ¿Qué mecanismos evolutivos han permitido o favorecido el surgimiento de tantos linajes?.

### 1.1.2 ¿Qué sabemos sobre la evolución de los procariontes?

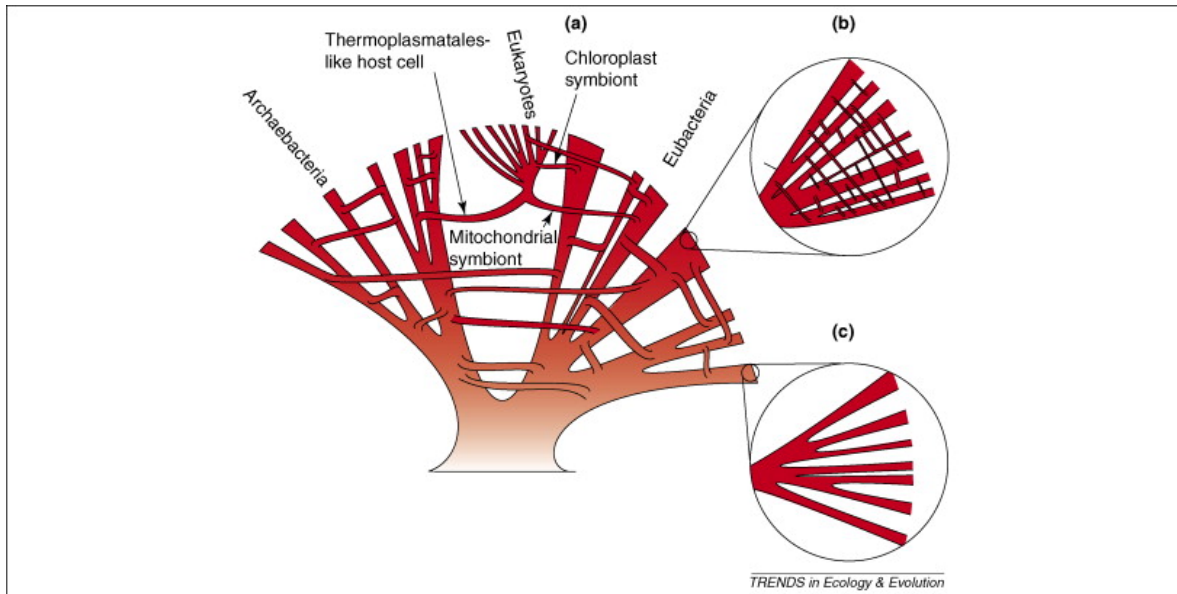
El análisis de la diversidad genética dentro de las poblaciones, ha sido hasta la fecha, la aproximación fundamental para determinar las fuerzas evolutivas que moldean las poblaciones de bacterias patógenas (Feil y Spratt 2001; Castillo *et al.*, 2005) y ambientales (Whitaker *et al.*, 2005; Silva *et al.*, 2005). Además de la comparación de genes específicos en una población, hoy en día el análisis de genomas completos es una herramienta importante para entender la evolución en procariontes (Peña *et al.*, 2010; Konstantinidis *et al.*, 2009; Cadillo-Quiroz *et al.*, 2012). En particular, la genómica de poblaciones es un acercamiento esencial, ya que a partir de la comparación de genomas entre individuos de la misma especie es posible determinar el pangenoma de la especie y explorar las fuerzas evolutivas involucradas en la divergencia genética entre linajes (Figura 1).

Asimismo, la diferenciación genética entre poblaciones también se puede detectar mediante métodos de genotipificación. Los patrones genotípicos cambian dependiendo de la divergencia a nivel de la secuencia debido al movimiento de elementos transponibles, inserciones o rearrreglos genómicos. Algunas técnicas utilizadas son los RAPDs (“Random Amplified Polymorphic DNA”), BOX-PCR (“BOX repetitive elements PCR”) y REP-PCR (“Repetitive extragenic palindromic PCR”), los cuales podrían detectar diferenciación más reciente que el análisis de genes específicos (Cho y Tiedje 2000; Foti *et al.* 2006; Sikorski y Nevo 2005).



Algunas de las preguntas que los microbiólogos evolutivos pretenden contestar a partir del análisis de la diversidad genética son: ¿Cómo definir una especie procarionte? y ¿Qué fuerzas evolutivas promueven el surgimiento de nuevos linajes microbianos?. En particular el estudio de los mecanismos de especiación, es uno de los temas fundamentales para entender el surgimiento y distribución de las poblaciones bacterianas, los fenómenos de patogenicidad y exploración de nuevos nichos ecológicos; así como las fuerzas evolutivas que mantienen o diluyen la cohesión al interior de las especies microbianas (Hanage *et al.*, 2006a).

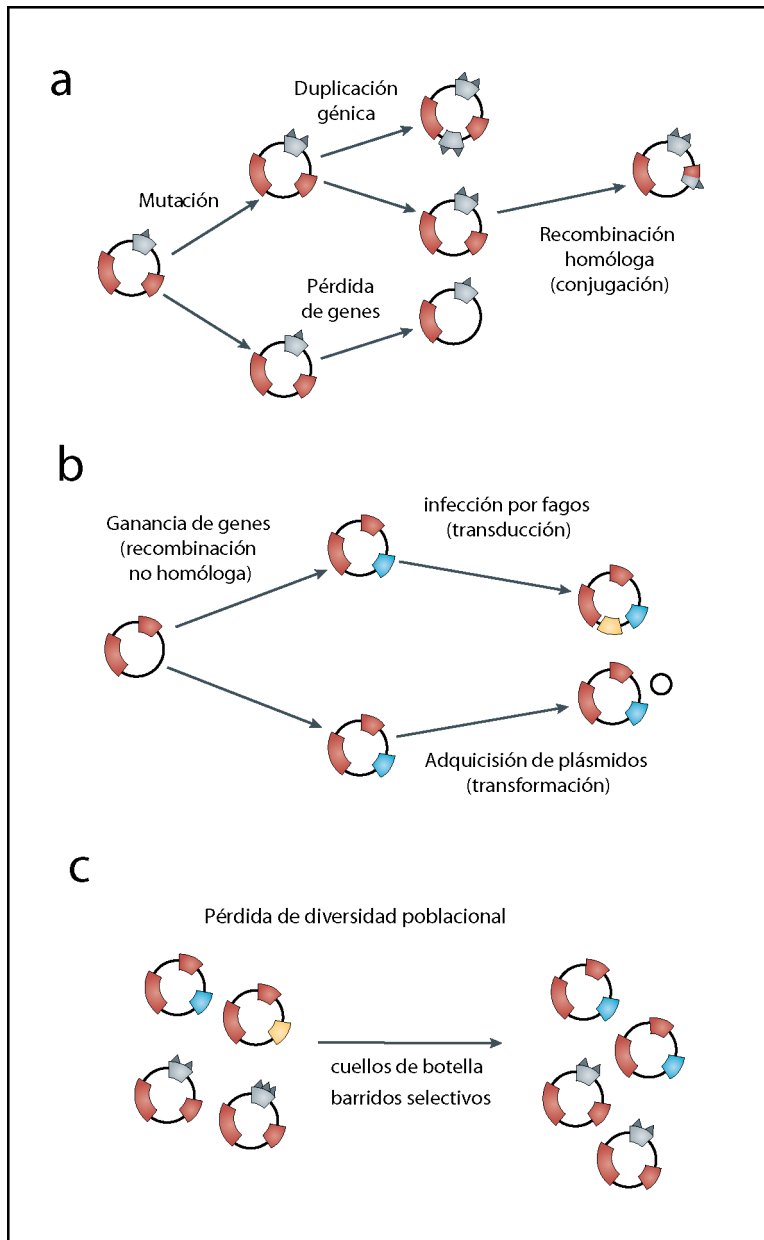
Las fuerzas evolutivas que generan y modulan la diversidad en las poblaciones naturales han sido ampliamente descritas en organismos diploides con reproducción sexual (macroorganismos principalmente). Los microorganismos procariontes se reproducen asexualmente, lo que los lleva a poseer linajes clonales. Con base en lo anterior, la fuente fundamental de la variabilidad en bacterias sería la mutación, sobre la cuál actúan la selección y la deriva para generar la diversidad y la adaptación microbiana (Cohan, 2002). Sin embargo, en las dos últimas décadas han surgido evidencias que muestran que la recombinación homóloga y la recombinación no homóloga (transferencia horizontal) son procesos muy frecuentes en las poblaciones bacterianas (Souza *et al.*, 1992; Istock *et al.*, 1992; Ochman *et al.*, 2000; Feil y Spratt 2001; Hanage *et al.*, 2005). Incluso, se ha llegado a proponer que la transferencia horizontal podría llegar a borrar las huellas filogenéticas entre linajes, dando lugar a una red de la vida en vez de un árbol de la vida (Figura 3; McInerney *et al.*, 2008). Sin embargo, más allá de la prevalencia de la recombinación entre y dentro de las poblaciones, diversos estudios muestran que es posible reconocer la presencia de unidades o grupos genéticos que nos permiten rastrear la historia evolutiva de los linajes procariontes (Hanage *et al.*, 2005; Papke *et al.*, 2007; Vos 2010).



**Figura 3**

Representación de la red de la vida. Los eucariotes se muestran como organismos quiméricos en donde la mitocondria y el cloroplasto tienen un origen diferente a los genes nucleares **(a)**. En los linajes bacterianos se pueden encontrar linajes altamente recombinantes o panmíticos **(b)** así como linajes totalmente clonales **(c)**. (Tomada de McInerney *et al.*, 2008).

Con base en lo anterior, la diversidad genética en poblaciones procariontes se genera fundamentalmente a partir de eventos de mutación y duplicación génica. Asimismo, estos microorganismos tienen la capacidad de adquirir material genético del exterior mediante conjugación, transformación y transducción (Jain *et al.*, 2002). Mediante estos tres mecanismos moleculares, los procariontes adquieren variación por medio de la recombinación homóloga entre organismos genéticamente cercanos. Por otra parte la recombinación no homóloga (transferencia horizontal) permite la adquisición de genes nuevos que provienen incluso de linajes procariontes muy distantes (Figura 4a, 4b; Medini *et al.*, 2008).



**Figura 4**

Mecanismos de ganancia y pérdida de diversidad al interior de las especies bacterianas. El cambio en la diversidad genética al interior de una población se puede generar **(a)** mediante mutación, duplicación o deleción génica, así como recombinación homóloga. **(b)** Por otra parte se pueden adquirir nuevos alelos mediante eventos de recombinación no homóloga. **(c)** Esta diversidad puede estar sujeta a eventos demográficos (deriva génica) y/o selectivos que purgan la diversidad. (modificada de Medini *et al.*, 2008).

Según la teoría evolutiva, la diversidad genética al interior de las poblaciones está a su vez sujeta a eventos que purgan la diversidad tales como la deriva y los barridos selectivos (Figura 4c). Con base en lo anterior, algunas de las preguntas fundamentales

para entender la especiación procarionte son ¿Cuáles son las barreras que limitan los mecanismos de recombinación en procariontes y permiten el surgimiento de linajes independientes? ¿Estas barreras son geográficas o ecológicas? ¿Cuál es el papel de la selección natural en la divergencia de las poblaciones naturales?. En el siguiente capítulo se abordarán las principales evidencias que contribuyen a entender los mecanismos de divergencia genética y especiación en procariontes.

## 1.2 Divergencia genética y especiación en procariontes

Como ya se mencionó en la sección anterior, la mutación y la recombinación son fuente de la amplia diversidad procarionte sobre la cual actúan las diferentes fuerzas evolutivas. Sin embargo, para entender el verdadero origen de las especies bacterianas, es esencial considerar el papel de las barreras ecológicas y geográficas en el proceso de divergencia genética entre poblaciones. Por medio de las técnicas moleculares recientes y de la genética de poblaciones hemos comenzado a entender algunos de los procesos involucrados en la especiación procarionte. Estos procesos pueden estar determinados por el aislamiento geográfico entre poblaciones (especiación alopátrica) así como por la selección del ambiente a nivel local (especiación ecológica) (Whitaker, 2006).

Para poder profundizar en los procesos evolutivos que llevan a la diversificación microbiana es necesario ampliar los análisis genéticos de poblaciones silvestres, ya que la mayor parte de la información genética multilocus se enfoca en el análisis de bacterias patógenas (Maiden, 2006). Una de las estrategias más utilizadas para entender los eventos de divergencia genética en las poblaciones ha sido la construcción de filogenias moleculares. Éstas nos permiten entender los procesos de diversificación y estructuración genética que dan como resultado los patrones biogeográficos que observamos hoy en día (Papke *et al.*, 2003; Jones *et al.*, 2006; Whitaker *et al.*, 2003).

Las evidencias encontradas a la fecha sugieren fuertemente que el aislamiento geográfico, ecológico y reproductivo, durante periodos largos y continuos, son elementos importantes para el surgimiento de endemismos microbianos en las poblaciones naturales (Souza *et al.*, 2008). En este capítulo de la introducción se hará una revisión de los principales estudios que han explorado estos temas y han contribuido al conocimiento de la evolución en procariontes.

### **1.2.1 El aislamiento geográfico (especiación alopátrica).**

Los estudios sobre la distribución geográfica de los macroorganismos se remontan a los estudios de los naturalistas del siglo XVIII. Más de dos siglos investigación sobre el tema muestran claramente la existencia de patrones biogeográficos que estructuran las poblaciones y comunidades de animales y plantas (revisado en Hughes *et al.*, 2006). Sin embargo en microorganismos, no es sino hasta el surgimiento de diversas técnicas moleculares, que podemos estudiar la distribución de los procariontes en el planeta. En la actualidad, existen todavía debates sobre si los microorganismos muestran patrones biogeográficos asociados a procesos históricos y ambientales como aquellos observados en macroorganismos o si son organismos de excepción.

La evolución divergente de poblaciones geográficamente aisladas (especiación alopátrica) ha sido la más tangible y menos controversial en macroorganismos, sin embargo su contribución en microorganismos ha sido más difícil de determinar (Whitaker, 2006). La visión clásica acerca de la distribución microbiana, plantea un modelo en donde las barreras para la dispersión no afectan a los organismos unicelulares debido a su pequeño tamaño, su plasticidad metabólica y su abundancia (Finlay, 2002). Por lo anterior, se ha propuesto que los patrones de distribución observados dependerán de la presión selectiva propia de cada lugar. En esta última década, diversos estudios en bacterias y arqueas han mostrado que existen tanto barreras geográficas como

ambientales que afectan los patrones de distribución y contribuyen de manera importante a los procesos de diversificación y especiación (Papke *et al.*, 2003, Whitaker *et al.*, 2003). Los patrones biogeográficos observados muestran una amplia gama de distribuciones microbianas, desde las especies cosmopolitas (*Pseudomonas* y *Nitrosococcus*) (Cho y Tiedje 2000; Ward y O'Mullan 2002 respectivamente) hasta poblaciones endémicas y altamente restringidas (*Sulfolobus* y *Synechococcus sp.*) (Whitaker *et al.*, 2003 y Papke *et al.*, 2003 respectivamente) (Tabla 1).

Las poblaciones microbianas cosmopolitas se observan cuando la tasa de dispersión y de homogeneización de la poza génica (vía recombinación homóloga) es mayor a la tasa de evolución local. Las tasas de dispersión, dependen de las capacidades físicas de transporte y la sobrevivencia de los organismos. Los microorganismos con distribuciones globales en general poseen movimiento activo (por ejemplo por flagelos) o pasivo facilitado por la formación de estructuras de resistencia como las endosporas, biopelículas protectoras o la asociación con macroorganismos migratorios (Whitaker, 2006). Además de la capacidad de dispersión de las especies microbianas, la estructura geográfica de cada hábitat es importante para su distribución. Existen entonces hábitats estructurados a manera de islas, con grandes distancias que impiden la migración de un lugar a otro, o ambientes continuos que favorecen la distribución ubicua. Es posible que la distribución espacial de los microorganismos también se encuentre influenciada por las capacidades metabólicas de las especies ya que la capacidad de desarrollarse en ciertos ambientes dependerá de los diversos nutrientes que cada organismo pueda metabolizar (Whitaker, 2006).

En la **Tabla 1** se muestran los principales trabajos que han analizado la estructura genética de poblaciones procariontes silvestres, no patógenas, utilizando diferentes herramientas y diferentes marcadores moleculares de alta resolución, con el fin de determinar el papel del ambiente o el aislamiento geográfico en la divergencia evolutiva.

Como se observa en esta tabla, las metodologías utilizadas para determinar la genética de las poblaciones ambientales han sido muy variadas, lo cuál dificulta en muchos casos la comparación entre lo observado en diferentes poblaciones. Sin embargo, es claro que se han encontrado patrones genéticos que sugieren una estructuración geográfica o ambiental como parte de un proceso de divergencia evolutiva.

Tabla1. Estudios de estructura poblacional asociada al ambiente o la geografía

G	A	N	Marcadores usados	Especie	Habitat	Escala	Referencia
	X		Electroforesis enzimática multilocus	<i>Burkholderia cepacia</i>	Arroyos de agua dulce	Local (50km)	Wise <i>et al.</i> 1995
X			16S–23S rRNA, ITS y box-PCR	<i>Pseudomonas spp.</i>	Suelo	Global	Cho y Tiedje. 2000
		X	16S rRNA y un locus codificante	<i>Nitrosococcus oceani</i>	Marino	Global	Ward y O’Mullan, 2002
X			16S rRNA y ocho loci codificantes	<i>Sulfolobus islandicus</i>	Manantiales geotermales	Global	Whitaker <i>et al.</i> , 2003
X			16S rRNA, ITS	<i>Synechococcus sp.</i>	Manantiales geotermales	Global	Papke <i>et al.</i> , 2003
		X	Cinco loci codificantes	<i>Photobacterium phosphoreum</i>	Marino	Global	Ast y Dunlap, 2005
	X		Tres loci codificantes y RAPDs	<i>Bacillus simplex</i>	Suelo	Local (40km)	Sikorski y Nevo, 2005
X			Tres loci codificantes dos loci simbióticos	<i>Rhizobium gallicum</i>	Nódulos en suelo	Global	Silva <i>et al.</i> , 2005
X			16S, ITS y tres loci codificantes	<i>Mastigocladus laminosus</i>	Manantiales geotermales	Local (50km)	Miller <i>et al.</i> , 2006
		X	26,517 pb de secuencias sobrelapadas	<i>Thermotoga sp.</i>	Ventilas hidrotermales	Global	Nesbo <i>et al.</i> , 2006
X			rep-PCR	<i>Thioalkalivibrio sp.</i>	Lagos haloalcalinos	Global	Foti <i>et al.</i> , 2006
		X	Cuatro loci codificantes	<i>Bradyrhizobium</i>	Nódulos en suelo	Global	Vinuesa <i>et al.</i> , 2008
X			Cinco loci codificantes	<i>Myxococcus xanthus</i>	Suelo	Global	Vos y Velicer, 2008
	X		Un locus y genotipificación por DGGE	<i>Desulfobulbus</i>	Sedimento en ambiente estuarino	Local (20km)	Oakley <i>et al.</i> , 2010
	X		Cuatro loci codificantes y 20 genomas	<i>Vibrio spp.</i>	Marino costero	Local (columna de agua)	Hunt <i>et al.</i> , 2008 Shapiro <i>et al.</i> , 2012

**Tabla 1**

Trabajos en donde se detectó el papel de factores geográficos (G) o ambientales (A) en los patrones genéticos en poblaciones procariontes. También han sido incluido trabajos en donde no se encontró ningún patrón geográfico y/o ambiental (N) (Actualizada y modificada de Whitaker, 2006).

Considerando las evidencias obtenidas en la última década, los patrones de diferenciación genética entre poblaciones están, en su mayoría, asociadas al aislamiento a grandes distancias geográficas. Asimismo, es interesante considerar que estos patrones son independientes del estilo de vida. Por ejemplo, existe aislamiento geográfico en poblaciones de arqueas y cianobacterias hipertermófilas que están finamente adaptadas a ambientes extremos (Whitaker *et al.*, 2003; Papke *et al.*, 2003). Por otra parte existen organismos generalistas como *Pseudomonas* o *Myxococcus* que también presentan patrones asociados a la distancia (Cho y Tiedje 2000, Vos y Velicer 2008). Interesantemente, no existe hasta la fecha ningún microorganismo marino que muestre una estructura genética asociada a la geografía. Lo anterior sugiere que el flujo en los ambientes marinos contrarresta los efectos del aislamiento por distancia que se observan en organismos terrestres y de aguas continentales.

En conclusión, las evidencias encontradas a la fecha sugieren fuertemente que el aislamiento geográfico (especiación alopátrica) ha permitido la evolución divergente de poblaciones de arqueas y bacterias.

### **1.2.2 La selección del ambiente (especiación ecológica).**

En la sección anterior se discutieron algunas de las evidencias que sugieren que las bacterias si tienen límites o barreras para la dispersión, y que éstos límites se observan principalmente a escalas continentales (Cho y Tiedje 2000; Whitaker *et al.*, 2003; Vos y Velicer 2008). Sin embargo, la selección natural se ha propuesto como la fuerza fundamental que promueve la adaptación a escalas locales (Shapiro *et al.*, 2009; Vos, 2010).

La importancia de la selección natural para explicar los patrones biogeográficos y la estructura de las comunidades bacterianas se remonta a los trabajos de L.G.M Baas Becking y su colaborador y contemporáneo M.W. Beijerinck (De Wit y Bouvier, 2006).



“*Todo está en todos lados pero el ambiente selecciona*” (Baas Becking, 1934) es hoy en día uno de las frases más citadas y representa uno de los cuestionamientos más interesantes en el área de la microbiología evolutiva.

A pesar de que sabemos que las condiciones ambientales (bióticas y abióticas) son fundamentales en el proceso de adaptación, existen muy pocos estudios que hayan encontrado pruebas de barreras ecológicas y/o ambientales que promuevan el aislamiento de poblaciones bacterianas. La falta de evidencias al respecto se debe por un lado a la falta de estudios evolutivos en poblaciones silvestres procariontes. Por otro lado, determinar el nicho ecológico así como las presiones selectivas en los diferentes ambientes, es una tarea compleja en el caso de los microorganismos.

Uno de los fenómenos menos estudiados es la evolución de microorganismos simbiotes. En la última década, se ha demostrado que los microorganismos asociados a macroorganismos presentan patrones genéticos y fenotípicos asociados a la distribución de sus hospederos (Falush *et al.*, 2003; Jones *et al.*, 2006). Tal es el caso de la distribución geográfica observada en el simbiote *Vibrio fischeri* con respecto a su hospedero, el calamar perteneciente al género *Euprymna* (Jones *et al.*, 2006). En este trabajo se sugiere que la estructura genética de *V.fischeri* está determinada por la interacción con su hospedero pero también por factores ambientales como son las corrientes marinas en la región estudiada. Estudios adicionales filogeográficos de bacterias simbiotes, son esenciales para entender los patrones genéticos y las fuerzas evolutivas involucradas en la evolución de los microorganismos simbiotes.

En contraste con los estudios de microorganismos simbiotes, en los años recientes se han realizado trabajos importantes sobre los factores ecológicos involucrados en la estructuración de las poblaciones microbianas de vida libre. Con base en estudios de genotipificación y análisis de secuencias multilocus, se han encontrado patrones de

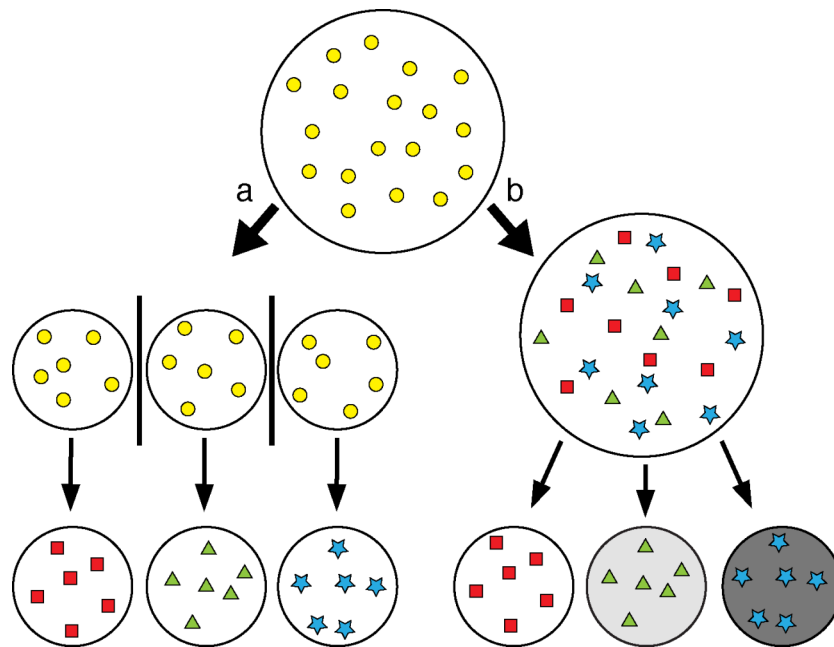
estructuración genética asociada a condiciones ambientales particulares. Tal es el caso de dos especies de bacterias de suelo (*Bacillus simplex* y *Desulfobulbus*), las cuales están estructuradas de acuerdo a un gradiente de radiación solar y de salinidad respectivamente (Sikorski y Nevo 2005; Sikorski y Nevo 2007; Oakley *et al.*, 2008). En ambos trabajos se sugiere que las condiciones ecológicas particulares promueven la diferenciación genética entre poblaciones y favorecen la especiación simpátrica. Además de las poblaciones bacterianas presentes en el suelo y los sedimentos, se han encontrado patrones de estructuración genética al interior de las poblaciones de vibrios marinos (Hunt *et al.*, 2008). En este trabajo, se analizaron 1000 aislados del género *Vibrio* y se determinó que existen subgrupos genéticos asociados a partículas planctónicas de diferentes tamaños. Recientemente, mediante la comparación de 20 genomas provenientes de estos aislados, se comprobó que la diferenciación genética es evidencia de un proceso de diferenciación ecológica reciente al interior de las poblaciones (Shapiro *et al.*, 2012). En este último estudio se detectaron genes únicos, así como polimorfismos únicos, asociados a cepas con nichos ecológicos diferentes. Lo anterior sugiere que las poblaciones microbianas están sujetas a procesos de divergencia genética como parte de un proceso de adaptación a ambientes y condiciones ecológicas diferentes.

Previo al análisis de genomas completos, era muy difícil encontrar señales de selección positiva entre poblaciones. Lo anterior se debe a que los marcadores más conservados utilizados en los análisis multilocus son generalmente genes constitutivos sujetos a selección purificadora.

La genómica comparada ha permitido detectar señales de selección positiva y de diferenciación ecológica al interior de las poblaciones (Bielawski *et al.*, 2004; Shapiro *et al.*, 2009; Shapiro *et al.*, 2012). Asimismo, la secuenciación de genomas individuales completos ha permitido detectar eventos de transferencia horizontal que están

asociados a procesos de adaptación a nichos ecológicos particulares. Tal es el caso de los estudios recientes con *Bacillus coahuilensis*, la bacteria hiperhalófila *Salinibacter ruber* y la arquea termoacidófila *Sulfolobus islandicus* (Alcaráz *et al.*, 2008; Peña *et al.*, 2010; Cadillo-Quiroz *et al.*, 2012).

Finalmente, los datos más recientes ya mencionados, sugieren fuertemente que la diferenciación genética entre poblaciones está determinada por factores ambientales (diferenciación ecológica) así como por barreras geográficas que limiten la dispersión (diferenciación alopátrica)(Figura 5). Los patrones biogeográficos que se observan en los organismos, son resultado de estos dos procesos de diferenciación, los cuales se distinguen mediante la partición de la variación en poblaciones ecológica o geográficamente aisladas. Es importante mencionar que estos patrones de diferenciación también están sujetos a los efectos de la deriva génica así como a los procesos de adaptación local.



**Figura 5**

Dos modelos de diferenciación genética a partir de una población inicial homogénea. **(a)** En la diferenciación alopátrica una población se separa en tres poblaciones aisladas por barreras geográficas. Este aislamiento permite la evolución independiente de cada población. **(b)** En la diferenciación ecológica la población homogénea evoluciona a una población heterogénea que diverge a partir de que cada genotipo explora un nicho ecológico independiente. (Adaptada de Whitaker, 2006).

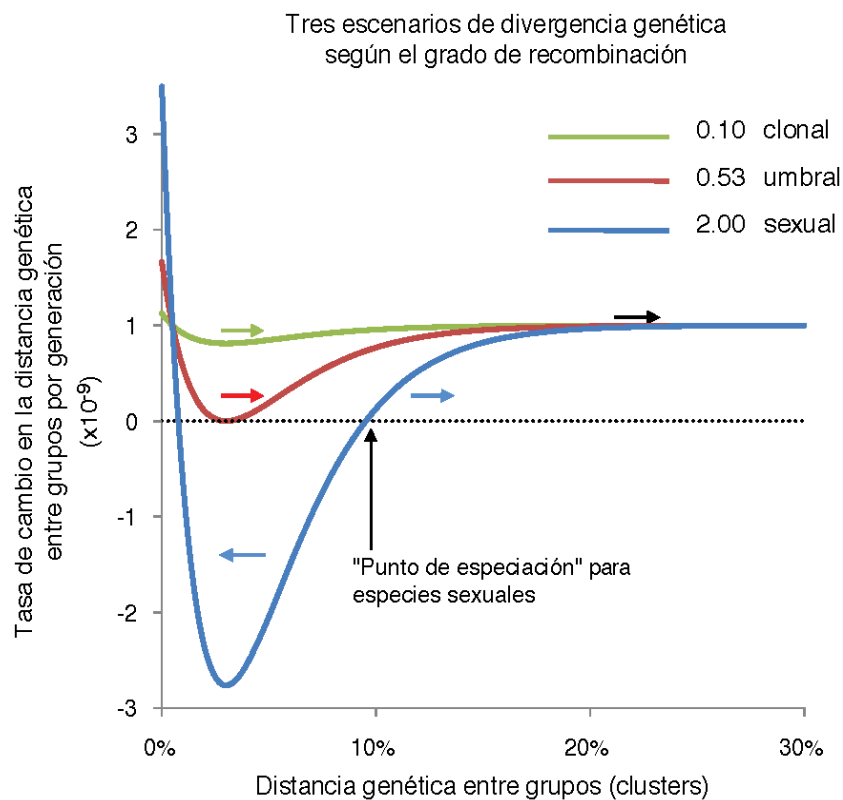
En particular para el caso de los microorganismos, las evidencias mencionadas en esta tesis, sugieren que a gran escala se logran observar patrones asociados al aislamiento geográfico. Estos patrones de diferenciación geográfica también se encuentran influenciados por eventos de deriva así como de adaptación local posterior a la separación geográfica de las poblaciones. Por otro lado, a escalas locales, las barreras ecológicas (dadas por las condiciones bióticas y abióticas) así como la exploración de nuevos nichos, serán factores determinantes para la diferenciación genética entre las poblaciones microbianas naturales.

### **1.2.3 El papel de la recombinación en la divergencia genética.**

Como se mencionó anteriormente, los procariontes son organismos que se reproducen de manera clonal. Debido a su forma de reproducción (asexual), la evolución microbiana suponía ser distinta de los organismos sexuales. Bajo este escenario, los microorganismos parecían estar sujetos a barridos selectivos en donde los ecotipos exitosos desplazan a los menos adaptados (Cohan y Perry 2007). Sin embargo, una gran cantidad de estudios recientes han demostrado que la recombinación (homóloga y no homóloga) es un mecanismo muy común y esencial para la generación de diversidad en procariontes (Gogarten *et al.*, 2002). Los eventos de recombinación han mostrado ser tan comunes dentro y entre las poblaciones que incluso se ha cuestionado profundamente la existencia de especies procariontes, consideradas como unidades evolutivas aisladas reproductivamente de otras (Hanage *et al.*, 2005; Doolittle y Papke 2006). Por otra parte, la transferencia horizontal es uno de los procesos esenciales para la adquisición de nuevas habilidades metabólicas que permiten a los microorganismos explorar nuevos nichos ecológicos.

Uno de los aspectos más interesantes de la evolución procarionte es que la recombinación homóloga puede ser considerada una fuerza que promueve la divergencia ecológica entre poblaciones y mantiene a su vez la cohesión al interior de las poblaciones (Papke *et al.*, 2004; Papke *et al.*, 2007; Cadillo-Quiroz *et al.*, 2012; Shapiro *et al.*, 2012). Desde los primeros trabajos utilizando isoenzimas (Maynard-Smith *et al.*, 1993) hasta los últimos análisis de secuencias multilocus (Vos y Didelot 2009), es claro que existen diferentes tasas de recombinación dependiendo de los linajes analizados. Por ejemplo, *Staphylococcus aureus* es considerada una especie clonal mientras que *Neisseria meningitidis* y *Helicobacteri pylori* son especies cuyas poblaciones son prácticamente panmícticas, es decir, altamente recombinantes (Feil *et al.*, 2003; Jolley *et al.*, 2005; Vos y Didelot 2009). Con base en lo anterior, la dificultad del análisis de la recombinación en microorganismos es que no existe una regla general, es decir, el grado de recombinación así como la prevalencia de los eventos es extremadamente variable. Esta variabilidad parece depender de las características genéticas intrínsecas de cada linaje así como del ambiente en el que se desarrollan.

Tomando en cuenta lo anterior, en el mundo microbiano pueden existir diferentes escenarios dependiendo del grado de recombinación homóloga en los linajes (Fraser *et al.*, 2009). En la figura 6 se representa la tasa en la que dos linajes divergen a lo largo del tiempo (eje y). Considerando esta tasa de cambio en la distancia genética entre dos linajes, los valores negativos indican la convergencia y homogeneización entre los linajes, mientras que los valores positivos indican la divergencia entre dos linajes. Bajo este escenario, los linajes más clonales tenderán a divergir más rápidamente a diferencia de los linajes que recombinan entre si, ya que la recombinación tiende a homogeneizar la poza génica.



**Figura 6**

Divergencia genética entre dos grupos a partir de dos modelos contrastantes (clonal vs recombinante).

Para tasas de recombinación bajas, la población se define como clonal y por lo tanto estos grupos se diferenciarán genéticamente (**Línea y flecha verde**). Conforme aumenta la tasa de recombinación, las fuerzas cohesivas retardarán el proceso de divergencia. Si la tasa de recombinación es mayor al umbral (**línea roja**) las poblaciones tenderán a convergir (**línea y flechas azules**). El punto de especiación estará dado por otros factores que eviten la convergencia genética y permitan la divergencia genética (modificada de Fraser *et al.*, 2009).

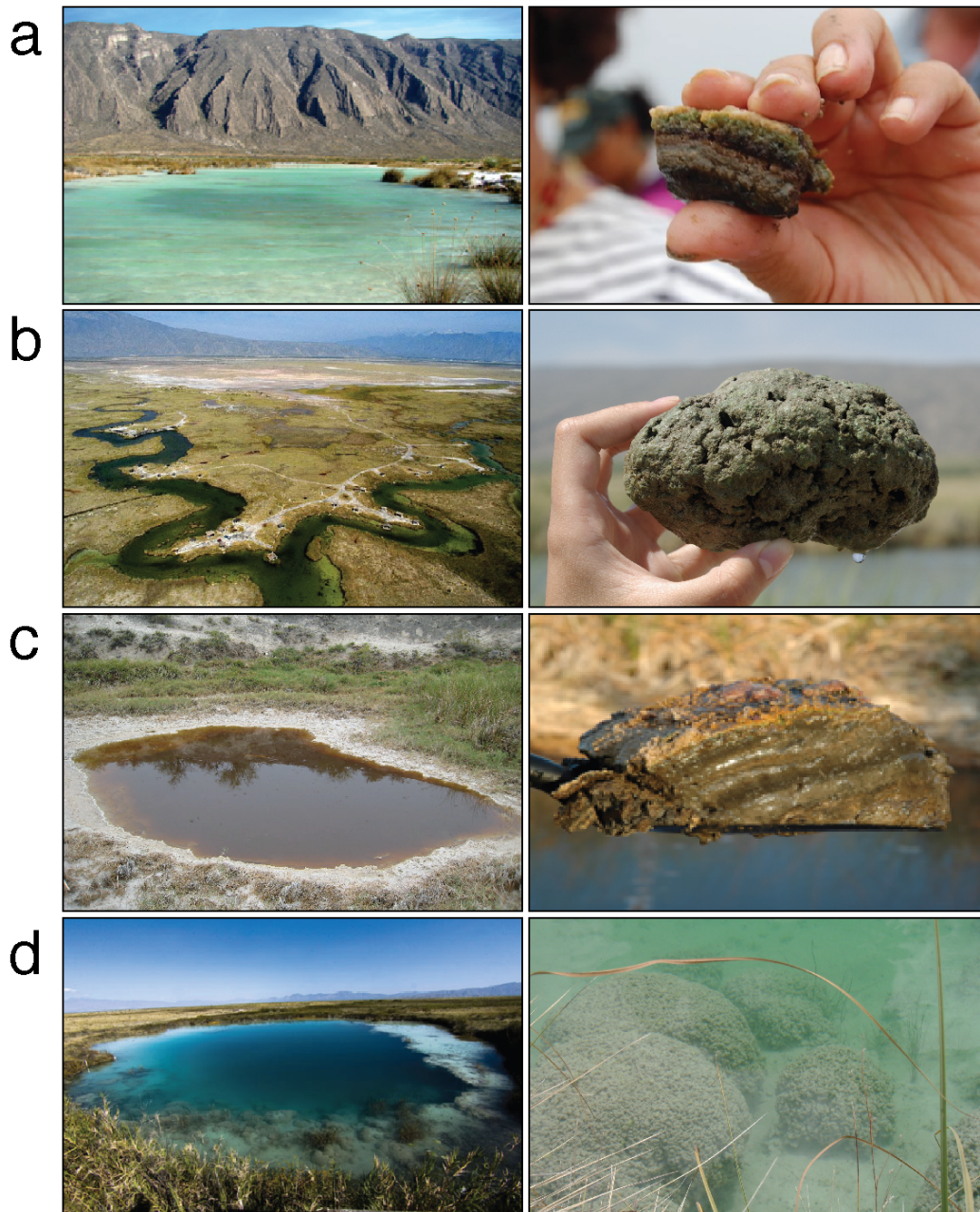
Considerando que la mayor parte de los estudios poblacionales en procariontes han detectado recombinación, el papel de las barreras ecológicas y geográficas es fundamental para entender los procesos que delimitan los eventos de recombinación y por ende promueven la divergencia genética y la especiación. De esta forma podremos explicar los mecanismos que dan lugar a la enorme diversidad de especies procariontes que existen en el planeta.

## 2. ANTECEDENTES

### 2.1 Diversidad y evolución bacteriana en Cuatro Ciénegas Coahuila.

El surgimiento de diversas técnicas moleculares en las últimas décadas ha permitido detectar la gran diversidad de procariontes presentes en todos los ecosistemas. En particular, el valle de Cuatro Ciénegas Coahuila (CCC) constituye uno de los sitios más interesantes para estudiar la ecología y evolución de linajes bacterianos. El valle de CCC es un humedal localizado en el desierto chihuahuense y 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 y recursos contrastantes (Elser *et al.*, 2005). En específico, CCC se caracteriza por tener la concentración más baja de fósforo de los sistemas acuáticos continentales estudiados hasta la fecha ( $[P] < 0.1\mu\text{g/ml}$ ; Garcia-Oliva datos no publicados; Elser *et al.*, 2005). A pesar de esta limitante, las comunidades bacterianas son muy abundantes y poseen una amplia diversidad de especies bacterianas endémicas (Souza *et al.* 2006; Cerritos *et al.*, 2008; Cerritos *et al.*, 2010; Escalante *et al.*, 2008; Escalante *et al.*, 2009).

Los estudios realizados en el valle de CCC, han mostrado que existe una enorme diversidad de comunidades bacterianas. Estas comunidades se encuentran presentes en el agua, pero también se observan comunidades altamente estructuradas como el caso de los tapetes microbianos y los estromatolitos (Figura 7). Incluso, se ha sugerido que las bacterias son el principal productor primario en estos ambientes (Souza *et al.*, 2006)



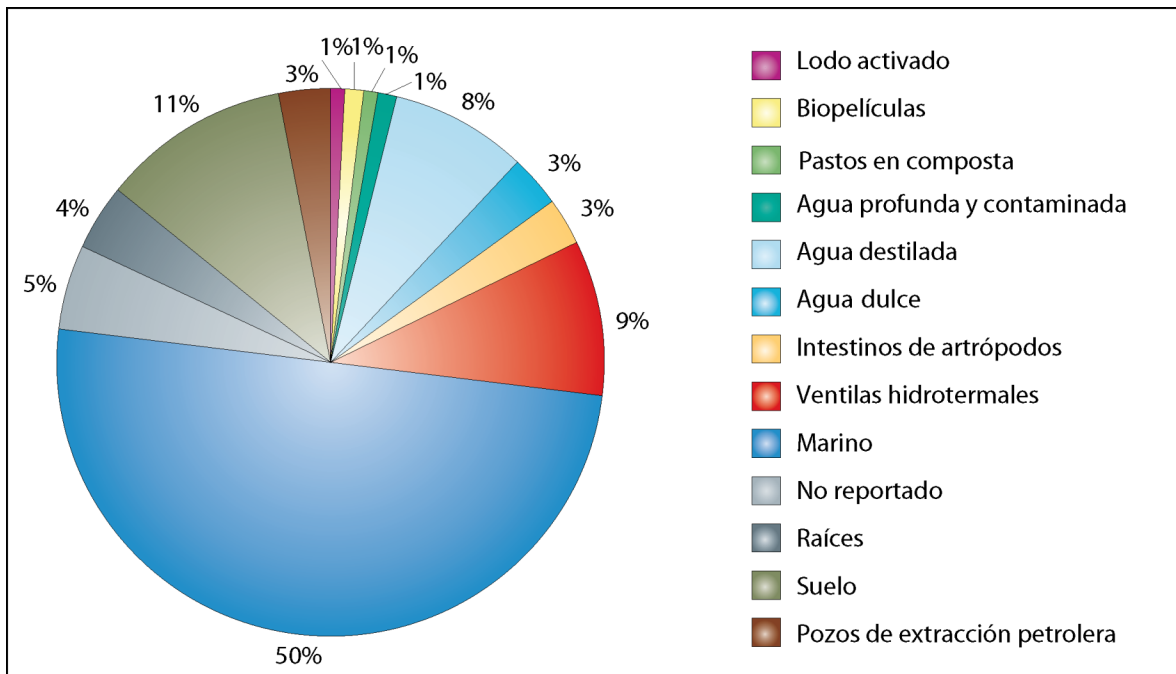
**Figura 7**

Algunas de las comunidades bacterianas presentes en los cuatro sistemas acuáticos principales del valle de CCC. (a) Churince y tapetes microbianos, (b) Rio Mezquites y oncolitos, (c) Los Hundidos y tapetes microbianos. (d) Pozas Azules y estromatolitos. (Fotos de Eria Rebollar y Valeria Souza).

La primera descripción de la estructura y composición de las comunidades bacterianas se obtuvo a partir de un amplio muestreo que abarcó los principales sistemas acuáticos dentro del valle y sus alrededores (Souza *et al.*, 2006). En este trabajo se detectó una enorme diversidad de filotipos bacterianos, los cuales en su



mayoría son cercanos filogenéticamente a bacterias marinas y/u originarias de ventilas hidrotermales (50% y 9% respectivamente). Asimismo los filogrupos asociados a agua dulce continental corresponden a sólo el 3% de la diversidad bacteriana (Figura 8). Este hallazgo permitió sugerir que estas comunidades bacterianas son reliquias del pasado marino de CCC, cuando éste se encontraba todavía bajo el mar hace 150 millones de años en el Jurásico Tardío (Wolaver *et al.*, 2008). Es interesante mencionar, que las comunidades virales presentes en los estromatolitos de CCC, también tienen una composición única y se asocian a linajes marinos (Desnues *et al.*, 2008).



**Figura 8**

Librería de clonas del *16S rRNA* del valle de CCC y zonas aledañas. La gráfica de pastel representa el porcentaje de secuencias con afiliación a diferentes ambientes (Modificada de Souza *et al.*, 2008).

La composición de las comunidades microbianas en CCC es extremadamente diversa pero además representa una composición única asociada a cada sistema acuático (alta diversidad  $\beta$ ; Escalante *et al.*, 2008). Asimismo, se ha encontrado que al interior de un

sistema acuático, las comunidades de bacterias gram positivas (firmicutes y actinobacterias) se distribuyen con respecto a un gradiente de salinidad (Cerritos *et al.*, 2010). Por otro lado, es interesante destacar que la mayor parte de las especies microbianas corresponden al dominio bacteriano y sólo una pequeña proporción al dominio arqueano (Souza *et al.*, 2006 ; Breitbart *et al.*, 2009).

La diversidad bacteriana encontrada en CCC no sólo radica en el número y abundancia de las especies encontradas, sino también en la diversidad de funciones metabólicas presentes en estos ecosistemas microbianos. Por ejemplo, los metagenomas de las comunidades presentes en los microbialitos (oncolitos y estromatolitos) de CCC, han mostrado que estas comunidades son extremadamente complejas en su metabolismo y se encuentran finamente adaptadas a las condiciones ambientales del sitio (Breitbart *et al.*, 2009). En particular, los microbialitos de Rio Mezquites y Pozas Azules, poseen un mayor contenido de genes asociados al transporte de membrana, al metabolismo del fósforo y a la señalización celular (Breitbart *et al.*, 2009).

Además de los estudios realizados en estromatolitos, recientemente se llevó a cabo la caracterización de los tapetes microbianos presentes en dos sistemas acuáticos contrastantes dentro del valle de CCC (Pozas Azules y Los Hundidos). Con base en un enfoque metagenómico, se analizó la estructura y composición de estos dos tapetes microbianos (Bonilla-Rosso *et al.*, en prensa). Asimismo se compararon las vías metabólicas presentes en cada una de las comunidades (Peimbert *et al.*, en prensa). Ambos estudios muestran por un lado la existencia de comunidades microbianas extremadamente distintas en términos de la composición y diversidad de especies. Por otro lado, ambos tapetes poseen una gran diversidad de funciones metabólicas que se asocian a las condiciones únicas de cada sistema acuático.

Además del análisis funcional de las comunidades en CCC, se han estudiado especies particulares que son únicas de CCC, como es el caso de *Bacillus coahuilensis* (Cerritos *et al.*, 2008 y Alcaráz *et al.*, 2008). El análisis detallado del genoma de esta especie ha mostrado increíbles adaptaciones a las condiciones locales. Por ejemplo, esta bacteria posee el genoma más pequeño dentro del género *Bacillus* reportado hasta la fecha. Por otra parte, *B. coahuilensis* tiene la capacidad de sintetizar sulfolípidos de membrana en vez de fosfolípidos. Los rasgos anteriores podrían representar una gran ventaja adaptativa en un ambiente en donde el fósforo es extremadamente bajo. Otra de las características interesantes del genoma de este firmicute, es que presenta genes adquiridos por transferencia horizontal asociados a la adaptación a las condiciones de CCC. En específico, los genes relacionados con la síntesis de sulfolípidos así como genes asociados a la detección de luz y tolerancia a la radiación, se asocian filogenéticamente a cianobacterias marinas (Alcaráz *et al.*, 2008).

Se ha sugerido que debido a las condiciones oligotróficas de CCC, las comunidades bacterianas asimilan DNA ambiental como fuente esencial de fósforo (Souza *et al.*, 2008). Lo anterior conllevaría a tener bajas tasas de recombinación no homóloga en las comunidades bacterianas. Sin embargo, el genoma de *B.coahuilensis* revela que pueden existir eventos de transferencia horizontal asociados a procesos de adaptación local.

Con base en los trabajos mencionados, el valle de CCC es un sitio ideal para el estudio de la evolución y la ecología de las comunidades microbianas silvestres. Por un lado es un sitio único por sus características geológicas y su heterogeneidad ambiental (Elser *et al.*, 2005; Wolaver *et al.*, 2008). Por otro lado, presenta una gran diversidad de comunidades microbianas que permiten contestar preguntas clave sobre la evolución de las poblaciones bacterianas. En específico, los diferentes ambientes acuáticos nos dan la oportunidad de entender cómo la estructura genética en procariontes puede estar determinada por diferentes condiciones ambientales y/o geográficas en las que a largo

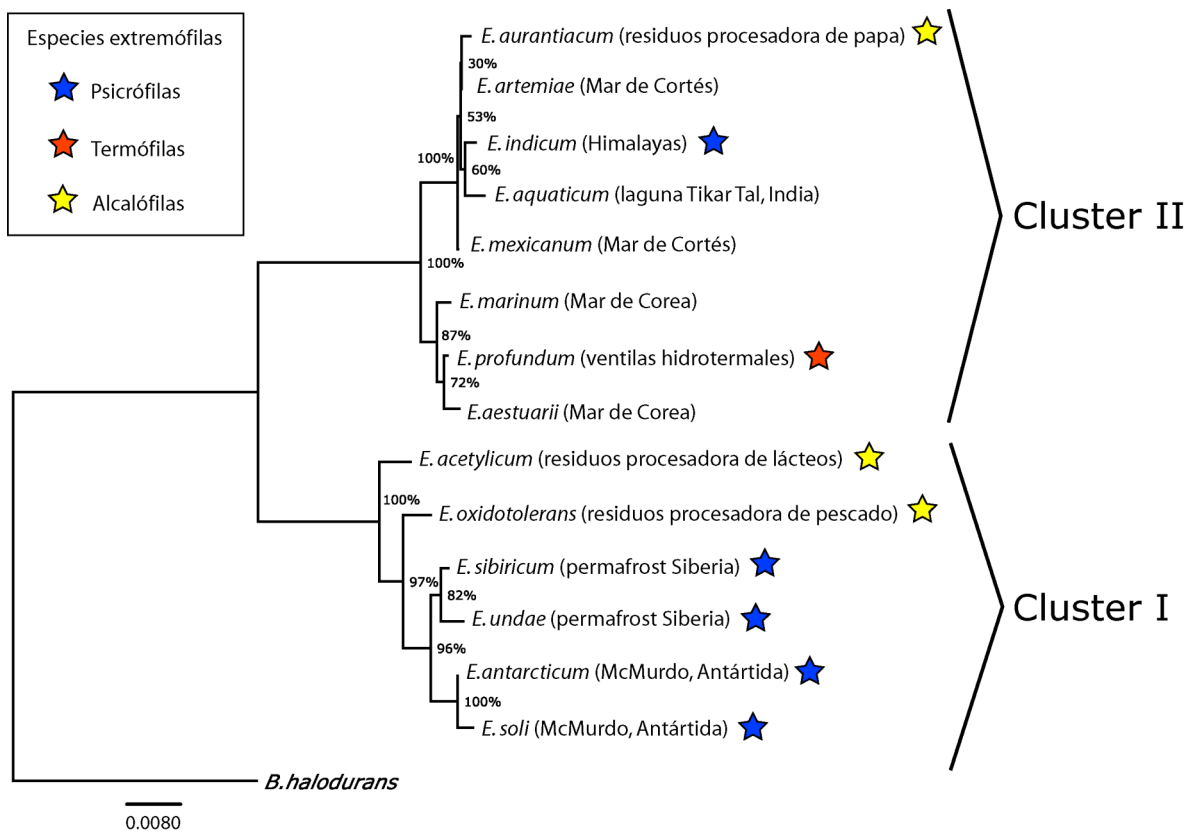
plazo se generan nuevas especies adaptadas a su entorno, generando un fuerte endemismo microbiano.

## 2.2 El género *Exiguobacterium*.

*Exiguobacterium* es uno de los géneros ampliamente distribuidos en el valle de CCC. Previo a este trabajo se habían identificado especies de este género por medio de muestreos de bacterias cultivables así como por medio de librerías de clonas (Souza *et al.*, 2006).

Los organismos del género *Exiguobacterium* son bacterias gram positivas anaerobias facultativas, filogenéticamente cercanos al género *Bacillus* (Farrow *et al.*, 1994). A diferencia del género *Bacillus*, las especies de este género no tienen la capacidad de formar esporas. Como parte de este proyecto se realizó una filogenia con base en la secuencia completa del gen *16srRNA*, utilizando las secuencias de las especies tipo reportadas hasta la fecha (Figura 9). Como se observa en esta figura, existen hasta la fecha 14 especies reportadas dentro del género, de las cuáles la mayoría son organismos provenientes de ambientes extremos. La relación filogenética entre las especies de *Exiguobacterium* reportadas utilizando el *16srRNA* (Figura 9) correlaciona también con el porcentaje de similitud a nivel del genoma completo (Vishnivetskaya *et al.*, 2009).

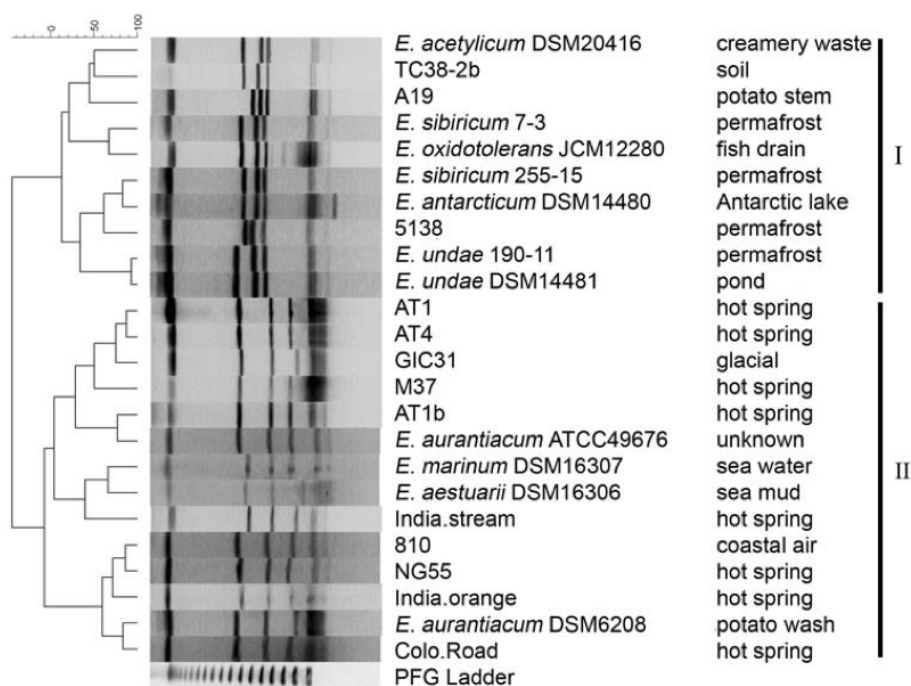
Como se observa en la Figura 9, las 14 especies del género reportadas hasta la fecha se distribuyen en dos grupos principales. El *cluster I* corresponde a organismos psicrófilos así como alcalófilos. Por otro lado, el *cluster II* está compuesto en su mayoría por bacterias marinas.



**Figura 9**

Filogenia NJ del gen *16srRNA* con las 14 especies reportadas de *Exiguobacterium*. El género se divide en dos grupos principales (*cluster I* y *II*).

Las especies y cepas aisladas de este género poseen un amplio rango de temperaturas óptimas de desarrollo, es decir existen especies psicrófilas (Rodrigues *et al.*, 2006), mesófilas (Kim *et al.*, 2005; López Cortés *et al.*, 2006) y termófilas (Crapart *et al.*, 2007). Estudios previos a cerca de la diversidad genotípica y fenotípica del género determinaron que los *clusters I* y *II* son grupos filogenéticamente independientes con un patrón genómico particular. Asimismo, las especies del *cluster I* y *II* se asocian a organismos provenientes de bajas y altas temperaturas respectivamente (Vishnivetskaya *et al.*, 2009; Figura 10).



**Figura 10**

Genotipificación por medio de electroforesis de campo pulsado (PFGE) para un grupo representativo de aislados de *Exiguobacterium*. La parte derecha de la imagen muestra el origen de cada uno de los aislados asociados al cluster I y II (tomada de Vishnivetskaya *et al.*, 2009).

Además de la descripción de las especies del género, existen muy pocos trabajos que exploren la biología de este grupo de bacterias (Rodrigues *et al.*, 2009; Vishnivetskaya *et al.*, 2009). Debido a que este género contiene un gran número de especies extremófilas, es interesante estudiar si existen patrones biogeográficos asociados a la biología de este grupo de organismos. Un trabajo reciente describió la distribución de distintas especies psicrofílicas de *Exiguobacterium* en ambientes fríos, templados y tropicales (Rodrigues *et al.*, 2009). Rodrigues y colaboradores encontraron que la distribución de miembros de este género se encuentra asociada a variables ambientales específicas tales como, salinidad, pH,  $[K^+]$  y  $[Cu^{2+}]$ . Por otra parte, aunque se detectaron especies psicrofílicas en climas templados y tropicales, la mayor diversidad de OTUs (Unidades Taxonómicas Operacionales) se localiza en los ambientes de bajas temperaturas como es la Antártida y Siberia. Este estudio sugiere que la distribución del género *Exiguobacterium* se asocia específicamente a ambientes con baja actividad de

agua ( $a_w$ ), es decir, ambientes en donde la disponibilidad de agua para las funciones biológicas es baja. Los ambientes con baja  $a_w$  son por ejemplo ambientes salinos o con baja temperaturas, en donde el agua es secuestrada por la presencia de iones o no se encuentra disponible en forma líquida debido a las bajas temperaturas (Grant 2004).

Con base en lo anterior, es claro que gran parte de los miembros de este género han sido aislados de ambientes con baja  $a_w$ , tales como el permafrost de Siberia (Rodrigues *et al.*, 2006), ambientes ricos en carbonatos (Vishniveskaya *et al.*, 2009) así como las salinas andinas (Ordoñez *et al.*, 2009).

Hasta la fecha existen dos genomas secuenciados del género, uno de los cuáles es psicrófilo (*E. sibiricum*; <http://genome.ornl.gov/microbial/exig/>) y el otro es termófilo aislado de una de las lagunas de Yellowstone (*E. sp AT1b*; Vishnivetskaya *et al.*, 2011). La comparación de estos genomas con otras especies dentro del género será esencial para entender la adaptación de estos organismos a ambientes de baja actividad de agua, así como al amplio rango de temperaturas en el que se desarrollan.

### 3. JUSTIFICACIÓN DEL PROYECTO

Uno de los principales objetivos de la microbiología evolutiva es entender ¿Porqué existen tantas especies o linajes de procariontes? y ¿Cuáles son los principales procesos involucrados en la divergencia genética y la especiación en estos organismos?. Para contestar a esta y otras preguntas es necesario ampliar nuestro conocimiento acerca de la diversidad genética y distribución de las poblaciones naturales microbianas, ya que hasta la fecha, los microorganismos más estudiados son los patógenos de plantas y animales (Maynard-Smith *et al.*, 2000; Sarkar y Guttman 2004; Hanage *et al.*, 2006b).

Uno de los sitios ideales para estudiar la evolución y ecología de poblaciones microbianas es justamente el valle de CCC. Esto se debe a que es un sitio con una gran cantidad de sistemas acuáticos que poseen una enorme diversidad de ambientes dominados por bacterias. Así mismo, cada uno de estos sistemas posee condiciones ambientales contrastantes entre ellas y únicas con respecto a otros sitios en el planeta. Las bacterias presentes en CCC son endémicas que han sido consideradas reliquias del pasado marino del sitio y poseen adaptaciones únicas a las condiciones particulares de este lugar.

*Exiguobacterium*, es uno de los géneros recurrentemente encontrados en las lagunas de CCC. Como se mencionó en los antecedentes, los miembros de este género tienen características muy interesantes en su biología y distribución geográfica, ya que en su mayoría son especies extremófilas adaptadas a un amplio rango de ambientes extremos. En particular, las especies psicrófilas del género tienen una distribución que correlaciona con parámetros como la salinidad.

En este proyecto se exploró la evolución de poblaciones de *Exiguobacterium* presentes en cuatro de los principales sistemas acuáticos dentro del valle de CCC. El



objetivo principal de este proyecto fue describir la diversidad genética y determinar el papel de las condiciones ambientales y/o geográficas involucradas en la divergencia genética de las poblaciones de *Exiguobacterium*. En particular, es interesante explorar la distribución de los genotipos de este género en ambientes con proporciones iónicas contrastantes a lo largo del valle. Este es el primer trabajo que describe la diversidad y evolución de poblaciones naturales de este grupo de organismos.

Los resultados más importantes de este proyecto se plasmaron en un primer artículo que se presenta en el **capítulo cuatro** de esta tesis. En ese capítulo, se demuestra que las poblaciones de *Exiguobacterium* en CCC, presentan un patrón de diferenciación asociado a dos nichos ecológicos contrastantes. De manera complementaria, la distribución de los linajes de *Exiguobacterium*, así como la genética de poblaciones de estos grupos, se describe en el **capítulo cinco**. Además del estudio evolutivo de las poblaciones de *Exiguobacterium*, consideramos importante profundizar en las características genotípicas y fenotípicas de estos linajes con el fin de describirlos como nuevas especies dentro del género. Esta información se encuentra detallada en el **capítulo seis** de esta tesis.

Esta tesis contribuye de manera importante al entendimiento de los factores ecológicos involucrados en la divergencia genética de las poblaciones bacterianas. En particular, se describe la presencia de una estructura genética al interior de tres linajes de *Exiguobacterium* en el valle de CCC. Esta estructura está asociada a dos nichos ecológicos contrastantes como es el agua y el sedimento. La identificación de subgrupos genéticos al interior de las poblaciones bacterianas apoya la propuesta de que las condiciones ambientales locales son fundamentales en divergencia de linajes bacterianos que promueven eventualmente la especiación. Bajo este escenario, es esencial estudiar las condiciones ambientales locales para explicar la gran diversidad, así como la coexistencia de todas las formas microbianas presentes en el planeta.

**4. ARTÍCULO 1:** “Diferenciación de nicho agua-sedimento en linajes marinos antiguos de *Exiguobacterium* endémicos del valle de Cuatro Ciénegas”.

#### 4.1 RESUMEN

En este trabajo se describió la historia evolutiva y la diferenciación ecológica del género *Exiguobacterium* en el valle de Cuatro Ciénegas Coahuila, México. Los microorganismos del género *Exiguobacterium* son bacterias halófilas y alcalófilas abundantes en los sistemas acuáticos del valle de Cuatro Ciénegas. Se aislaron 183 cepas de *Exiguobacterium* provenientes de los cuatro sistemas acuáticos dentro del valle. De esta colección, se obtuvo la secuencia completa del gen *16srRNA*, así como la secuencia parcial de cuatro genes constitutivos (*citC*, *rpoB*, *recA* y *hsp70*). Con base en la secuencias concatenadas de los genes, se definieron tres grupos filogenéticamente cercanos a especies marinas y termofílicas del género. Estos tres filogrupos no mostraron ninguna asociación con un sistema acuático ni con una salinidad particular. El análisis de la reconstrucción filogenética indicó la presencia de subgrupos de aislados asociados a sedimento o agua. Esta observación fue confirmada por medio de los análisis Unifrac y AdaptML. Lo anterior sugiere un patrón de diferenciación genética asociado a los ambientes de sedimento y agua. Estos resultados muestran que la diferenciación de nicho es uno de los principales procesos involucrados en la estructura de las poblaciones procariontes, llevando así a la divergencia evolutiva entre linajes. Así mismo, este es el primer trabajo que define la diferenciación ecológica en linajes bacterianos asociados a ambientes de sedimento y agua.

# Water–sediment niche differentiation in ancient marine lineages of *Exiguobacterium* endemic to the Cuatro Ciénegas Basin

Eria A. Rebollar,<sup>1</sup> Morena Avitia,<sup>1</sup> Luis E. Eguiarte,<sup>1</sup> Andrea González-González,<sup>1</sup> Lucy Mora,<sup>2</sup> Germán Bonilla-Rosso<sup>1</sup> and Valeria Souza<sup>1\*</sup>

<sup>1</sup>Laboratorio de Evolución Molecular y Experimental, Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City, Mexico.

<sup>2</sup>Instituto de Geología, Universidad Nacional Autónoma de México, Mexico City, Mexico.

## Summary

The evolutionary history and ecological differentiation of the genus *Exiguobacterium* was characterized within natural communities from the Cuatro Ciénegas Basin, Mexico. *Exiguobacterium* comprises both halophilic and alkaliphilic bacteria that are abundant among the aquatic systems of the Cuatro Ciénegas Basin. We obtained complete sequences of the 16S rRNA gene and partial sequences of four housekeeping genes (*citC*, *rpoB*, *recA* and *hsp70*) in 183 *Exiguobacterium* isolates retrieved from distinct aquatic systems. We defined three main phylogroups that are closely related to marine and thermophilic species of the genus. These phylogroups were neither specific to a given aquatic system nor to a particular salinity. Phylogenetic reconstruction indicated the presence of several small clusters within the phylogroups. These clusters consisted of isolates predominantly retrieved from sediment or water. Unifrac and AdaptML analyses confirmed this observation, pointing towards a clear pattern of differentiation linked to either sediment or water habitats. Our results are in line with the concept that niche differentiation is one of the main factors shaping prokaryotic populations and leading to evolutionary divergence.

## Introduction

Prokaryotes are both the most diverse and the most abundant organisms on earth (Horner-Devine *et al.*, 2004).

Received 10 June, 2011; accepted 25 April, 2012. \*For correspondence. E-mail souza@servidor.unam.mx; Tel. (+52) 55 5622 9006; Fax (+52) 55 5622 8995.

Owing to the immense metabolic plasticity of these organisms, they inhabit a vast and contrasting range of ecosystems. Most studies concerning differentiation among prokaryotic lineages refer to pathogenic species (Hanage *et al.*, 2006). However, pathogens represent only a small fraction of microbial diversity and therefore it is difficult to extrapolate these findings to the evolution of all free-living microorganisms (Fraser *et al.*, 2009).

Over the last decade, some studies have shown that genetic differentiation of free-living microbial populations can be promoted by geographical isolation or ecological processes (Hughes *et al.*, 2006; Whitaker, 2006). Geographical isolation at continental scales has been shown to occur among microbial populations that exhibit a wide range of lifestyles (Cho and Tiedje, 2000; Papke *et al.*, 2003; Whitaker *et al.*, 2003; Foti *et al.*, 2006; Vos and Velicer, 2008). This therefore suggests that prokaryotes have a limited capacity for dispersal over large geographical distances.

In addition, local adaptation has been proposed as one of the main evolutionary processes affecting microbial populations. Evidence for this is limited but suggests clear patterns of genetic divergence associated with specific environmental conditions (Sikorski and Nevo, 2005; Hunt *et al.*, 2008; Oakley *et al.*, 2010). Genetic differentiation, caused by local adaptation, would therefore lead to population divergence and the eventual emergence of new prokaryotic species. In line with the distribution patterns of macroorganisms, population structure as a consequence of ecological differentiation could explain some of the observed biogeographical patterns among bacteria, as well as endemisms (Gray *et al.*, 2007; Oakley *et al.*, 2010).

The Cuatro Ciénegas Basin (CCB) is located in the Chihuahuan Desert of north central Mexico. Despite being a desert, this area comprises several aquatic systems that range from ephemeral ponds to permanent lakes, rivers and springs (Minckley, 1969). All these aquatic systems are oligotrophic due to low concentrations of phosphorous and they differ according to specific physico-chemical parameters (Dinger, 2001; Elser *et al.*, 2005). In addition, previous studies have shown that aquatic communities of bacteria and viruses vary among the aquatic systems of the CCB (Desnues *et al.*, 2008; Escalante *et al.*, 2008; Cerritos *et al.*, 2010).

*Exiguobacterium* is one of the genera encountered in different water systems throughout the CCB. The members of this genus are alkaliphilic, non-spore-forming bacteria closely related to the genus *Bacillus* (Farrow *et al.*, 1994). This genus is so far comprised of 14 species that have been isolated from different environments such as industrial wastes (Collins *et al.*, 1983; Farrow *et al.*, 1994), freshwater (Fruhling *et al.*, 2002; Raichand *et al.*, 2011) and diverse marine environments (Kim *et al.*, 2005; Rodrigues *et al.*, 2006; Crapart *et al.*, 2007).

A recent study of *Exiguobacterium* diversity has shown that this genus is divided into two main phylogenetic clusters. *Cluster I* is dominated by temperate and cold-adapted strains whereas *cluster II* is essentially composed of alkaliphilic species, with a marine origin and/or from high-temperature habitats (Vishnivetskaya *et al.*, 2009). Concerning the biogeography of *Exiguobacterium*, the distribution of isolates from this genus has been correlated with physico-chemical factors such as pH and salinity (Rodrigues *et al.*, 2009).

Populations of *Exiguobacterium* from the CCB represent an opportunity to explore the ecological factors that influence the evolution of bacterial populations. Specifically, the heterogeneous nature of these aquatic systems allows the determination of particular environmental conditions involved in the genetic divergence of *Exiguobacterium* populations.

In this study, we characterized the salinity of the four main aquatic systems of the CCB. Based on a multilocus approach, we described three main phylogenetic lineages of *Exiguobacterium* isolated from these four aquatic systems. The genetic structure and the extent of homologous recombination of these lineages were determined. Furthermore, we explored the possibility of genetic clustering associated with particular environmental conditions, thus providing evidence of niche partitioning. Our results suggest that the genetic differentiation of *Exiguobacterium* is shaped by the ecological factors associated with different environments. Specifically, niches linked to sediment and water environments appear to facilitate the genetic divergence between *Exiguobacterium* isolates.

## Results

### *Ionic concentration and Exiguobacterium isolation from the CCB aquatic systems*

To determine the salinity of the four main aquatic systems in the valley, the concentration of eight ions ( $Mg^{2+}$ ,  $K^+$ ,  $Na^+$ ,  $Ca^{2+}$ ,  $SO_4^{2-}$ ,  $Cl^-$ ,  $HCO_3^-$ ,  $CO_3^{2-}$ ) was obtained (Table S1). These aquatic systems called Churince (CH), Rio Mezquites (RM), Los Huididos (LH) and Pozas Azules (PA) are geographically separated and have contrasting hydrological features (Fig. 1; *Experimental procedures*).

In addition to the salinity, the pH of each site was determined. Nevertheless, no significant differences were identified (data not shown) and therefore this variable was discarded from further analyses.

Once the ionic concentration was determined, these variables were reduced by a principal component analysis (Fig. S1). To determine differences in salinity between systems, a linear mixed model was fitted for each principal component (Fig. S2). Based on this analysis, the main differences between systems depend upon the ionic content of the water from RM and LH. These values are very high relative to all other values (Fig. S2a).

In comparison with other saline systems (Oren, 2006), the water from RM and LH can be considered hypersaline with a unique ionic proportion. The rest of the CCB environments are considered saline environments, with an ionic composition similar to seawater (Fig. S3).

In parallel with the measurements of the ionic composition, both the sediment and water were sampled extensively at each site to obtain a total of 183 *Exiguobacterium* isolates (Table 1). These isolates represent 8% of the total colonies obtained during the isolation procedures. Given the distribution of the isolates, it is clear that *Exiguobacterium* is widespread throughout the CCB; however, these bacteria seem to be more abundant in systems like RM and less represented in others such as PA (Fig. 1).

### *Phylogenetic structure and recombination of Exiguobacterium at CCB*

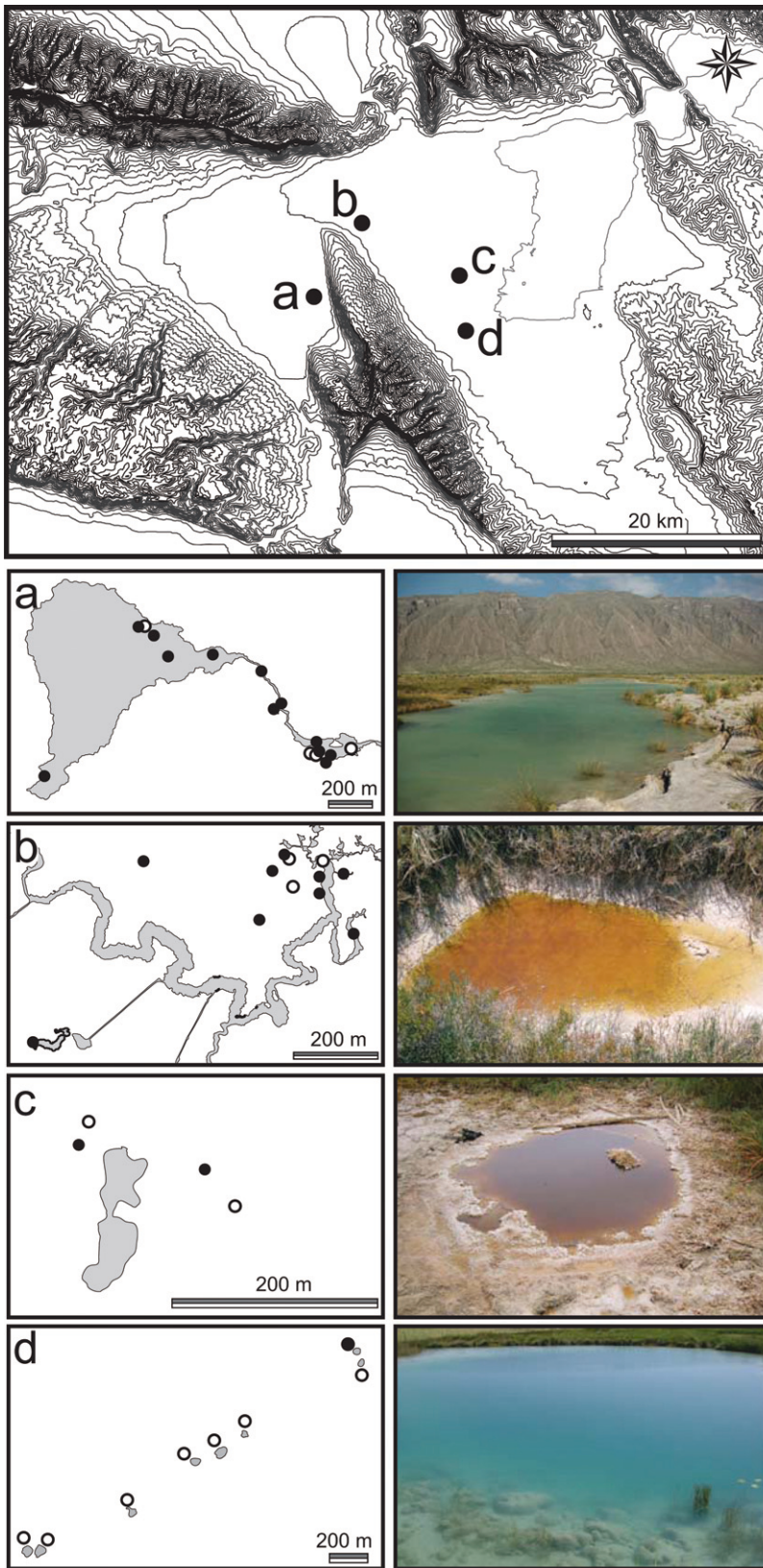
Previous studies indicate that the *Exiguobacterium* genus is divided into two main clusters (Vishnivetskaya *et al.*, 2009). Phylogenetic analyses based on *16srRNA* gene sequences imply that *Exiguobacterium* isolates from the CCB are closely related to *cluster II* species (Fig. 2a). Across the collection of isolates, these sequences share at least 98.5% similarity to each other.

Partial sequences from the four housekeeping genes (*citC*, *hsp70*, *recA* and *rpoB*) were used to reconstruct the phylogenetic history of each gene. The individual gene tree topologies defined the same three phylogroups (data not shown). As a consequence, the partial sequences of the four housekeeping genes were concatenated (Fig. 2b).

Considering the topology of the multilocus tree, *Exiguobacterium* isolates consist of at least three main phylogroups (F1, F2 and F3). These groups represent at least a 6% genetic divergence between clusters. Furthermore, the presence of phylogroups F1, F2 and F3 was confirmed by a STRUCTURE analysis (Pritchard *et al.*, 2000) (Fig. 2c).

In order to determine the genetic diversity of each phylogroup, a population genetic analysis approach was carried out. Nucleotide diversity indices showed contrasting diversity values between the phylogroups. Specifically, the lowest nucleotide diversity values are from the largest





**Fig. 1.** Map of the Cuatro Ciénegas Basin (CCB) with the location of the water systems analysed (top image): (a) Churince (CH), (b) Rio Mezquites (RM), (c) Los Hundidos (LH), (d) Pozas Azules (PA). Black circles show sampling sites in which *Exiguobacterium* isolates were found and empty circles show sites in which *Exiguobacterium* were not found. Next to a, b, c and d are photographs of each system.

**Table 1.** Number of *Exiguobacterium* isolates obtained in this study.

Aquatic system	No. of isolates	Sediment isolates	Water isolates	ESS/TSS
Churince (CH)	42	24	17	13/17
Rio Mezquites (RM)	93	21	73	9/12
Los Hundidos (LH)	35	32	3	2/4
Pozas Azules (PA)	13	13	0	1/8
Total	183	90	93	25/41

Second column shows the number of isolates from each water system and the total number of isolates. Third and fourth columns show the number of isolates from sediment or water. ESS means the number of sampling sites in which *Exiguobacterium* isolates were found and TSS means the total number of sampling sites in which we looked for isolates.

phylogroup F1, which is also reflected by the negative Tajima's *D*-values (Table S2).

To better understand the phylogenetic structure of *Exiguobacterium* at the CCB, we explored the presence of homologous recombination within and between the phylogroups.

First, the pairwise homoplasmy index ( $\Phi_W$ ; Bruen *et al.*, 2006) indicated that there was neither recombination within genes between phylogroups, nor any recombination within F1 and F3. In contrast, in phylogroup F2, *citC* and *recA* genes showed statistically significant evidence for homologous recombination (Table 2).

We additionally tested whether the alleles within the *Exiguobacterium* collection were in linkage equilibrium using the standardized index of association ( $I_A^S$ ; Haubold

**Table 2.** *P*-values obtained from the  $\Phi_W$  (pairwise homoplasmy index) of four housekeeping genes (Huson and Bryant, 2006).

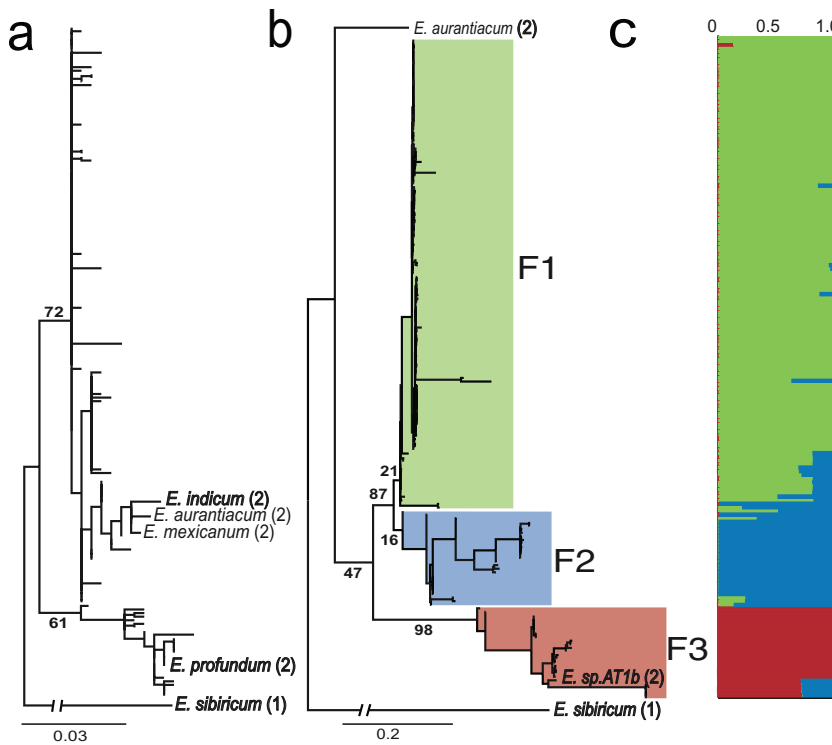
Genetic marker	$\Phi_W$ (pairwise homoplasmy index)			
	Total isolates (183)	F1 isolates (130)	F2 isolates (28)	F3 isolates (25)
<i>citC</i>	0.318	0.317	0.001*	0.346
<i>hsp70</i>	0.380	0.927	0.184	0.290
<i>recA</i>	0.629	1.0	0.028*	0.404
<i>rpoB</i>	0.274	1.0	0.828	0.108

The asterisk (\*) shows significant *P*-values < 0.05.

and Hudson, 2000).  $P_{MC}$ -values obtained for  $I_A^S$  were significant in all cases, so the null hypothesis of linkage equilibrium was rejected (Table 3). These result imply that homologous recombination between alleles has not been sufficient to achieve equilibrium between the alleles.

To further explore the extent of homologous recombination over mutation in the evolutionary history of these phylogroups we performed a ClonalFrame analysis (Didelot and Falush, 2007). From this analysis we obtained the frequency at which recombination and mutation events occur ( $\rho/\theta$ ) and the ratio of probabilities that a given polymorphism is explained by recombination or mutation ( $r/m$ ). This latter statistic is therefore a measure of how important the effect of recombination is relative to mutation in the diversification of the sample.

$\rho/\theta = 0.342$  (0.238 – 0.4775) shows that recombination is approximately three times less frequent than mutation

**Fig. 2.** Phylogenetic relationships of the *Exiguobacterium* isolates from the CCB. Numbers at each node show bootstrap values.

a. ML phylogeny of the complete *16srRNA* gene with previously reported species. The numbers next to each reference species correspond to *cluster I* (1) and *cluster II* (2) of the *Exiguobacterium* genus. F1 and F2 together are at least 99% similar for this gene.

b. ML phylogeny of the concatenated partial housekeeping genes (*citC*, *hsp70*, *recA*, *rpoB*) with reference strains from *cluster I* (1) and *cluster II* (2). F1, F2 and F3 phylogroups are highlighted with different colours.

c. STRUCTURE analysis for three populations ( $K = 3$ ) derived from the four concatenated genes. Each isolate is depicted by a horizontal line in the plot. Each population (F1, F2 and F3) is represented by a different colour. X-axis shows the proportion of membership of the sample in each of the three populations.

**Table 3.** Linkage equilibrium analysis of the multilocus data using LIAN (Haubold and Hudson, 2000).

	$I_A^S$ (standardized index of association)			
	Total isolates (183)	F1 isolates (130)	F2 isolates (28)	F3 isolates (25)
H	0.8493 (0.0326)	0.7261 (0.0626)	0.8873 (0.0123)	0.8585 (0.0088)
VD	0.9758	1.2582	0.9741	1.2649
Ve	0.4992	0.7485	0.3981	0.4851
$I_A^S$	0.3183	0.2270	0.4822	0.5358
$L_{MC}$	0.5173	0.7785	0.4507	0.5489
$P_{MC}$	$< 1.00 \times 10^{-05*}$	$< 1.00 \times 10^{-05*}$	$< 1.00 \times 10^{-05*}$	$< 1.00 \times 10^{-05*}$

The asterisk (\*) shows significant  $P$ -values  $< 0.05$ .

H is mean genetic diversity with its correspondent standard deviation, VD is the observed mismatch variance, Ve is the expected mismatch variance,  $I_A^S$  is the standardized index of association,  $L_{MC}$  is the Monte Carlo simulated 5% critical value and  $P_{MC}$  is the Monte Carlo simulated significance testing the null hypothesis of linkage equilibrium ( $H_0$ :  $VD = Ve$ ).

across the sample. Although recombination seems to be less frequent than mutation,  $r/m = 4.146$  (2.926 – 5.633) shows that more polymorphisms in this population can be explained by recombination rather than mutation.

In addition,  $\rho/\theta$ - and  $r/m$ -values were calculated for each phylogroup.  $\rho/\theta$ -values were 0.29, 0.33 and 0.19 in F1, F2 and F3 respectively. Moreover,  $r/m$ -values were 3.75, 4.16 and 2.52 for F1, F2 and F3 respectively. Considering these statistics, it is clear that each phylogroup is in line with the overall trend from the complete collection. The highest  $r/m$ -value was identified in F2 where the probability of recombination events was linked to *citC* and *recA* genes (data not shown). The latter is consistent with the significant values obtained with the  $\Phi_w$  for these genes in phylogroup F2 (Table 2).

All the previous analyses suggest that although recombination is less frequent than mutation, homologous recombination has played an important role in the divergence between phylogroups along the evolutionary history of *Exiguobacterium* in the CCB.

#### The influence of the environment on the population structure of *Exiguobacterium*

Based on the phylogenetic approach we defined three main phylogroups (F1, F2 and F3; Fig. 2b). These genetic clusters are unevenly distributed along the valley (Fig. 3a). With the exception of LH water, F1 is present in the four aquatic systems. On the other hand, phylogroups F2 and F3 are predominantly from CH and RM. This distribution, however, does not correlate with the salinity of each system. For example, all phylogroups contain isolates from high-salinity environments (water from RM or LH). In addition, sediment isolates from PA and LH are almost exclusively from F1, despite the clear differences in carbonates and calcium concentration between these two sites (Fig. S2b and c).

Based on the previous observations, the distribution of the isolates from each phylogroup is non-random;

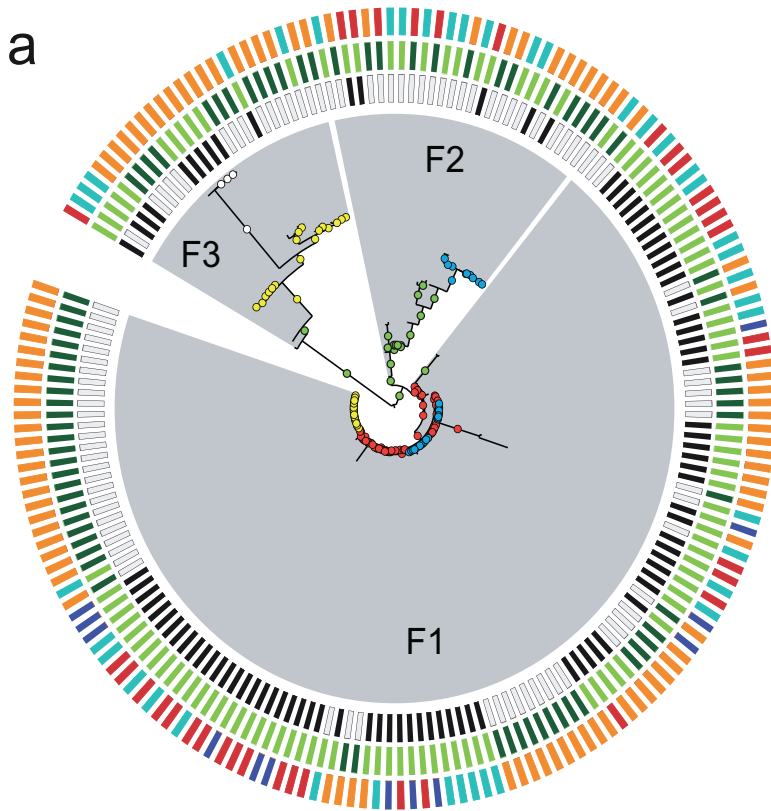
however, this pattern is not determined by salinity or by the distance between systems, as confirmed by a Mantel test (data not shown).

In addition to the distribution of the isolates within the valley, the multilocus tree revealed subclusters within each phylogroup. Most of these observed subclusters were composed of isolates from sediment or water only (Fig. 3a). In contrast, subclusters do not show a specific association with a particular aquatic system (data not shown).

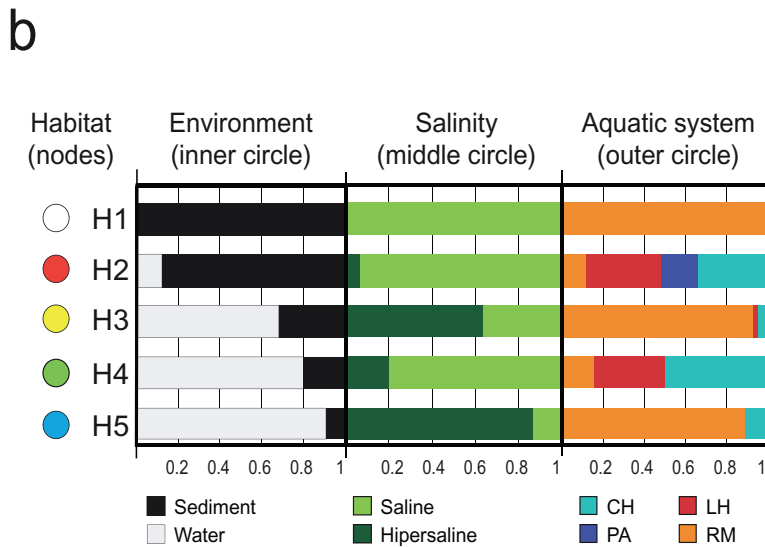
To assess the probability of a significant relationship between the genetic clusters and their corresponding origin, a Unifrac analysis was performed (Lozupone *et al.*, 2006). Both the weighted Unifrac significance test (Lozupone and Knight, 2005) and the parsimony  $P$ -test (Martin, 2002) indicated significant  $P$ -values with respect to either sediment or water partitioning ( $P = 0.01$  and  $P < 0.01$  respectively). Moreover, a significant relationship between saline and hypersaline environments was determined only from the  $P$ -test ( $P < 0.01$ ). Nevertheless, there was no relationship with respect to the aquatic systems or year of isolation ( $P > 0.05$ ).

The above results suggest the existence of ecologically differentiated clusters within populations of *Exiguobacterium*. This differentiation is associated with either sediment or water environments. Furthermore, salinity is also an important factor for the observed ecological divergence. In order to evaluate if this is an ongoing process of niche partitioning, we performed an AdaptML analysis (Hunt *et al.*, 2008). This program has been successfully used to define ecologically differentiated populations of free-living marine *Vibrio* strains and other bacterial species (Hunt *et al.*, 2008; Oakley *et al.*, 2010). The 'projected habitats' obtained by this approach suggest the existence of genetic clusters associated with a combination of environmental conditions. The organisms from each projected habitat can be considered ecotypes within a given population.

We obtained five 'projected habitats' that are represented by various colours at the nodes of the phylogeny



**Fig. 3.** AdaptML analysis of *Exiguobacterium* populations from the CCB.  
 a. Concatenated ML phylogeny showing the origin of each of the isolates: sediment or water (inner circle), salinity (middle circle) and aquatic system (outer circle). The coloured circles at the nodes of the phylogeny show the five habitats inferred by the program.  
 b. Proportion of the isolates in each of the five inferred habitats according to: environment, salinity and aquatic system.



(Fig. 3a). In some cases, habitats (such as H4 and H5) were present in different phylogroups. However, some of these habitats were predicted at the base of the phylogeny (H4) and others are present at the most recent branches of the tree (H3 and H5). Considering the origin of the isolates, it is clear that the proposed habitats are not restricted to a specific aquatic system (Fig. 3b).

On the other hand, H1 and H2 are dominated by isolates from saline sediments. Given the low number of

isolates from H1, any further extrapolation should be viewed with caution. Habitat H3 is an intermediate environment in which there are isolates from sediment, water, saline and hypersaline environments. In addition, isolates from H4 and H5 are predominantly from water environments. These aquatic habitats are differentially composed by isolates living in two contrasting salinities. Overall, the five proposed habitats are clearly structured according to a gradient of sediment to water partitioning.



## Discussion

In this study we characterized three genetically differentiated populations of *Exiguobacterium* within the CCB. These populations are closely related to the previously described *cluster II* in which most of the reported species have a marine and/or thermotolerant origin (Vishnivetskaya *et al.*, 2009).

This finding is in good agreement with previous diversity analyses within the CCB that describe a large number of phylotypes affiliated with marine or hydrothermal vent bacterial groups (Souza *et al.*, 2006). The aquatic communities of the CCB represent relics of marine bacterial lineages from the ancient sea that covered Coahuila at least 150 million years ago (Souza *et al.*, 2006). The results obtained in this study show that salinity of the CCB systems remain close to marine ionic proportions. In agreement with this finding, *Exiguobacterium* populations can be considered part of this marine ancestry.

Based on the concatenated phylogeny and the population genetic estimates, F1, F2 and F3 display contrasting diversity levels. Specifically, F1 appears to be in the process of a demographic expansion in which there is a few dominant haplotypes within the population. Furthermore, F1 is actually the most widespread phylogroup in the valley. This suggests that this phylogroup is adapted to a wide range of local conditions within the CCB. Furthermore, F1 is not closely related to any reported species within this genus, and therefore this recent lineage could be considered an endemic population from the CCB.

Additional analyses indicate that recombination within genes was detected only in F2. Furthermore, the standardized index of association revealed that alleles are in linkage disequilibrium. However, ClonalFrame analysis shows that although infrequent, recombination events are relevant in the context of both the population structure and differentiation between lineages of *Exiguobacterium*.

Nonetheless, other evolutionary forces such as genetic drift and natural selection may also play a role in the genetic divergence of these lineages. However, the relative importance of either of these evolutionary forces would require the use of less conserved markers than those used in this study.

Concerning the distribution of the phylogroups along the valley, it is evident that there is no clear relationship between their distribution and a particular aquatic system, salinity or geographic distance between sites. As shown in other bacterial lineages (Cho and Tiedje, 2000; Vos and Velicer, 2008), the absence of an isolation by distance pattern shows that *Exiguobacterium* isolates do not encounter dispersal limits at local scales.

A previous study addressing the biogeography of *Exiguobacterium* found that the distribution of this genus was correlated to physico-chemical parameters such as salin-

ity and pH (Rodrigues *et al.*, 2009). Rodrigues and colleagues proposed that this pattern could be explained by the capacity of the genus to successfully grow within low water activity ( $a_w$ ) environments, such as Siberian permafrost or marine sediments (Grant, 2004).

As shown in this study, the aquatic systems of the CCB can be classified as saline or hypersaline environments (with low  $a_w$ -values). The prevalence of *Exiguobacterium* isolates within sediment and water from the CCB can be explained by their adaptation to salinity, as one of the main physiological characteristics of this genus (Rodrigues *et al.*, 2006; Crapart *et al.*, 2007).

As mentioned previously, the evolutionary processes promoting the emergence of new bacterial species are not well understood. It has been shown that bacteria tend to form genetic clusters that could arise by niche differentiation (Fraser *et al.*, 2009). This latter process has been proposed as part of the ecotype model in which an adaptive mutant is able to exploit a new ecological niche (Cohan and Koeppel, 2008).

In order to detect ecological differentiation events within a population, it is essential to detect informative environmental parameters or niches (Polz *et al.*, 2006). Based on the concatenated sequence phylogeny, we have been able to detect genetic clusters that are associated with sediment or water habitats, thus representing a fine-scale differentiation process associated with niche partitioning. Overall, our results are in agreement with the ecotype model, in which genetic clusters differentiate through a niche-partitioning process.

The genetic differentiation between water and sediment isolates has been confirmed by Unifrac and parsimony *P*-tests. Furthermore, AdaptML results suggest the presence of habitats or ecotypes determined by a sediment–water gradient. On the other hand, isolates from water-dominated habitats are differentiated into saline and hypersaline environments. This latter observation suggests the existence of two different ecotypes from water-dominated habitats.

The presence of ‘projected habitats’ within more than one phylogenetic cluster suggests that the ecological differentiation process is occurring simultaneously in different phylogroups. An alternative interpretation of this result is that the ecotypes present in more than one group constitute the ancestral habitat. However, the ecotypes projected by the AdaptML suggest the presence of an ancestral habitat (H4) that is present at the base of the phylogeny and that consists of isolates from water with a similar salinity to seawater. Furthermore, habitats such as H3 and H5 are present in more than one phylogroup and consist of isolates from the most recent branches of the tree (Fig. 3).

Recent studies concerning the evolution of bacterial populations have demonstrated niche differentiation

events in marine, sediment or soil bacteria (Sikorski and Nevo, 2005; Hunt *et al.*, 2008; Oakley *et al.*, 2010). This work supports the previous observations in other bacterial lineages and represents the first study to show ecological divergence associated with water and sediment habitats at the population level.

What are the main factors that could therefore promote ecological divergence in sediment and water environments? A water body represents a less structured environment with a homogenous distribution of resources and low cell density. In contrast, sediment or microbial mats are highly structured environments with heterogeneous distribution of resources and higher cell density. High cell density environments promote a series of competition and/or cooperation interactions that confer different challenges to coexisting species and the overall community.

To define the particular differences between the sediments and water bodies of the CCB, it would be necessary to measure several environmental and biological parameters. Specifically, nutrient and oxygen concentrations, in addition to cell density, would provide critical data to support the differentiation of these two habitats. Moreover, these measurements would facilitate the identification of specific parameters involved in the ecological differentiation within *Exiguobacterium* populations. The sediment and water characterization of the CCB water bodies is currently being addressed in our laboratory.

The association of these genetic clusters to specific habitats with contrasting ecological conditions strongly suggests that we are witnessing a local adaptation process. However, further comparative genomic analysis of sediment and water isolates would allow us to detect specific signatures that could account for this ecological differentiation process.

## Experimental procedures

### *Study site description and sample collection*

We analysed four aquatic systems within the CCB: CH, RM, LH and PA. These aquatic systems are geographically separated and have different hydrological features described in previous works (Minckley, 1969). The RM and LH ponds are the most fluctuating environments; both consist of small desiccation ponds in which the water volume reduces dramatically during summer and increases during winter. The RM ponds are water bodies fed by the Mezquites River. The LH ponds are instead fed by a subterranean river that emerges from several sinkholes. Site CH consists of two evaporating lagoons interconnected by a small river and fed a permanent water spring. Finally, site PA consists of several permanent deep ponds that do not experience fluctuation in terms of water volume.

To retrieve *Exiguobacterium* strains at the CCB, surface water and sediment samples were obtained from the four water systems along three summer seasons (2007, 2008 and

2009). Sediment samples correspond to the first 20 cm beneath the water column. Considering the characteristics of the aquatic systems, an equivalent sampling effort was performed within each system every year. In PA we collected samples from the eight main ponds; LH consists of four small desiccation ponds so we collected samples in all of them. In RM we collected all desiccation ponds that surrounded the river and finally in CH we obtained samples from all sites throughout the system (considering the large lagoon, the connecting river and the intermediate lagoon). Despite the equivalent sampling effort, *Exiguobacterium* isolates were more readily cultivated from some aquatic systems such as RM and CH, and very difficult to retrieve from other systems such as PA. Specifically, isolates were obtained in 25 out of the 41 sites in which samples were collected.

Water and sediment samples were diluted using saline solution 1% NaCl at different concentrations (1:10, 1:100, 1:1000 and 1:10000). Two hundred microlitres of each dilution were incubated in 10 marine agar media (Difco 2216) Petri dishes at 37°C. Considering all dilutions, we incubated 80 Petri dishes per site (40 with water samples and 40 with sediment samples). After 24 h incubation, *Exiguobacterium* strains grew on the Petri dishes as bright orange colonies. *Exiguobacterium* strains represent 8% of the total colonies and 40% of all the orange colonies. All orange colonies were purified by single-colony isolation on marine agar to obtain a single-clone plate. Finally, *Exiguobacterium* isolates were stored at -80°C in marine agar media with 20% (w/vol) glycerol. This same procedure was performed along 3 years; however, final isolation numbers were different in each system. For example, in RM we obtained water and sediment isolates in the 3 years of collection; however, in PA we only obtained 13 sediment isolates in 2008.

### *Ionic concentration and statistical analyses, molecular methods and genetic analyses*

See Appendix S1 in *Supporting information*.

### *Molecular markers*

The *16srRNA* gene was PCR amplified using primers 27F and 1492R (Lane, 1991). In addition, four housekeeping genes (*citC*, *recA*, *hsp70* and *rpoB*) were amplified using previously reported primers specifically designed for *Exiguobacterium* strains (Rodrigues *et al.*, 2006). These four genes are evenly distributed throughout the two available sequenced genomes of the genus (*Exiguobacterium sibiricum* and *Exiguobacterium AT1b*). At the time we conducted this study only *E. sibiricum* genome was available. We also designed additional primers for *csp*, *gltX* and *gyrB* according to the *E. sibiricum* genome and previous reports (Rodrigues *et al.*, 2006); however, we could not obtain specific products for these genes so we did not include them for the final analyses.

### *Phylogenetic analysis*

*16srRNA* sequences from *Exiguobacterium* strains were aligned and edited using Bioedit (Hall, 1999). Reported

16srRNA sequences from five *Exiguobacterium* species were included as references and were obtained from the GenBank database: *E. aurantiacum* (DQ019166), *E. mexicanum* (AM072764), *E. indicum* (DQ497799), *E. profundum* (AY818050) and *E. sibiricum* (CP001022). A Jmodel Test analysis was used to obtain a specific substitution model for this gene: TIM1+I+G (Posada, 2008). The sequence alignment of 1302 bp was used to obtain a phylogenetic reconstruction by a maximum likelihood (ML) approach using PhyML v.3.0 (Guindon and Gascuel, 2003). Bootstrapping was performed with 1000 iterations.

The partial *citC*, *hsp70*, *recA* and *rpoB* gene sequences yielded alignments of 663, 657, 604, 611 nucleotides respectively. Jmodel Test was used to obtain a specific DNA substitution model per gene and for the concatenated alignment: *citC* (K80+G), *hsp70* (TPM2+G), *recA* (K80+G), *rpoB* (K80+I) and concatenated genes (TIM2+I+G). Phylogenetic relationships of individual genes and concatenated genes were reconstructed using PhyML v.3.0 (Guindon and Gascuel, 2003) with the corresponding DNA substitution models and 1000 bootstrap iterations. Three reference *Exiguobacterium* strains obtained from the GenBank database were included in the concatenated sequence phylogenetic reconstructions: *E. aurantiacum*, *E. sibiricum* and *Exiguobacterium* sp. AT1b.

The Bayesian clustering program STRUCTURE v2.2 (Pritchard *et al.*, 2000) was used to infer the population structure without a predefined population subdivision. For the selected  $K = 3$  value, the individual membership coefficient ( $q_{ind}$ ) was evaluated and values were plotted.

### Recombination analyses

To detect recombination within housekeeping genes the  $\Phi_w$  test (pairwise homoplasy index) was performed using SplitsTree program (Huson and Bryant, 2006). To evaluate the extent of linkage equilibrium we obtained the  $I_A^S$  (standardized index of association) using LIAN 3.0 software (Haubold and Hudson, 2000). ClonalFrame analysis (v1.2) was obtained out of 400 000 iterations with a burn-in of 400 000. Results were analysed with the correspondent graphical interface called GUI (Didelot and Falush, 2007). Based on the output ClonalFrame file,  $\rho/\theta$  and  $r/m$  were calculated for the complete collection and for each phylogroup.  $\rho/\theta$  is the ratio of rates at which recombination and mutation occur and  $r/m$  is the ratio of probabilities that a nucleotide is changed as the result of recombination relative to point mutation.

### Environmental correlation analyses

To determine whether phylogenetic clustering is driven by specific environmental parameters, we used the Unifrac online tool (Lozupone *et al.*, 2006). This method uses an input tree with branch lengths and the environmental characteristics of each isolate. In this case the ML tree of the concatenated genes was used. Different environmental parameters were evaluated independently: aquatic system, sediment or water origin, salinity and the year of isolation. Unifrac metric (Lozupone and Knight, 2005) and Parsimony  $P$ -tests (Martin, 2002) were calculated based on the previous specifications.

To detect putative ecotypes based on the phylogenetic clustering of the concatenated sequences we used the AdaptML online server (Hunt *et al.*, 2008). The analysis was performed using a concatenated ML phylogeny of the *Exiguobacterium* collection according to its corresponding origin: aquatic system, environment, salinity and isolation year. The model was run with default parameters and 100 randomized topologies. AdaptML circular tree figures were drawn using the online iTOL v2.0 (Letunic and Bork, 2011). The results shown in Fig. 3 do not include isolation year because the time component did not influence the projected habitats retrieved by the program.

### Acknowledgements

We thank Dr Rebeca Aguirre for the statistical advisory along the project. We thank Dr Daniel Piñero, Santiago Ramírez-Barahona and Dr Christine Rooks for their constructive reviews of the manuscript. We also thank Laura Espinosa and Dr Erika Aguirre for technical support during the development of the project. We finally thank Dr Xavier Didelot for his advice in the calculation of  $r/m$  per lineage. This paper is part of the doctoral research of the first author, who thanks the Doctorado en Ciencias Biomédicas (Universidad Nacional Autónoma de México) and CONACYT (Grant No. 170206) for financial support. The project was supported by Grants 0023459-SEMARNAT/CONACYT, 57507 SEP CONACyT and WWF-Alianza Carlos Slim and by grants of DGAPA and UC-Mexus during sabbatical of Dr Valeria Souza and Dr Luis E. Eguiarte.

### References

- Bruen, T.C., Philippe, H., and Bryant, D. (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics* **172**: 2665–2681.
- Cerritos, R., Eguiarte, L.E., Avitia, M., Siefert, J., Travisano, M., Rodríguez-Verdugo, A., and Souza, V. (2010) Diversity of culturable thermo-resistant aquatic bacteria along an environmental gradient in Cuatro Ciénegas, Coahuila, México. *Antonie Van Leeuwenhoek* **99**: 1–16.
- Cho, J., and Tiedje, J.M. (2000) Biogeography and degree of endemism of fluorescent *Pseudomonas* strains in soil. *Appl Environ Microbiol* **66**: 5448–5456.
- Cohan, F.M., and Koeppel, A.F. (2008) The origins of ecological diversity in prokaryotes. *Curr Biol* **18**: R1024–R1034.
- Collins, M.D., Lund, B.M., Farrow, J.A.E., and Schleifer, K.H. (1983) Chemotaxonomic study of an alkalophilic bacterium, *Exiguobacterium aurantiacum* gen. Nov., sp. nov. *J Gen Microbiol* **129**: 2037–2042.
- Crapart, S., Fardeau, M.L., Cayol, J.L., Thomas, P., Sery, C., Ollivier, B., and Combet-Blanc, Y. (2007) *Exiguobacterium profundum* sp. nov., a moderately thermophilic, lactic acid-producing bacterium isolated from a deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* **57**: 287–292.
- Desnues, C., Rodríguez-Brito, B., Rayhawk, S., Kelley, S., Tran, T., Haynes, M., *et al.* (2008) Biodiversity and biogeography of phages in modern stromatolites and thrombolites. *Nature* **452**: 340–343.



- Didelot, X., and Falush, D. (2007) Inference of bacterial microevolution using multilocus sequence data. *Genetics* **175**: 1251–1266.
- Dinger, E. (2001) Aquatic invertebrates of Cuatro Ciénegas, Coahuila, Mexico and effects of fish on stromatolite invertebrate assemblages. MS Thesis. Flagstaff, AZ, USA: Northern Arizona University.
- Elser, J., Schampel, J.H., Garcia-Pichel, F., Wade, B.D., Souza, V., Eguiarte, L.E., et al. (2005) Effects of phosphorus enrichment and grazing snails on modern stromatolitic microbial communities. *Freshw Biol* **50**: 1808–1825.
- Escalante, A.E., Eguiarte, L.E., Espinosa-Asuar, L., Forney, L.J., Noguez, A.M., and Souza, V. (2008) Diversity of aquatic prokaryotic communities in the Cuatro Ciénegas Basin. *FEMS Microbiol Ecol* **65**: 50–60.
- Farrow, J.A., Wallbanks, E.S., and Collins, M.D. (1994) Phylogenetic interrelationships of round spore-forming bacilli containing cell walls based on lysine and the non-spore-forming genera *Caryophanon*, *Exiguobacterium*, *Kurthia* and *Planococcus*. *Int J Syst Bacteriol* **44**: 74–82.
- Foti, M., Ma, S., Sorokin, D.Y., Rademaker, J.L.W., Kuenen, J.G., and Muyzer, G. (2006) Genetic diversity and biogeography of haloalkaliphilic sulphuroxidizing bacteria belonging to the genus *Thioalkalivibrio*. *FEMS Microbiol Ecol* **56**: 95–101.
- Fraser, C., Alm, E.J., Polz, M.F., Spratt, B.G., and Hanage, W.P. (2009) The bacterial species challenge: making sense of genetic and ecological diversity. *Science* **323**: 741–746.
- Fruhling, A., Schumann, P., Hippe, H., Straubler, B., and Stackebrandt, E. (2002) *Exiguobacterium undae* sp. nov. and *Exiguobacterium antarcticum* sp. nov. *Int J Syst Evol Microbiol* **52**: 1171–1176.
- Grant, W.D. (2004) Life at low water activity. *Philos Trans R Soc Lond B Biol Sci* **359**: 1249–1267.
- Gray, N.D., Brown, A., Nelson, D.R., Pickup, R.W., Rowan, A.K., and Head, I.M. (2007) The biogeographical distribution of closely related freshwater sediment bacteria is determined by environmental selection. *ISME J* **1**: 596–605.
- Guindon, S., and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**: 696–704.
- Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**: 95–98.
- Hanage, W.P., Fraser, C., and Spratt, B.G. (2006) Sequences, sequence clusters and bacterial species. *Philos Trans R Soc Lond B Biol Sci* **361**: 1917–1927.
- Haubold, B., and Hudson, R. (2000) LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Bioinformatics* **16**: 847–848.
- Horner-Devine, M.C., Carney, K.M., and Bohannon, B.J.M. (2004) An ecological perspective on bacterial biodiversity. *Proc Biol Sci* **271**: 113–122.
- Hughes, J.B., Bohannon, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., et al. (2006) Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4**: 102–112.
- Hunt, D., David, L.A., Gevers, D., Preheim, S.P., Alm, E.J., and Polz, M.F. (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* **320**: 1081–1085.
- Huson, D.H., and Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**: 254–267.
- Kim, I., Lee, M., Jung, S., Song, J.J., Oh, T., and Yoon, J. (2005) *Exiguobacterium aestuarii* sp. nov. and *Exiguobacterium marinum* sp. nov. isolated from tidal flat of the Yellow Sea in Korea. *Int J Syst Evol Microbiol* **55**: 885–889.
- Lane, D.J. (1991) 16S/23S rRNA sequencing. In *Nucleic Acid Techniques in Bacterial Systematics*. Stackebrandt, E., and Goodfellow, M. (eds). Chichester, UK: Wiley, pp. 115–175.
- Letunic, I., and Bork, P. (2011) Interactive Tree of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* **39**: W475–W478.
- Lozupone, C., and Knight, R. (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71**: 8228–8235.
- Lozupone, C., Hamady, M., and Knight, R. (2006) UniFrac – an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* **7**: 371.
- Martin, A.P. (2002) Phylogenetic approaches for describing and comparing the diversity of microbial communities. *Appl Environ Microbiol* **68**: 3673–3682.
- Minckley, W.L. (1969) Environments of the Bolson of Cuatro Ciénegas, Coahuila, Mexico, with special reference to the aquatic biota. University of Texas, El Paso. Science Series 2: 1–65.
- Oakley, B.B., Carbonero, F., van der Gast, C.J., Hawkins, R.J., and Purdy, K.J. (2010) Evolutionary divergence and biogeography of sympatric niche-differentiated bacterial populations. *ISME J* **4**: 448–497.
- Oren, A. (2006) Chapter 1.9: Life at high salt concentrations. In *Prokaryotes: A Handbook on the Biology of Bacteria*, 3rd edn., Vol. 2. Dworkin, M. (ed.). Minneapolis, MN, USA: Springer, pp. 263–282.
- Papke, R.T., Ramsing, N.B., Bateson, M.M., and Ward, D.M. (2003) Geographical isolation in hot spring cyanobacteria. *Environ Microbiol* **5**: 650–659.
- Polz, M.F., Hunt, D.E., Preheim, S.P., and Weinreich, D.W. (2006) Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. *Philos Trans R Soc Lond B Biol Sci* **361**: 2009–2021.
- Posada, D. (2008) jModelTest: phylogenetic model averaging. *Mol Biol Evol* **25**: 1253–1256.
- Pritchard, J.K., Stephens, M., and Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Raichand, R., Pareek, S., Singh, N.K., and Mayilraj, S. (2011) *Exiguobacterium aquaticum* sp. nov., a new member of the genus *Exiguobacterium*. *Int J Syst Evol Microbiol*. doi:10.1099/ijs.0.035790-0. Epub ahead of print.
- Rodrigues, D.F., Goris, J., Vishnivetskaya, T., Gilichinsky, D., Thomashow, M.F., and Tiedje, J.M. (2006) Characterization of *Exiguobacterium* isolates from the Siberian permafrost. Description of *Exiguobacterium sibiricum* sp. nov. *Extremophiles* **10**: 285–294.
- Rodrigues, D.F., Jesus, E., Ayala-Del-Río, H., Pellizari, V., Gilichinsky, D., Sepulveda-Torres, L., and Tiedje, J.M. (2009) Biogeography of two cold-adapted genera: *Psychrobacter* and *Exiguobacterium*. *ISME J* **3**: 658–665.
- Sikorski, J., and Nevo, E. (2005) Adaptation and incipient sympatric speciation of *Bacillus simplex* under microcli-

- matic contrast at 'Evolution Canyons' I and II, Israel. *Proc Natl Acad Sci USA* **102**: 15924–15929.
- Souza, V., Espinosa-Asuar, L., Escalante, A.E., Eguiarte, L.E., Farmer, J., Forney, L., *et al.* (2006) An endangered oasis of aquatic microbial biodiversity in the Chihuahuan desert. *Proc Natl Acad Sci USA* **103**: 6565–6570.
- Vishnivetskaya, T.A., Kathariou, S., and Tiedje, J.M. (2009) The *Exiguobacterium* genus: biodiversity and biogeography. *Extremophiles* **13**: 541–555.
- Vos, M., and Velicer, G.J. (2008) Isolation by distance in the spore-forming soil bacterium *Myxococcus xanthus*. *Curr Biol* **18**: 386–391.
- Whitaker, R.J. (2006) Allopatric origins of microbial species. *Philos Trans R Soc Lond B Biol Sci* **361**: 1975–1984.
- Whitaker, R.J., Grogan, D.W., and Taylor, J.W. (2003) Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* **301**: 976–978.

### Supporting information

Additional Supporting Information may be found in the online version of this article:

**Fig. S1.** Principal component analysis of the ionic concentration in the CCB aquatic systems.

a. Eigenvalues and percentage of the variance explained by each of the principal components (PCs). Components in red are the ones with the highest eigenvalues and contribution of variance percentage.

b. Plot showing the coefficient values for the three PCs.

**Fig. S2.** Estimated means obtained by linear mixed models of the three principal components (PCs). X-axis shows the aquatic systems analysed (CH, RM, LH, PA). Y-axes show the estimated means of each principal component. Sediment and water are shown in black and white circles respectively.

(a), (b) and (c) show the respective estimated mean values for PC1, PC2 and PC3 and their correspondent ions.

**Fig. S3.** Principal component analysis of the ionic concentration at the CCB aquatic systems and other saline environments worldwide (Oren, 2006). Top right table shows the eigenvalues and percentage variance explained by each of the principal components (PCs). Components in red are the ones with the highest eigenvalues and contribution of variance percentage. Plots show the four CCB sites (CH, RM, LH, PA) and the seven worldwide sites against the values for the three PCs in pairwise combinations. Blue and red labels mean water (w) and sediment (s) values from the CCB systems respectively. Seawater KCl means seawater at onset on KCl precipitation, and the same for seawater NaCl.

**Table S1.** Ionic composition of the water and sediment of each aquatic system. Top table shows the concentration of the cations and bottom table shows the concentration of the anions (average and standard deviation). Red numbers show the highest ionic values that differ by at least an order of magnitude.

**Table S2.** Genetic diversity of the phylogroups F1, F2 and F3 with five molecular markers.  $\pi$  is Nei's pairwise nucleotide diversity,  $\theta$  is Waterson's theta per site =  $2Ne\mu$ . Numbers in parenthesis show standard deviation values. H = number of haplotypes, S = number of segregant sites/total nucleotide number, Hd = haplotypic diversity. The two last columns show Tajima's D-values with its respective P-value. Red numbers show the significant values for this test (Rozas and Rozas, 1999).

**Appendix S1.** Supplementary methods.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## SUPPLEMENTARY METHODS

### Ionic concentration and statistical analyses

Sediment and water samples from 14 sites within four aquatic systems (CH, RM, LH, PA) were collected on the summer of 2009. Of the 14 sites that were sampled, three were from CH, six from RM, two from LH, and three from PA. In each site, triplicate samples were obtained from sediment and column water.

To determine the ionic concentration of the samples, sediment samples were incubated with distilled water for 19h at 25°C with continuous shaking. This procedure allows the mobilisation of the available ions within the sediment. Liquid samples from sediment and water were filtered with 0.22 nitrocellulose membranes. Determination of anions ( $\text{CO}_3^{2-}$ ,  $\text{HCO}_3^-$ ,  $\text{Cl}^-$  and  $\text{SO}_4^{2-}$ ) and cations ( $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Na}^+$  and  $\text{K}^+$ ) was performed with the following procedures.

The anions  $\text{CO}_3^{2-}$  and  $\text{HCO}_3^-$  were determined by HCl 0.01N titration.  $\text{Cl}^-$  and  $\text{SO}_4^{2-}$  concentration were determined by liquid chromatography (HPLC). A 432-conductivity detector (Waters) and an IC-Pack HR-Waters column were used. The mobile phase was a mixture of sodium borate (1.3M) and acetonitrile (12%) at pH 8.5.  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  concentrations were obtained by atomic absorption spectroscopy using a Perkin Elmer 3110 equipment.  $\text{Na}^+$  and  $\text{K}^+$  were determined by flamometry using a Corning 400 device.

A Principal Component Analysis was performed with SPSS Statistics v18 (IBM Corporation). The input data consisted of eight variables representing the different ions measured. These measures corresponded to 84 observations given by the triplicate water and sediment samples obtained at the 14 sites. For each principal component a linear mixed model (LMM) was fitted with Stata v11.0 (StataCorp LP). This was performed to determine if its score values are associated with the aquatic system (RM, LH, CH and PA) and the type of environment (sediment or water). The mixed models included two fixed effects: 1) aquatic system (CH, RM,

LH or PA) and 2) environment (sediment or water) plus a random effect for each site to account for the spatial correlation between the three samples taken at each site.

## **Molecular methods**

DNA was extracted from 183 isolates using a 'DNA easy tissue Kit' (Qiagen) following the manufacturer instructions. The *16srRNA* gene was PCR amplified following previous procedures (Cerritos *et al.*, 2008). The 27F fragment was used to determine the identity of the isolates by comparison with other *Exiguobacterium* sequences from the NCBI database using BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) (Altschul *et al.*, 1990). The sequenced, *16srRNA* fragments were assembled using Bioedit v7.0.9 (Hall, 1999).

The partial sequence of the genes *citC*, *recA*, *hsp70* and *rpoB* was PCR amplified by standard procedures. The annealing temperatures for each primer sets were 52°C for *citC*, 47°C for *recA* and *hsp70*, and 53°C for *rpoB*.

All PCR products were sequenced at University of Washington's High Throughput Sequencing Solutions Laboratory. All the sequence data have been submitted to the GenBank database with the following accession numbers: *16srRNA* (JF345522-JF345559, JF345561-JF345566, JF345568, JF345571-JF345609, JF345611-JF345643, JF345645-JF345694, JF345696-JF345713), *citC* (JF916988-JF917080, JF952020-JF952109), *hsp70* (JF952110-JF952292), *recA* (JF952293- JF952475) and *rpoB* (JF952476- JF952658).

## **Genetic analyses**

Population genetics analyses were performed for the five genes mentioned previously. Nucleotide diversity indices  $\pi$  and  $\theta$ , as well as Tajima's *D* values were obtained with DnaSP v5.10.01 (Rozas and Rozas, 1999) according to equations from Nei (1987) and Tajima (1989).

In order to determine the DNA similarity percentage between isolates, a genetic euclidian distance matrix was obtained from the *16srRNA* gene alignment. Sequences were grouped using the furthest-neighbour method at different similarity cut-offs. This allowed us to define the similarity percentage between phylogroups. The same procedure was performed with the concatenated sequence alignment in order to determine the similarity percentage between phylogroups. All the previous analyses were performed with mothur v.1.21.1 (Schloss *et al.*, 2009).

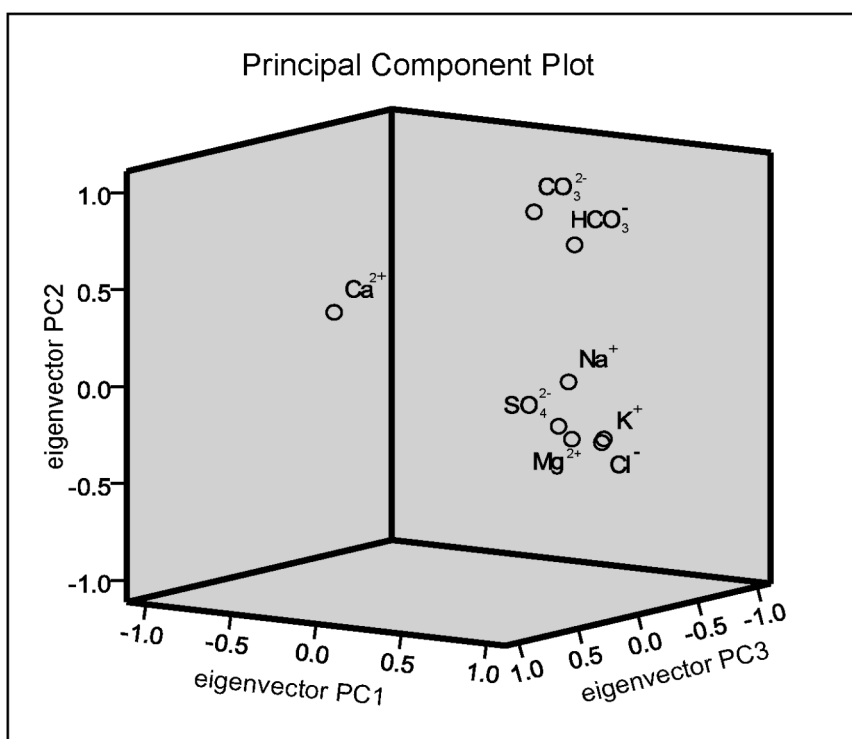


## SUPPORTING FIGURES

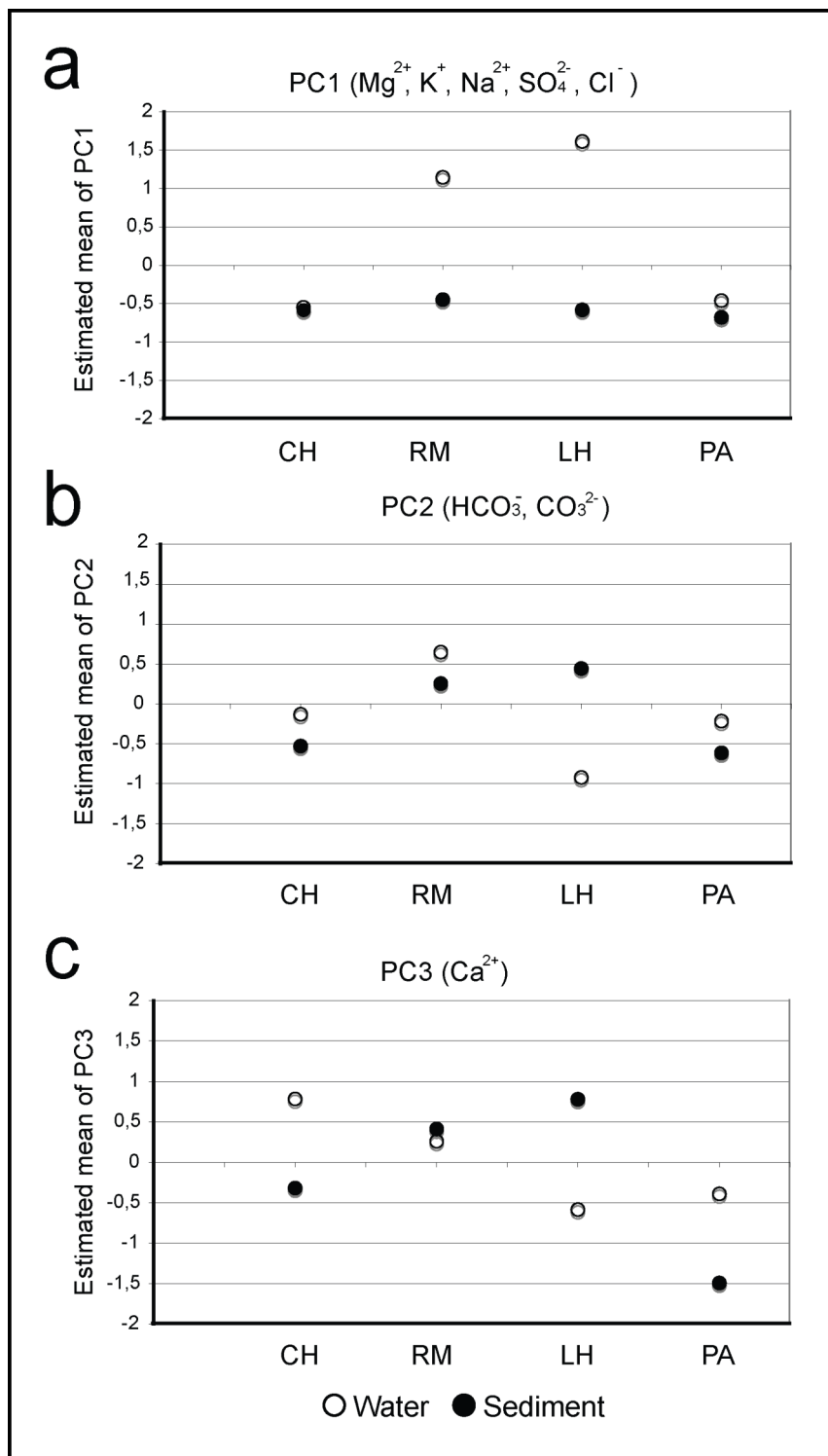
**a**

Principal Component	Initial eigenvalues and total variance explained		
	Total	% of variance	Cumulative variance %
1	3.857	48.214	48.214
2	1.629	20.357	68.571
3	.955	11.939	80.510
4	.756	9.447	89.958
5	.434	5.422	95.380
6	.200	2.495	97.875
7	.104	1.294	99.170
8	.066	.830	100.000

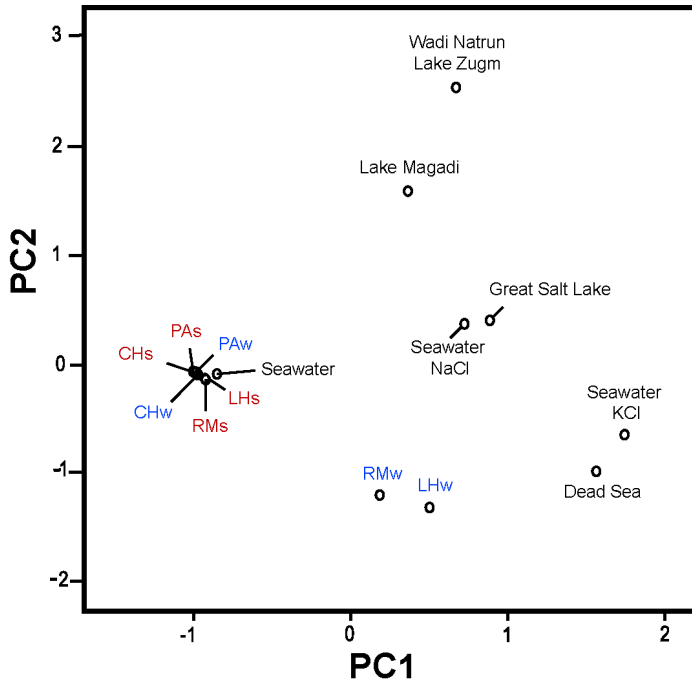
**b**



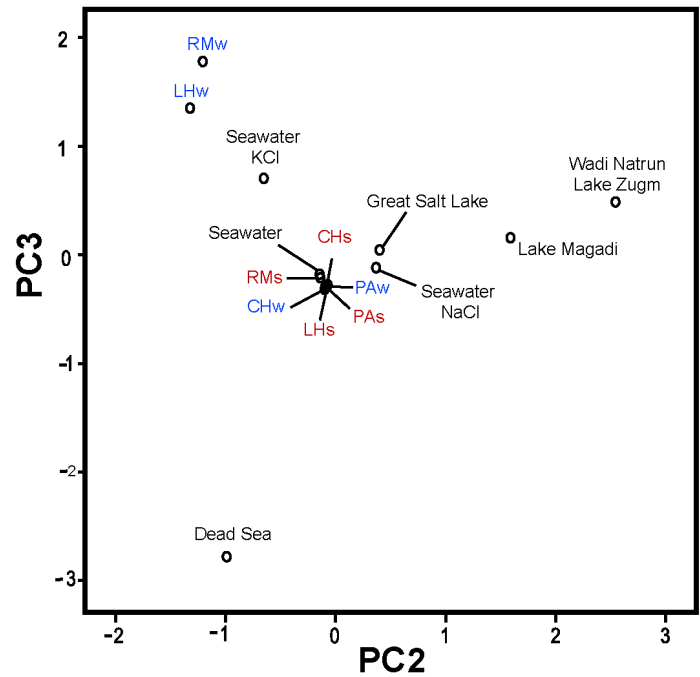
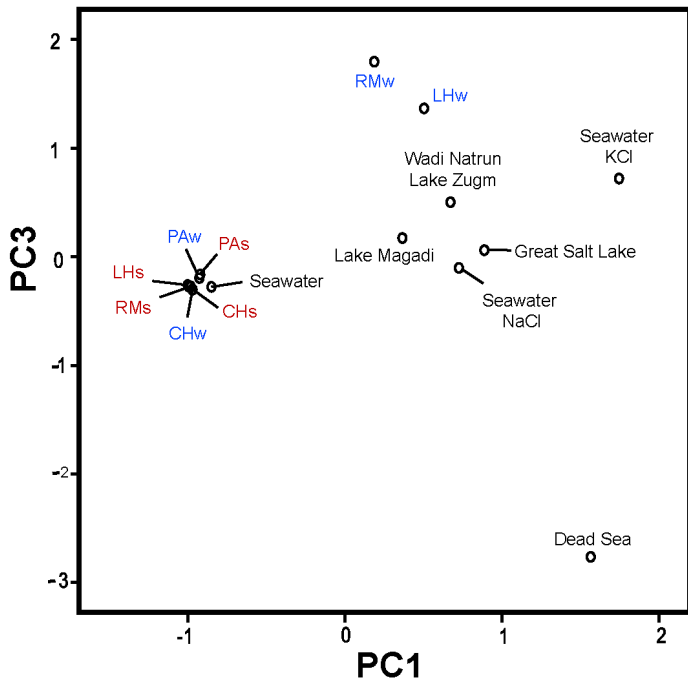
**Figure S1.** Principal Component Analysis of the ionic concentration in the CCB aquatic systems. **(a)** Eigenvalues and percentage of the variance explained by each of the PC. Components in red are the ones with the highest eigenvalues and contribution of variance percentage. **(b)** Plot showing the coefficient values for the three PCs.



**Figure S2.** Estimated means obtained by linear mixed models of the three principal components. X axis shows the aquatic systems analyzed (CH, RM, LH, PA). Y axis show the estimated means of each principal component. Sediment and water are shown in black and white circles respectively. (a), (b) and (c) show the respective estimated mean values for PC1, PC2 and PC3 and their correspondent ions.



Principal Component	Initial eigenvalues and total variance explained		
	Total	% of variance	Cumulative variance %
1	2.992	42.744	42.744
2	1.899	27.128	69.872
3	1.314	18.770	88.642
4	.570	8.145	96.787
5	.156	2.224	99.011
6	.047	.676	99.687
7	.022	.313	100.000



**Figure S3.** Principal Component Analysis of the ionic concentration at the CCB aquatic systems and other saline environments worldwide (Oren, 2006). **Top right table** shows the eigenvalues and percentage variance explained by each of the PCs. Components in red are the ones with the highest eigenvalues and contribution of variance percentage. **Plots** show the four CCB sites (CH, RM, LH, PA) and the seven worldwide sites against the values for the three PCs in pairwise combinations. **Blue** and **red** labels mean water (**w**) and sediment (**s**) values from the CCB systems respectively. **Seawater KCl** means seawater at onset on KCl precipitation, and the same for **Seawater NaCl**

## SUPPORTING TABLES

Aquatic system (Environment)	Na <sup>+</sup> (mg/L)		Mg <sup>2+</sup> (mg/L)		Ca <sup>2+</sup> (mg/L)		K <sup>+</sup> (mg/L)	
	Average	St Dev	Average	St Dev	Average	St Dev	Average	St Dev
CH (water)	159,53	59,2	192,50	19,5	583,97	86,6	12,41	3,2
CH (sediment)	42,19	11,4	22,55	36,1	229,73	196,6	2,78	2,9
PA (water)	483,18	13,5	91,44	2,0	302,25	1,1	7,63	2,4
PA (sediment)	22,71	12,9	1,43	0,7	37,43	6,3	0,96	0,4
LH (water)	<b>4496,32</b>	1409,0	<b>54866,67</b>	6411,1	268,86	90,6	<b>2417,74</b>	2591,2
LH (sediment)	226,92	240,7	1082,59	664,5	576,08	54,2	11,24	7,7
RM (water)	<b>6644,51</b>	5211,8	<b>39667,97</b>	42214,5	365,67	51,8	<b>1160,28</b>	1751,0
RM (sediment)	331,59	327,2	1322,00	1125,3	504,96	71,6	7,65	7,0

Aquatic system (Environment)	Cl <sup>-</sup> (mg/L)		SO <sub>4</sub> <sup>2-</sup> (mg/L)		CO <sub>3</sub> <sup>2-</sup> /HCO <sub>3</sub> <sup>-</sup> (mg/L)	
	Average	St Dev	Average	St Dev	Average	St Dev
CH (water)	131,74	41,9	1326,41	435,8	85,51	80,2
CH (sediment)	53,25	40,1	477,99	422,3	32,86	7,4
PA (water)	93,29	3,3	882,76	35,7	302,83	35,2
PA (sediment)	10,47	3,4	85,43	27,6	28,39	24,8
LH (water)	<b>66372,92</b>	6849,3	<b>56740,52</b>	16029,0	310,30	117,7
LH (sediment)	954,44	654,3	5308,16	2254,2	215,69	172,4
RM (water)	<b>14945,47</b>	15071,9	<b>74302,64</b>	65398,5	392,31	80,7
RM (sediment)	684,75	452,4	6153,24	4738,5	173,60	127,3

**Table S1.** Ionic composition of the water and sediment of each aquatic system. **Top table** shows the concentration of the cations and **bottom table** shows the concentration of the anions (average and standard deviation). Red numbers show the highest ionic values that differ by at least an order of magnitude.

F1=130 isolates							
Genetic marker	$\pi$	$\theta$	H	S	Hd	Tajima's <i>D</i>	P value
<i>16s</i>	0.00104(0.00011)	0.00454(0.00130)	27	32/1293	0.622	-2.43513	P < 0.01
<i>citC</i>	0.00596(0.00091)	0.01458(0.00385)	12	54/679	0.809	-1.88586	P < 0.05
<i>hsp70</i>	0.00371(0.00125)	0.02291(0.00576)	18	84/671	0.671	-2.69481	P < 0.001
<i>recA</i>	0.01613(0.00281)	0.03998(0.00977)	27	129/593	0.838	-2.10852	P < 0.05
<i>rpoB</i>	0.00147(0.00025)	0.00662(0.00207)	12	21/581	0.571	-2.19411	P < 0.01
F2=28 isolates							
Genetic marker	$\pi$	$\theta$	H	S	Hd	Tajima's <i>D</i>	P value
<i>16s</i>	0.00127(0.00015)	0.00118(0.00059)	8	6/1302	0.849	-0.25166	P > 0.10
<i>citC</i>	0.03630(0.00258)	0.02611(0.00863)	12	69/679	0.899	1.06831	P > 0.10
<i>hsp70</i>	0.05028(0.00398)	0.03590(0.01166)	11	95/680	0.841	1.22598	P > 0.10
<i>recA</i>	0.03138(0.00151)	0.02220(0.00748)	14	52/602	0.878	1.36796	P > 0.10
<i>rpoB</i>	0.02012(0.00229)	0.01929(0.00659)	13	44/586	0.878	-0.00946	P > 0.10
F3= 25 isolates							
Genetic marker	$\pi$	$\theta$	H	S	Hd	Tajima's <i>D</i>	P value
<i>16s</i>	0.00282(0.00038)	0.00478(0.00176)	18	24/1302	0.946	-1.57724	P > 0.10
<i>citC</i>	0.04924(0.00638)	0.03407(0.01108)	9	92/694	0.854	1.43159	P > 0.10
<i>hsp70</i>	0.02873(0.00538)	0.02613(0.00868)	11	71/705	0.886	0.32502	P > 0.10
<i>recA</i>	0.04508(0.00733)	0.03476(0.01136)	9	84/621	0.833	0.86258	P > 0.10
<i>rpoB</i>	0.01368(0.00280)	0.01162(0.00423)	9	27/609	0.809	0.65401	P > 0.10

**Table S2.** Genetic diversity of the phylogroups F1, F2 and F3 with five molecular markers.  $\pi$  is Nei's pairwise nucleotide diversity,  $\theta$  is Waterson's theta per site =  $2Ne\mu$ . Numbers in parenthesis show standard deviation values. H = Number of haplotypes, S = number of segregant sites/total nucleotide number, Hd = Haplotypic diversity. The two last columns show Tajima's *D* values with its respective *p*-value. Red numbers show the significant values for this test (Rozas and Rozas 1999).

## REFERENCES

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990). Basic local alignment search tool. *J Mol Biol* 215:403-10.
- Cerritos, R., Vinuesa, P., Eguiarte, L.E., Herrera-Estrella, L., Alcaraz-Peraza, L.D., Arvizu-Gómez, J.L., *et al.* (2008). *Bacillus coahuilensis* sp. nov., a moderately halophilic species from a desiccation lagoon in the Cuatro Ciénegas Valley in Coahuila, Mexico. *Int J Syst Evol Microbiol* 58:919-923.
- Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia Univ. Press, New York. USA.
- Rozas, J., Rozas, R. (1999). DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15: 174-5.
- Schloss P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., *et al.* (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537-41.
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.

## 5. DIVERSIDAD GENÉTICA Y DISTRIBUCIÓN DE *Exiguobacterium* EN CCC.

Con base en la colección de aislados que se obtuvo en este proyecto, fue posible describir la presencia de tres linajes o filogrupos del género *Exiguobacterium*. Como parte del proyecto de doctorado se utilizaron herramientas de la genética de poblaciones para describir la dinámica poblacional de estos filogrupos en contraste con otros géneros bacterianos descritos en CCC. Así mismo, se describió la distribución de estos filogrupos, así como de cada uno de los genotipos, en los cuatro sistemas principales de CCC. Ambos resultados se describen a continuación como parte complementaria de los resultados plasmados en el **capítulo 4**.

### 5.1 Genética de poblaciones comparada: el caso de *Exiguobacterium* con respecto a otros géneros en CCC.

A partir de la filogenia obtenida de cuatro genes constitutivos (*citC*, *hsp70*, *recA* y *rpoB*) se pudieron definir tres linajes genéticos del género *Exiguobacterium*. Estos linajes se confirmaron por medio de un análisis de STRUCTURE (Capítulo 4). Así mismo se realizaron dos pruebas de diferenciación genética con el objetivo de evaluar si existe estructura genética entre estos tres filogrupos. Como se observa en la Tabla 2, los valores de diferenciación con ambos índices confirman la observación de que estos filogrupos se encuentran genéticamente diferenciados.

**Tabla 2. Índices de diferenciación genética entre filogrupos de *Exiguobacterium*.**

$N_{st}$ pareada / $F_{st}$ pareada	F1(130)	F2(28)	F3(25)
F1(130)	-	0.59255	0.80797
F2(28)	0.58936	-	0.70442
F3(25)	0.79848	0.68936	-

**Tabla 2**

Índices de diferenciación genética pareada entre filogrupos. Los números en paréntesis indican el número de muestra. Los valores sobre la diagonal muestran los valores de  $N_{st}$  pareada ( $p < 0.0001$ ) (Lynch y Crease, 1990). Los valores por debajo de la diagonal muestran los valores de  $F_{st}$  pareada ( $p < 0.0001$ ) (Hudson *et al.*, 1992).

Por otra parte, se calcularon los estimados de diferenciación genética entre los cuatro sistemas acuáticos. Como se muestra en la Tabla 3, los índices de diferenciación entre estos sitios es muy baja, lo cuál indica que no existe una estructura poblacional asociada al origen de los aislados. Asimismo, se realizó una prueba de Mantel con el fin de evaluar si los genotipos de esta muestra reflejaban algún patrón de aislamiento por distancia. Esta prueba fue no significativa, lo cuál refleja la ausencia de patrones asociados a la localización geográfica de los sistemas acuáticos.

**Tabla 3. Índices de diferenciación genética entre sistemas acuáticos**

$N_{st}$ / $F_{st}$	Churince	Rio Mesquites	Los Hundidos	Pozas Azules
<b>Churince (42)</b>	-			
<b>Rio Mesquites (93)</b>	0.01774/ 0.01818	-		
<b>Los Hundidos (35)</b>	0.02616/ 0.02526	0.09149/ 0.08937	-	
<b>Pozas Azules (13)</b>	0.12731/ 0.12793	0.13856/ 0.13697	0.08170/ 0.08123	-

**Tabla 3**

Índices de diferenciación genética pareada entre sistemas acuáticos. Los números en paréntesis indican el número de muestra. Los valores sobre la diagonal muestran los valores de  $N_{st}$  pareada ( $p < 0.0001$ ) (Lynch y Crease, 1990). Los valores por debajo de la diagonal muestran los valores de  $F_{st}$  pareada ( $p < 0.0001$ ) (Hudson *et al.*, 1992).



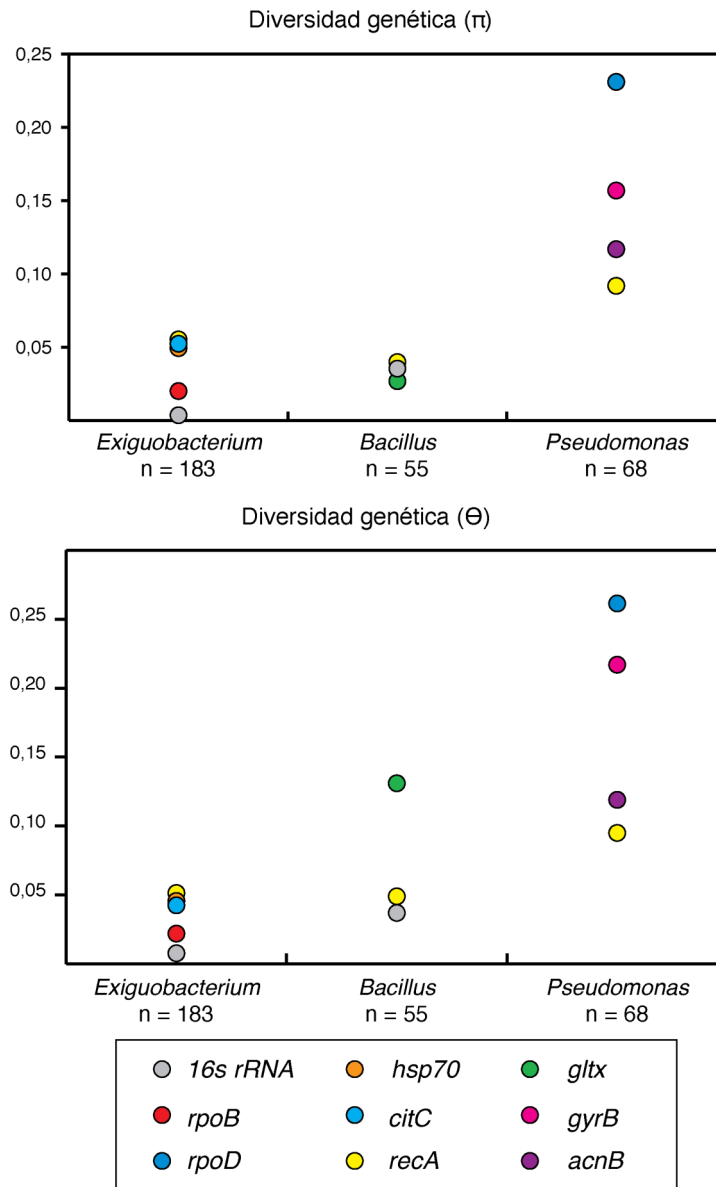
Con el fin de entender la enorme diversidad de especies procariontes presentes en el planeta, es necesario describir la diversidad genética presente en las poblaciones naturales de bacterias. Por lo anterior, previos a este trabajo, se han realizado estudios de genética de poblaciones de dos de los géneros más abundantes en el valle de CCC: *Pseudomonas* y *Bacillus* (Escalante, 2008; Cerritos, 2009). Es interesante contrastar estos trabajos previos con los resultados obtenidos en esta tesis, y de esta manera tener una perspectiva sobre la genética de las poblaciones naturales presentes en CCC. *Pseudomonas* es un género de bacterias gram negativas con una enorme plasticidad metabólica y genética (Spiers *et al.*, 2000; Rius *et al.*, 2001). Por otra parte el género *Bacillus* son bacterias gram positivas formadoras de esporas y con una amplio rango de estilos de vida. Ambos géneros son considerados cosmopolitas ya que se encuentran presentes en todo tipo de ambientes (acuáticos y terrestres) incluyendo asociaciones simbióticas con otros organismos (plantas y animales)(Raymond *et al.*, 2010; Maughan y Van Der Auwera 2011). Por otra parte, *Exiguobacterium* es un género de bacterias gram-positivas que es filogenéticamente cercano a *Bacillus* (Farrow *et al.*, 1994). Sin embargo, las especies reportadas hasta la fecha no forman esporas y en su mayoría son especies extremófilas con respecto a diferentes parámetros ambientales (**Capítulo 2.2**). Con base en lo anterior, resulta interesante hacer la comparación entre tres géneros con biología contrastantes pero que se encuentran ampliamente distribuidos en el valle de CCC.

Los muestreos de *Pseudomonas* realizaron en el sistema acuático de Churince con un número de muestra de aproximadamente 68 aislados. Por otra parte, el muestreo de *Bacillus* se llevo a cabo en el sistema Churince y Rio Mezquites con un número de muestra de 55 cepas. El muestreo realizado a lo largo de este proyecto involucró el muestreo de cuatro sistemas acuáticos (Churince, Rio Mezquites, Los Hundidos y Pozas Azules) obteniendo un número de muestra de aproximadamente el triple de los muestreo

anteriores (n=183). Al momento de comparar los índices de diversidad nucleotídica es importante considerar esta última observación.

En primer lugar se comparó la diversidad genética ( $\pi$  y  $\theta$ ) de *Pseudomonas*, *Bacillus* y *Exiguobacterium* con base en los datos obtenidos en estudios previos y en este trabajo (Escalante, 2008; Cerritos, 2009; Rebollar *et al.*, 2012). Es importante destacar que se utilizaron marcadores moleculares diferentes para cada género, sin embargo todos ellos corresponden a genes constitutivos conservados en todos los linajes bacterianos y que son comúnmente usados en análisis multilocus.

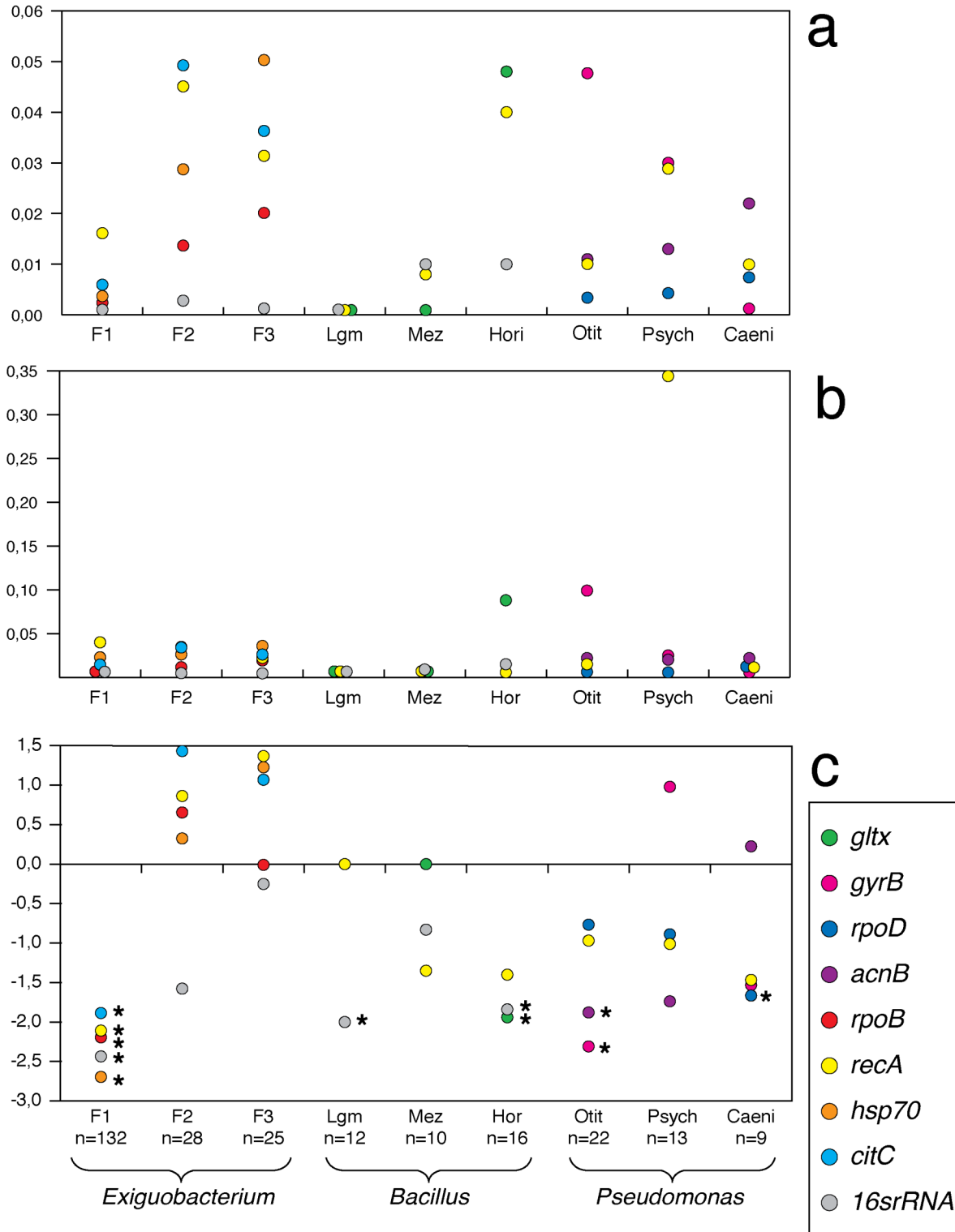
Como se observa en la Figura 11, la diversidad genética en *Pseudomonas* ( $\pi$  y  $\theta$ ) es mayor a la diversidad en *Bacillus* y *Exiguobacterium*. Lo anterior, se debe posiblemente a que existe una mayor diversidad de linajes de *Pseudomonas* que de los otros dos géneros, a pesar de que este género sólo fue muestreado en un sistema acuático y con mucho menor número de muestra que *Exiguobacterium*. Asimismo, la varianza de la diversidad  $\theta$  utilizando diferentes marcadores, es menor en *Exiguobacterium* a la que se observa en los otros dos géneros.



**Figura 11**

Análisis de diversidad genética de tres géneros bacterianos a partir de diferentes marcadores moleculares.  $\pi$  diversidad nucleotídica pareada de Nei,  $\theta$  de Waterson por sitio =  $2Ne$  (Rozas y Rozas 1999).

Con el objetivo de describir la diversidad genética al interior de las especies, se graficó la diversidad nucleotídica  $\pi$  y  $\theta$  de los tres linajes principales dentro de cada género (Figura 12a y 12b).



**Figura 12**

Índices de diversidad nucleotídica y D de Tajima para los tres linajes principales de *Exiguobacterium*, *Bacillus* y *Pseudomonas* del valle de CCC. (a)  $\pi$  diversidad nucleotídica pareada de Nei, (b)  $\theta$  de Waterson por sitio =  $2Ne$ , (c) D de Tajima (Rozas y Rozas 1999). \* Indica los valores significativos para la prueba D de Tajima ( $<0.05$ ).

Con respecto a los linajes de *Exiguobacterium*, es claro que **F1** tiene los valores más bajos de diversidad nucleotídica pareada ( $\pi$ ) (Figura 12a). Lo anterior sugiere que existen unos cuantos haplotipos abundantes y pocos haplotipos raros en la muestra. Una tendencia similar se observa en los linajes de *Bacillus* en donde **Lgm** y **Mez** tienen valores mucho más bajos que **Hori**. Finalmente la diversidad nucleotídica al interior de los linajes de *Pseudomonas* presenta niveles semejantes a excepción de *gyrB* en **Otit**. Los valores de diversidad  $\theta$  en todos los linajes es muy semejante a excepción de *gyrB* en **Otit** y *gltX* en **Hor**. Es importante considerar que la diversidad  $\pi$  es un parámetro que es sensible a la frecuencia de cada una de las secuencias mientras que  $\theta$  es más sensible a la presencia de alelos raros en la muestra (Hedrick, 2005).

Considerando lo anterior ambas medidas son importantes para describir la diversidad genética de un grupo. La prueba de D de Tajima es un parámetro útil para comparar ambos estimados de diversidad ya que bajo el modelo de alelos infinitos, en un escenario de neutralidad, el valor de  $\pi$  y  $\theta$  tienen la misma magnitud (Hedrick, 2005). Esta prueba ha sido ampliamente utilizada para detectar eventos de selección positiva o negativa (fuera de la neutralidad) pero también es un indicador importante del comportamiento demográfico de una población (Shapiro *et al.*, 2009). Como se observa en la figura 12c la mayor parte de los valores de la prueba de Tajima son negativos, lo cual sugiere que estos marcadores se encuentran sujetos a selección purificadora. La anterior interpretación es congruente con las características de los marcadores utilizados, ya que en todos los casos, son genes constitutivos que se definen por tener tasas bajas de mutación y alto grado de conservación.

En el caso de *Exiguobacterium*, es interesante observar que **F2** y **F3** tienen altos valores de  $\pi$  y por lo tanto poseen valores positivos de la prueba de D de Tajima. Este comportamiento contrasta con lo observado en **F1**. Este último filogrupo tiene valores

negativos y significativos para la prueba de Tajima en todos los marcadores, lo cuál sugiere que este linaje se encuentra en proceso de expansión demográfica.

Considerando todos los datos anteriores, es claro que existen diferencias en la diversidad nucleotídica entre los tres géneros. Sin embargo al analizar los principales linajes (o especies) dentro de cada muestra, los valores de diversidad son variados pero se encuentran dentro de un mismo rango de valores. Asimismo, estos valores de diversidad son semejantes a la diversidad genética de bacterias patógenas y ambientales de otras partes del mundo. Tal es el caso del género *Vibrio* (Bisharat *et al.*, 2007; Keymer y Boehm 2011), *Neisseria* (Perez-Lozada *et al.*, 2005), *Legionella* (Coscollá y González-Candela 2007), *Pseudomonas* (Sarkar y Guttman 2004) y diversas especies de *Bacillus* (Sorokin *et al.*, 2006, Zwick *et al.*, 2011).

Otro de los aspectos interesantes a evaluar es el grado de recombinación presentes en cada uno de los tres géneros (entre y dentro de los linajes). La recombinación en *Exiguobacterium* fue abordada previamente (Capítulo 4). Por otro lado, el número de muestra de los linajes de *Pseudomonas* y *Bacillus* no son suficientemente grandes para realizar análisis profundos acerca de la recombinación al interior de estos grupos. Sin embargo, los análisis preliminares que se han realizado sugieren que ambos grupos son clonales (Escalante, 2008; Cerritos, 2009). Lo anterior es congruente con estudios sobre el grado de recombinación de diversas especies dentro de estos dos géneros (Vos y Didelot 2009).

## 5.2 La distribución de los filogrupos de *Exiguobacterium* en el valle.

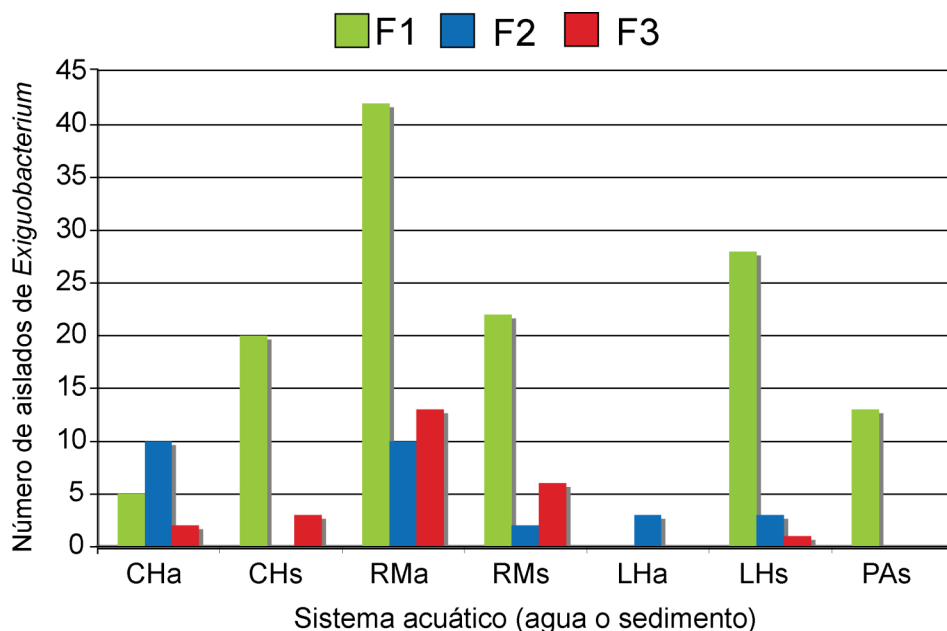
Estudios previos acerca de la distribución de especies bacterianas en CCC han revelado que existen claras diferencias en la composición bacteriana entre sitios. En particular, por

medio de T-RFLPs (“Terminal Restriction Fragment Length Polymorphism”) y librerías de clonas se determinó que existen claras diferencias en la composición de las comunidades bacterianas entre diferentes sistemas acuáticos (alta diversidad  $\beta$ ; Escalante *et al.*, 2008).

Por otra parte, se han realizado estudios detallados sobre la diversidad bacteriana en el sistema acuático Churince. Este sistema está compuesto por dos lagunas de desecación alimentadas por un manantial. La distribución de estos cuerpos de agua se caracteriza por presentar gradientes pronunciados de salinidad, pH, temperatura y oxígeno disuelto (Cerritos *et al.*, 2010). Se tomaron muestras de bacterias cultivables gram positivas de este sistema acuático y se determinó que la composición de especies cambia conforme el gradiente ambiental. En particular, se determinó que la salinidad es el componente ambiental que explica gran parte de la distribución de la diversidad microbiana (Cerritos *et al.*, 2010). En este mismo sistema acuático se encontró que existe una abundancia diferencial de especies de *Pseudomonas* en Laguna Grande, Intermedia y el manantial (Escalante, 2008). Sin embargo, en este caso no se encontró una distribución asociada a alguno de los parámetros ambientales medidos. Interesantemente la presencia de algunas de las especies de *Pseudomonas* de este sistema acuático presentan un patrón estacional diferencial (Rodríguez-Verdugo, 2008).

Tomando en cuenta todos estos antecedentes, así como la pregunta principal de este proyecto, era importante determinar la distribución de *Exiguobacterium* en el valle. En particular, estudios previos de *Exiguobacterium* han sugerido que las especies de este género se distribuyen con respecto a parámetros ambientales como la salinidad y el pH (Rodríguez *et al.*, 2009). Por lo mismo, es interesante determinar si al interior de los linajes existe una estructuración asociada a algún parámetro ambiental, como la salinidad.

Como parte esencial del este proyecto, se determinó la presencia de los filogrupos en diferentes sistemas acuáticos con el fin de determinar si la salinidad o el aislamiento geográfico juegan un papel importante en la distribución de estos linajes.



**Figura 13**

Distribución de los aislados en los cuatro sistemas acuáticos de CCC. Cada color representa un filogrupo. Los subíndices **a** y **s** significan agua y sedimento respectivamente.

En primer lugar, se analizó la distribución de los filogrupos F1, F2 y F3 en los cuatro sistemas acuáticos. Como se observa en la Figura 13, F1 presenta la distribución más amplia en los cuatro sistemas acuáticos analizados. De hecho el único ambiente en donde F1 no se encontró fue en el agua de LH. Por otra parte, F2 se encuentra presente en todos los ambientes a excepción del sedimento de CH y PA. Finalmente F3, se localiza principalmente en RM aunque también se detectó en el agua de CH, así como en el agua y sedimento de LH. A pesar de que esta distribución no es homogénea entre los sistemas, ésta no correlaciona con la salinidad de los ambientes. Como se describió en el **capítulo 4**, la concentración de sales en el agua de LH y RM es mucho mayor que el resto de los ambientes. Sin embargo, la presencia de los filogrupos en estos dos



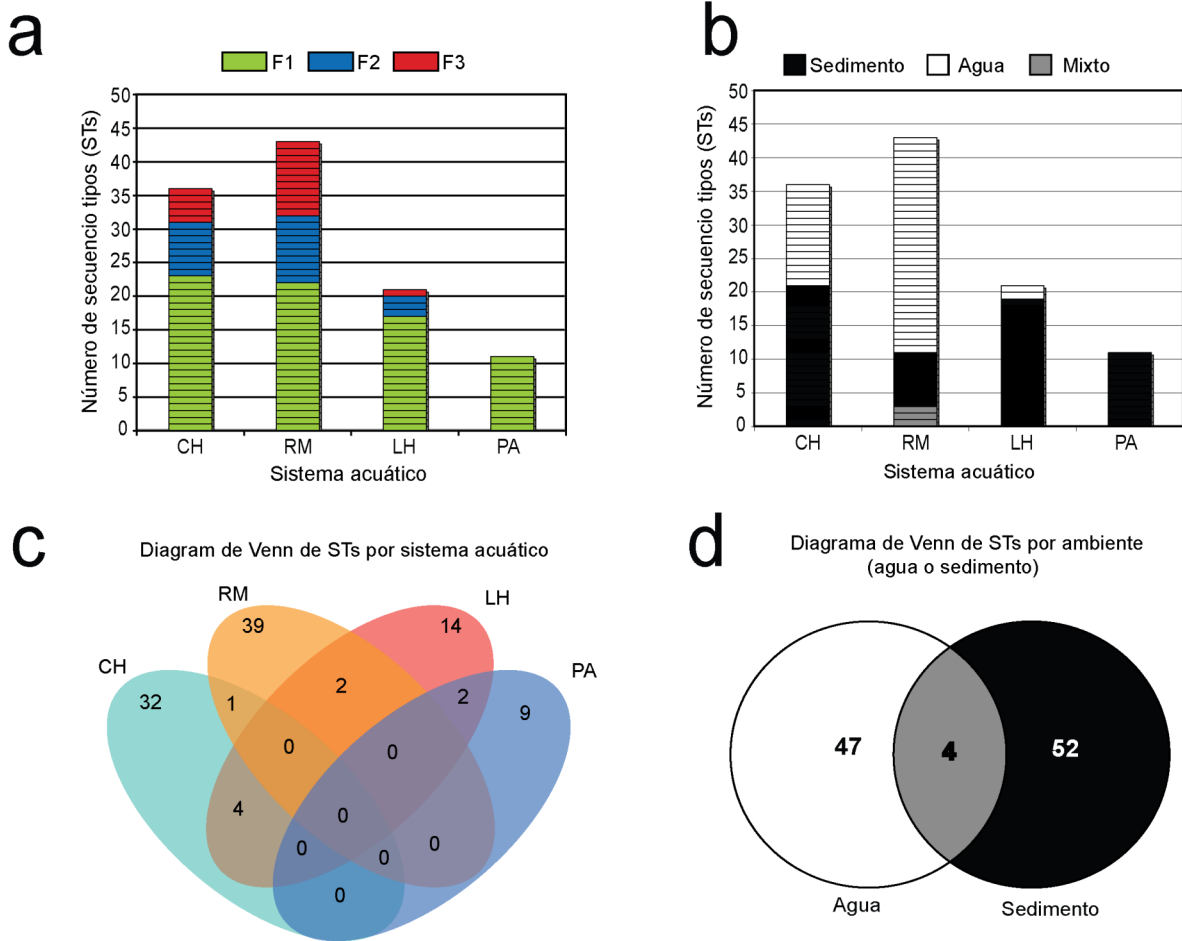
sistemas es totalmente distinta: En el agua de RM se encuentran presentes los tres filogrupos, siendo F1 el más abundante, mientras que en el agua de LH, sólo se detectó la presencia de F2.

Con el objetivo de evaluar a mayor profundidad la distribución de cada uno de los aislados se decidió analizar la abundancia y distribución de los secuencia tipos (STs) presentes en toda la muestra. Los STs se obtuvieron con base en la secuencia concatenada de cuatro genes constitutivos (*citC*, *hsp70*, *recA* y *rpoB*) y se definieron como secuencias únicas por medio de un análisis de haplotipos utilizando DNAsp (Rozas y Rozas 1999). A partir de este análisis se definieron 103 STs de los 183 aislados de *Exiguobacterium*.

Como se muestra en la Figura 14a, la distribución de los filogrupos es diferente entre los sistemas acuáticos, siendo F1 el filogrupo cuya distribución es más amplia. Es importante aclarar que aunque PA es el sistema con menor número de STs (y aislados), las cepas se obtuvieron a partir del mismo esfuerzo de muestreo en todos los sistemas (Capítulo 4). Lo anterior sugiere que la abundancia de *Exiguobacterium* en PA es mucho menor que en el resto de los sistemas.

Además de la distribución de los STs según el linaje filogenético al que pertenecen, se comparó la distribución de éstos con respecto a si provenían de agua, sedimento o de un ambiente mixto (Figura 14b). Como se puede observar en esta figura la abundancia de STs que provienen de STs de sedimento o de agua tampoco es homogénea. Por un lado los sistemas PA y LH están compuestos principalmente por STs de sedimento. Por otro lado, RM tiene una mayor abundancia de STs de agua. Finalmente CH, es un sistema en donde los STs de agua y de sedimento están presentes de manera equitativa. Considerando lo anterior, es claro que la presencia y abundancia de los STs de sedimento

y agua, tampoco correlaciona con la salinidad de cada uno de estos ambientes. Asimismo, la distancia geográfica entre los sistemas tampoco tiene un papel relevante en la distribución de estos STs (Capítulo 4).



**Figura 14**

Distribución de los 103 STs de *Exiguobacterium* en el valle de CCC. **(a)** Abundancia y distribución de STs en los cuatro sistemas acuáticos analizados según el filogrupo al que pertenecen. **(b)** Abundancia y distribución de STs en los cuatro sistemas acuáticos analizados según el ambiente al que pertenecen (sedimento, agua o mixto). **(c)** Diagrama de Venn representando el número de STs por sistema acuático y el número de éstos que se comparten entre sistemas. **(d)** Diagrama de Venn representando el número de STs únicos de sedimento y agua así como el número de STs compartido entre estos dos ambientes.

Otro de los resultados interesantes de este análisis es que de los 103 STs de la muestra, sólo 9 de ellos son compartidos entre dos de los sistemas (Figura 14c) y

ninguno de estos se comparte por tres o más sistemas acuáticos. Lo anterior sugiere que la composición de STs en cada sistema es único a pesar de que no existe un aislamiento geográfico entre los aislados de cada sistema. Como se mencionó anteriormente, CCC se caracteriza por tener una alta diversidad beta al comparar las comunidades microbianas en diferentes sistemas dentro del valle (Escalante *et al.*, 2008). Es interesante que incluso al interior de un sólo género bacteriano, se observe una composición única de STs por sistema acuático. Es claro que este patrón no correlaciona ni con la distancia geográfica ni con la salinidad de cada uno de los sistemas. Por lo tanto, es posible que esta composición diferencial de STs esté asociada a otras variables como la abundancia de nutrientes, así como a las relaciones de competencia y cooperación entre los organismos de un mismo sistema acuático.

Finalmente, otro de los aspectos interesantes de este análisis es la comparación de STs entre los ambientes de sedimento y agua (Figura 14d). De los 103 STs de la muestra, sólo cuatro son compartidos entre agua y sedimento. Lo anterior indica que la composición de aislados en estos dos ambientes es única e independiente. Este resultado correlaciona perfectamente con la presencia de subgrupos genéticos al interior cada uno de los tres filogrupos, sugiriendo así, la existencia de un proceso de diferenciación genética y ecológica asociado a los ambientes de sedimento y agua (ver **Capítulo 4**).

**6. ARTÍCULO 2:** “*Exiguobacterium halophilus* sp. nov. y *Exiguobacterium coahuilensis* sp. nov., dos nuevas especies halófilas aisladas del valle de Cuatro Ciénegas Coahuila, Mexico.\*

### 6.1 RESUMEN

En este trabajo se describieron dos cepas bacterianas, 11-28<sup>T</sup> y EPVM<sup>T</sup>, que fueron aisladas de dos sistemas acuáticos diferentes dentro del valle de Cuatro Ciénegas Coahuila, México. Estas cepas son bacterias gram-positivas, esféricas (1.0 x 1.2-1.5 µm) y sin movimiento. Ambas cepas crecen entre 28 y 37°C, en una salinidad de 0.11-10% (w/v) NaCl y a un pH de 7.0-9.5. La reconstrucción filogenética del gen *16srRNA* mostró que las cepas 11-28<sup>T</sup> y EPVM<sup>T</sup> pertenecen al género *Exiguobacterium* y son filogenéticamente cercanas a *Exiguobacterium mexicanum* (99.9% y 99.7% respectivamente). El análisis filogenético de las secuencias concatenadas de cuatro genes codificantes (*citC*, *hsp70*, *recA* and *rpoB*) muestra que estas dos cepas son linajes filogenéticos distintos. Los ácidos grasos que se encuentran en mayor proporción en las cepas 11-28<sup>T</sup> y EPVM<sup>T</sup> son iso-C<sub>13:0</sub> y anteiso-C<sub>13:0</sub> (11.13/25.11% y 10.81/27.34% respectivamente). La proporción del resto de los ácidos grasos en ambos aislados es diferente con respecto a las cepas tipo más cercanas. Estas cepas tienen un porcentaje de GC de 50.80% en la cepa 11-28<sup>T</sup> y 51.40% en la cepa EPVM<sup>T</sup>. La similitud DNA-DNA

---

\* Este artículo se encuentra en preparación y los nombres de las especies son nombres propuestos sujetos a cambios.

entre 11-28<sup>T</sup> y EPVM<sup>T</sup> es del 63%. Así mismo, en comparación con *E. mexicanum* 8N<sup>T</sup>, 11-28<sup>T</sup> y EPVM<sup>T</sup> tienen un porcentaje de similitud del 60% y 61% respectivamente. Con base en los análisis filogenéticos y fenotípicos proponemos que estas cepas deben ser consideradas como especies nuevas pertenecientes al género *Exiguobacterium*. Para la cepa 11-28<sup>T</sup> (= LMG 26555<sup>T</sup> = CECT 7978<sup>T</sup>) se propone el nombre “*Exiguobacterium halophilus* sp. nov.” El nombre propuesto para la cepa EPVM<sup>T</sup> (= LMG 26554<sup>T</sup> = CECT 7977<sup>T</sup>) es “*Exiguobacterium coahuilensis* sp. nov.”

*Exiguobacterium halophilus* sp. nov., and *Exiguobacterium coahuilensis* sp. nov., two novel halophilic species isolated from the Cuatro Ciénegas Basin, Mexico.

Eria A. Rebollar<sup>1</sup>, Luis E. Eguiarte<sup>1</sup> and Valeria Souza<sup>1\*</sup>

<sup>1</sup>Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Apartado Postal 70-275, México D.F. 04510, México.

\*Corresponding author: Valeria Souza Saldívar

Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Circuito exterior s/n anexo al Jardín Botánico Exterior. Apartado Postal 70-275, México D.F. 04510, México.

Tel. (+52 55 56229006) Fax: (+52 55 56228995)

E-mail: [souza@servidor.unam.mx](mailto:souza@servidor.unam.mx)

Running title: Two novel halophilic species of *Exiguobacterium*.

Contents Category: New Taxa (Firmicutes and Related Organisms).

Abbreviations: CCB (Cuatro Ciénegas Basin) and MAM (Marine Agar Media).

The GenBank accession numbers for the 16S rRNA, *citC*, *hsp70*, *recA* and *rpoB* gene sequences of *Exiguobacterium halophilus* strain 11-28<sup>T</sup> are JF345614, JF917037, JF952159, JF952342 and JF952525 respectively.

The GenBank accession numbers for the 16S rRNA, *citC*, *hsp70*, *recA* and *rpoB* gene sequences of *Exiguobacterium coahuilensis* strain EPVM<sup>T</sup> are JF345660, JF952074, JF952257, JF952430 and JF952623 respectively.

1 **ABSTRACT**

2 Two bacterial strains designated 11-28<sup>T</sup> and EPVM<sup>T</sup> were isolated from two different  
3 aquatic systems in the Cuatro Ciénegas Basin, Mexico. These strains are gram-  
4 positive, nonmotile coccoid rods (1.0 x 1.2-1.5 µm). Both strains exhibited growth  
5 between 28 and 37°C, between 0.11% and 10% (w/v) NaCl and with pH 7.0 to 9.5.  
6 Phylogenetic reconstructions based on the 16S rRNA gene sequences showed that  
7 strains 11-28<sup>T</sup> and EPVM<sup>T</sup> belong to the *Exiguobacterium* genus and are closely related  
8 to *Exiguobacterium mexicanum* (99.9% and 99.7% respectively). Phylogenetic analyses  
9 of concatenated sequences of four genes (*citC*, *hsp70*, *recA* and *rpoB*) revealed that  
10 11-28<sup>T</sup> and EPVM<sup>T</sup> are distinct phylogenetic lineages. The major fatty acids present in  
11 11-28<sup>T</sup> and EPVM<sup>T</sup> were iso-C<sub>13:0</sub> and anteiso-C<sub>13:0</sub> fatty acids (11.13/25.11% and  
12 10.81/27.34% respectively). The composition of the rest of the fatty acids in both  
13 isolates was different in comparison to closely related type isolates. The GC content  
14 was 50.80% in strain 11-28<sup>T</sup> and 51.40% in strain EPVM<sup>T</sup>. DNA-DNA relatedness  
15 between 11-28<sup>T</sup> and EPVM<sup>T</sup> was 63%. Furthermore, compared to *E. mexicanum* 8N<sup>T</sup>,  
16 11-28<sup>T</sup> and EPVM<sup>T</sup> had a DNA relatedness of 60% and 61% respectively. Based on the  
17 phylogenetic and phenotypic analyses we propose that these strains should be  
18 considered as new species within the genus *Exiguobacterium*. The name proposed for  
19 the strain 11-28<sup>T</sup> (= LMG 26555<sup>T</sup> = CECT 7978<sup>T</sup>) is *Exiguobacterium halophilus* sp. nov.  
20 The name proposed for the strain EPVM<sup>T</sup> (= LMG 26554<sup>T</sup> = CECT 7977<sup>T</sup>) is  
21 *Exiguobacterium coahuilensis* sp. nov.

22

23

24

25 The first species of the genus *Exiguobacterium* was described by Collins and  
26 colleagues in 1983 (Collins *et al.*, 1983). This study determined that *Exiguobacterium*  
27 were gram-positive, alkaliphilic, non-spore forming bacteria. Further phylogenetic  
28 studies analyzing the 16S rRNA gene confirmed the new genus as a distinct clade  
29 closely related to the *bacilli* 2 cluster (Farrow *et al.*, 1994). At the time of writing, 14  
30 species of the genus have been reported: *Exiguobacterium aurantiacum* (Collins *et al.*,  
31 1983), *Exiguobacterium acetylicum* (Farrow *et al.*, 1994), *Exiguobacterium undae* and  
32 *Exiguobacterium antarcticum* (Fruhling *et al.*, 2002), *Exiguobacterium oxidotolerans*  
33 (Yumoto *et al.*, 2004), *Exiguobacterium aestuarii* and *Exiguobacterium marinum* (Kim *et*  
34 *al.*, 2005), *Exiguobacterium mexicanum* and *Exiguobacterium artemiae* (López-Cortés  
35 *et al.*, 2006), *Exiguobacterium sibiricum* (Rodrigues *et al.*, 2006), *Exiguobacterium*  
36 *indicum* (Chaturvedi & Shivaji, 2006), *Exiguobacterium profundum* (Crapart *et al.*,  
37 2007), *Exiguobacterium soli* (Chaturvedi *et al.*, 2008) and the recently described  
38 *Exiguobacterium aquaticum* (Raichand *et al.*, 2011).

39 *Exiguobacterium* species have been isolated from different environments such as  
40 industrial wastes (Collins *et al.*, 1983; Farrow *et al.*, 1994; Yumoto *et al.*, 2004),  
41 freshwater (Fruhling *et al.*, 2002; Raichand *et al.*, 2011), marine water (Kim *et al.*, 2005;  
42 Crapart *et al.*, 2007) as well as permafrost and glaciers (Rodrigues *et al.*, 2006;  
43 Chaturvedi & Shivaji, 2006; Chaturvedi *et al.*, 2008). Even though *Exiguobacterium*  
44 species live in contrasting environments, all of them share the feature of being  
45 alkaliphilic and halotolerant. In addition, phylogenetic analysis using the 16S rRNA  
46 sequence have shown the presence of two species clusters with contrasting  
47 physiological features such as the capacity to reduce nitrate into nitrite (Crapart *et al.*,  
48 2007).



49 Several strains of *Exiguobacterium* have been isolated from the Cuatro Ciénegas  
50 Basin (CCB) in Coahuila, Mexico (Rebollar *et al.*, 2012). The CCB is part of the  
51 Chihuahuan desert located at the north eastern region of Mexico. This basin comprises  
52 many aquatic environments that have contrasting physicochemical conditions and  
53 harbour unique bacterial communities (Souza *et al.*, 2006). Moreover, several new  
54 species belonging to different genera had been isolated from these aquatic  
55 environments (Cerritos *et al.*, 2008; Escalante *et al.*, 2009; Avitia *et al.*, *submitted*).

56 In this work we characterised two *Exiguobacterium* isolates (11-28<sup>T</sup> and EPVM<sup>T</sup>)  
57 from two aquatic systems in the CCB. These isolates possess several molecular and  
58 physiological features that are different from those previously reported species.  
59 Therefore we propose these strains to be considered as new species within the  
60 *Exiguobacterium* genus.

61 Isolate 11-28<sup>T</sup> was recovered from water of a desiccation pond near Rio Mezquites  
62 (26°55.180' N, 102°06.340' W). Isolate EPVM<sup>T</sup> was obtained from sediment of a  
63 desiccation pond in Los Hundidos system (26°52'17" N, 102°01'11.3" W). Both  
64 desiccation ponds are unique hypersaline environments, with high concentrations of ions  
65 Mg<sup>+</sup>, Na<sup>+</sup>, K<sup>+</sup>, SO<sub>4</sub><sup>-</sup> and Cl<sup>-</sup> in comparison to seawater ionic proportions (Rebollar *et al.*,  
66 2012).

67 Water and sediment samples were diluted using saline solution of 1% (w/v) NaCl  
68 at different concentrations (1:10, 1:100, 1:1000 and 1:10000). From each dilution, 200  
69 µl were incubated in Marine Agar Media (MAM; Difco 2216) at 37°C under aerobic  
70 conditions. After 24h of incubation, *Exiguobacterium* isolates grew on the Petri dishes  
71 as bright orange colonies. All colonies were purified by single colony isolation on MAM  
72 to obtain a single clone plate. Once isolated, cell morphology and gram-staining

73 analyses for each of the two isolates were performed by the BCCM™/LMG consortium.  
74 Both isolates were gram-positive, nonmotile coccoid rods (1.0 x 1.2-1.5 µm) present as  
75 single cells or in pairs.

76 The growth of 11-28<sup>T</sup> and EPVM<sup>T</sup> was tested in two independent pH and salinity  
77 gradients using modified MAM (from pH 4.5 to 9.5 and from 0.11% to 30% (w/v) NaCl  
78 respectively; Table 1). In each condition growth was measured in triplicates. Both  
79 isolates grew in saline conditions up to 10% (w/v) NaCl (1.7M) and in alkaline conditions  
80 up to pH 9.5. These results are in agreement with the data obtained for the 14 type  
81 species previously reported for the genus (Table1).

82 DNA was extracted from 11-28<sup>T</sup> and EPVM<sup>T</sup> using a 'DNA easy tissue Kit'  
83 (Qiagen) following the manufacturer's instructions. The 16S rRNA gene was PCR  
84 amplified using primers 27F and 1492R (Lane, 1991). PCR amplification was performed  
85 following previously reported procedures (Cerritos *et al.*, 2008). The 27F fragment was  
86 used to determine the taxonomic affiliation of the isolates by direct comparison with 16S  
87 rRNA gene sequences from the NCBI database. This procedure was performed using  
88 BLAST (Altschul *et al.*, 1990). In addition, the sequenced 16S rRNA gene fragments  
89 were assembled using Bioedit v7.0.9 (Hall, 1999).

90 Furthermore, four housekeeping genes (*citC*, *recA*, *hsp70* and *rpoB*) were  
91 amplified using previously reported primers specifically designed for *Exiguobacterium*  
92 strains (Rodrigues *et al.*, 2006). The partial sequences of these genes were amplified by  
93 PCR under standard procedures. The annealing temperatures for each primer sets was:  
94 52°C for *citC*, 47°C for *recA* and *hsp70*, and 53°C for *rpoB*. All PCR products were  
95 sequenced at the University of Washington's High Throughput Sequencing Solutions  
96 Laboratory.

97 The 16S rRNA gene sequences from *Exiguobacterium* strains were aligned and  
98 edited using Bioedit (Hall, 1999). Sequences for the majority of the *Exiguobacterium*  
99 species were included in the alignment. The sequence alignment was used to obtain a  
100 ML phylogeny with PhyML v.3.0 program (Fig 1; Guindon & Gascuel, 2003).

101 The 16S rRNA gene phylogeny showed that isolates 11-28<sup>T</sup> and EPVM<sup>T</sup> are  
102 closely related to a subgroup of the *Exiguobacterium* genus that includes *E. mexicanum*  
103 8N<sup>T</sup>, *E. aurantiacum* DSM 6208<sup>T</sup> and the recently reported *E. aquaticum* IMTB-3094<sup>T</sup>. In  
104 order to determine the genetic similarity between each of the two isolates (11-28<sup>T</sup> and  
105 EPVM<sup>T</sup>) and the type species of the genus, a 16S rRNA gene sequence identity matrix  
106 was obtained (Table S1). The type species that have the highest similarity with isolates  
107 11-28<sup>T</sup> and EPVM<sup>T</sup> are *E. mexicanum* 8N<sup>T</sup> (99.9% and 99.7%), *E. aurantiacum* DSM  
108 6208<sup>T</sup> (99.7 and 99.6%) and *E. aquaticum* IMTB-3094<sup>T</sup> (99.5% and 99.3%).

109 In addition to the 16S rRNA gene sequences, partial sequences of four  
110 housekeeping genes (*citC*, *hsp70*, *recA* and *rpoB*) were concatenated and aligned using  
111 Bioedit (Hall, 1999). The concatenated alignment was used to obtain a ML phylogeny  
112 with PhyML v3.0 program (Fig 2). This phylogeny shows that isolates 11-28<sup>T</sup> and  
113 EPVM<sup>T</sup> are clearly differentiated into two distinct lineages.

114 To explore the physiological differences between the new isolates and the more  
115 closely related *Exiguobacterium* type species, the carbon utilisation and acid production  
116 of each of the isolates were measured (Table 1). To perform this analysis, strains were  
117 cultured by triplicates in modified MAM in which only one carbon source was included.  
118 Growth and acid production was evaluated at 24, 48 and 72h after the inoculation. In  
119 addition, GC content, catalase and oxidase reactions were determined by the  
120 BCCM™/LMG consortium. As seen in Table 1, isolates 11-28<sup>T</sup> and EPVM<sup>T</sup> have unique

121 carbon utilisation profiles, which are different from those of closely related type species  
122 profiles.

123 Furthermore, the fatty acid composition of each isolate was characterised (Table  
124 2). For this analysis, cells were grown for 24h at 28°C on MAM. Inoculation and  
125 harvesting of cells, as well as the extraction and analysis were performed according to  
126 the recommendations of the commercial identification system MIDI (Microbial  
127 Identification System, Inc., Delaware U.S.A.). The whole-cell fatty acid composition was  
128 determined by gas chromatography. All fatty acid analyses were performed by the  
129 BCCM™/LMG consortium. As in other *Exiguobacterium* types species, isolates 11-28<sup>T</sup>  
130 and EPVM<sup>T</sup> have a high amount of iso-C<sub>13:0</sub> and anteiso-C<sub>13:0</sub> fatty acids (11.13/25.11%  
131 and 10.81/27.34% respectively). However, the composition of the least common fatty  
132 acids in both isolates was different in comparison to closely related type species.

133 DNA-DNA hybridizations were performed to determine the DNA relatedness  
134 between isolates 11-28<sup>T</sup>, EPVM<sup>T</sup> and *E. mexicanum* 8N<sup>T</sup>. Hybridizations were  
135 performed with 50% formamide at 43°C, according to previous modifications of the  
136 method (Ezaki *et al.*, 1989). The DNA-DNA relatedness percentages represent the  
137 average of at least five hybridizations tests. Hybridizations were standardised and  
138 performed by the BCCM™/LMG consortium. The DNA-DNA relatedness percentage  
139 between isolates 11-28<sup>T</sup> and EPVM<sup>T</sup> was 63%. Additionally, 11-28<sup>T</sup> and EPVM<sup>T</sup> shared  
140 60% and 61% of DNA content with *E. mexicanum* 8N<sup>T</sup> respectively. On the basis of the  
141 DNA-DNA relatedness values, isolates 11-28<sup>T</sup> and EPVM<sup>T</sup> should be considered as  
142 distinct species.

143

144

145 **Description of *Exiguobacterium halophilus* sp. nov.**

146 Description of *Exiguobacterium halophilus* (ha.lo'phi.lus. Gr. n. *hals halos*, salt; Gr. adj.  
147 *philos*, loving; N.L. masc. adj. *halophilus*, salt-loving.).

148

149 Cells are nonmotile, coccoid rods (1.0 x 1.2-1.5 µm), gram-positive, oxidase and  
150 catalase positive. Colonies are bright orange, shiny and circular. Growth occurs after  
151 incubation for 24-48h in MAM between 28 and 37°C. Optimal growth occurs at pH 9,  
152 however growth is observed from pH 7 to pH 9.5. Growth occurs from NaCl 0.11% to  
153 10% (w/v). Carbon source utilisation and acid production are indicated in Table 1. The  
154 inability of this strain to use glucose, trehalose and sucrose as carbon sources, as well  
155 as the ability to use rhamnose, are distinctive characteristics of this species. The DNA  
156 G+C content is 50.80%. The four major fatty acids in the type strain are anteiso-C<sub>13:0</sub>  
157 (25.22%), iso-C<sub>17:1-10c</sub> (14.66%), iso-C<sub>15:0</sub> (14.06%) and iso-C<sub>13:0</sub> (11.13%). This profile  
158 is distinctive of this strain in comparison to previously reported species. The type strain  
159 11-28<sup>T</sup> (=LMG 26555<sup>T</sup>=CECT 7978<sup>T</sup>) was isolated from a hypersaline desiccation pond  
160 at the Rio Mezquites aquatic system in the CCB, Mexico.

161

162 **Description of *Exiguobacterium coahuilensis* sp. nov.**

163 *Exiguobacterium coahuilensis* (co.a.hui.len'sis. N.L. masc. adj. coahuilensis in reference  
164 to Coahuila, the state in Mexico where the type strain was collected).

165

166 Cells are nonmotile, coccoid rods (1.0 x 1.2-1.5 µm), Gram-positive, oxidase weakly  
167 positive and catalase positive. Colonies are bright orange, shiny and circular. Growth  
168 occurs after incubation for 24-48h in MAM between 28 and 37°C. Optimal growth occurs

169 at pH 9, however growth is observed from pH 7 to pH 9.5. Growth occurs from 0.11% to  
170 10% (w/v) NaCl. Carbon source utilisation and acid production are indicated in Table 1.  
171 Specifically, acid production from lactose is a distinctive characteristic of this species.  
172 The DNA G+C content is 51.40%. The four major fatty acids in the type strain are  
173 anteiso-C<sub>13:0</sub> (27.34%), iso-C<sub>17:1-10c</sub> (15.22%), iso-C<sub>15:0</sub> (14.74%) and iso-C<sub>13:0</sub> (10.81%).  
174 This profile is distinctive of this strain in comparison to previously reported species. The  
175 type strain EPVM<sup>T</sup> (=LMG 26554<sup>T</sup> = CECT 7977<sup>T</sup>) was isolated from a desiccation pond  
176 at Los Hundidos aquatic system in the CCB, Mexico.

177

## 178 **ACKNOWLEDGMENTS**

179 We thank Laura Espinosa-Asuar and Dr. Erika Aguirre for technical support during  
180 the development of the project. We also thank Morena Avitia, Rodrigo González-  
181 Chauvet, German Bonilla-Rosso and Santiago Ramírez-Barahona for their support  
182 during field work. This paper is part of the doctoral research of the first author, who  
183 thanks the Doctorado en Ciencias Biomédicas (Universidad Nacional Autónoma de  
184 México) and CONACYT (Grant No.170206) for financial support. The project was  
185 supported by grants 0023459-SEMARNAT/CONACYT and 57507 SEP/CONACyT.

186

## 187 **REFERENCES**

188 **Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990).** Basic local  
189 alignment search tool. *J Mol Biol* **215**, 403-10.

190

191 **Avitia, M., Cerritos R., Caballero-Mellado, J., Martínez-Aguilar, L., Molina-Torres,**  
192 **J., Ramírez-Chávez, E., Gudiño, M., Eguiarte, L.E. & Souza, V. (2012).** *Bacillus*

193 *cuatrocienegasensis* sp. nov., *Bacillus mexicanum* sp. nov., *Bacillus minckleyi* sp. nov.  
194 and *Bacillus churincensis* sp. nov., halophilic alkaliphilic species from the Cuatro  
195 Cienegas Basin in Coahuila, Mexico. *Submitted to Int J Syst Evol Microbiol*.

196

197 **Cerritos, R., Vinuesa, P., Eguiarte, L.E., Herrera-Estrella, L., Alcaraz-Peraza, L.D.,**  
198 **Arvizu-Gomez, J.L., Olmedo, G., Ramirez, E., Siefert, J.L. & other authors. (2008).**  
199 *Bacillus coahuilensis* sp. nov., a moderately halophilic species from a desiccation  
200 lagoon in the Cuatro Cienegas Valley in Coahuila, Mexico. *Int J Syst Evol Microbiol* **58**,  
201 919-923.

202

203 **Chaturvedi, P. & Shivaji, S. (2006).** *Exiguobacterium indicum* sp. nov., a psychrophilic  
204 bacterium from the Hamta glacier of the Himalayan mountain ranges of India. *Int J Syst*  
205 *Evol Microbiol.* **56**, 2765-70.

206

207 **Chaturvedi, P., Prabahar, V., Manorama, R., Pindi, P.K., Bhadra, B., Begum, Z. &**  
208 **Shivaji, S. (2008).** *Exiguobacterium soli* sp. nov., a psychrophilic bacterium from the  
209 McMurdo Dry Valleys, Antarctica. *Int J Syst Evol Microbiol* **58**, 2447-2453.

210

211 **Collins, M.D., Lund, B.M., Farrow, J.A.E. & Schleifer, K.H. (1983).** Chemotaxonomic  
212 study of an alkalophilic bacterium, *Exiguobacterium aurantiacum* gen. Nov., sp. nov. *J*  
213 *Gen Microbiol.* **129**, 2037-2042.

214

215 **Crapart, S., Fardeau, M.L., Cayol, J.L., Thomas, P., Sery, C., Ollivier, B. & Combet-**  
216 **Blanc, Y. (2007).** *Exiguobacterium profundum* sp. nov., a moderately thermophilic,

217 lactic acid-producing bacterium isolated from a deep-sea hydrothermal vent. *Int J Syst*  
218 *Evol Microbiol.* **57**, 287-92.

219

220 **Escalante, A.E., Caballero-Mellado, J., Martinez-Aguilar, L., Rodriguez-Verdugo,**  
221 **A., Gonzalez-Gonzalez A., Toribio-Jimenez, J. & Souza V. (2009).** *Pseudomonas*  
222 *cuatrocieneegasensis* sp. nov., isolated from an evaporating lagoon in the Cuatro  
223 Cienegas valley in Coahuila, Mexico. *Int J Syst Evol Microbiol* **59**, 1416-1420.

224

225 **Ezaki, T., Hashimoto, Y., & Yabuuchi, E. (1989).** Fluorometric deoxyribonucleic acid-  
226 deoxyribonucleic acid hybridization in microdilution wells as an alternative to membrane  
227 filter hybridization in which radioisotopes are used to determine genetic relatedness  
228 among bacterial strains. *Int. J. Syst. Bacteriol.* **39**, 224-229

229

230 **Farrow, J.A., Wallbanks, E.S. & Collins, M.D. (1994).** Phylogenetic interrelationships  
231 of round spore-forming bacilli containing cell walls based on lysine and the non-spore-  
232 forming genera *Caryophanon*, *Exiguobacterium*, *Kurthia* and *Planococcus*. *Int J Syst*  
233 *Bacteriol* **44**, 74-82.

234

235 **Fruhling, A., Schumann, P., Hippe, H., Straubler, B. & Stackebrandt, E. (2002).**  
236 *Exiguobacterium undae* sp. nov. and *Exiguobacterium antarcticum* sp. nov. *Int J Syst*  
237 *Evol Microbiol* **52**, 1171-1176.

238

239 **Guindon, S. & Gascuel, O. (2003).** A simple, fast, and accurate algorithm to estimate  
240 large phylogenies by maximum likelihood. *Syst Biol* **52**, 696-704.



241

242 **Hall, T.A. (1999).** bioedit: a user-friendly biological sequence alignment editor and  
243 analysis program for Windows 95/98/NT. *Nucl Acid Symp series* 95–98.

244

245 **Kim, I., Lee, M., Jung, S., Song, J.J., Oh, T. & Yoon, J. (2005).** *Exiguobacterium*  
246 *aestuarii* sp. nov. and *Exiguobacterium marinum* sp. nov. isolated from tidal flat of the  
247 Yellow Sea in Korea. *Int J Syst Evol Microbiol.* **55**, 885-889.

248

249 **Lane, D.J. (1991).** 16S/23S rRNA sequencing. In: *Nucleic Acid Techniques in Bacterial*  
250 *Systematics*, pp. 115–175. Edited by Stackebrandt, E. & Goodfellow, M. Ed Wiley:  
251 Chichester, UK.

252

253 **López-Cortés, A., Schumann, P., Pukall, R. & Stackebrandt, E. (2006).**  
254 *Exiguobacterium mexicanum* sp. nov. and *Exiguobacterium artemiae* sp. nov. isolated  
255 from the brine shrimp *Artemia franciscana*. *Syst Appl Microbiol* **29**, 183-190.

256

257 **Posada, D. (2008).** jModelTest: phylogenetic model averaging. *Mol Biol Evol* **25**, 1253-6.

258

259 **Raichand, R., Pareek, S., Singh, N.K. & Mayilraj, S. (2011).** *Exiguobacterium*  
260 *aquaticum* sp. nov., a new member of the genus *Exiguobacterium*. *Int J Syst Evol*  
261 *Microbiol*, Nov 4 [Epub ahead of print].

262

263 **Rebollar, E.A., Avitia, M., Eguiarte, L.E., González-González, A., Mora, L., Bonilla-**  
264 **Rosso, G. & Souza, V. (2012).** Water-sediment niche differentiation in ancient marine

265 lineages of *Exiguobacterium* endemic to the Cuatro Ciénegas Basin. *Environ Microbiol.*  
266 doi: 10.1111/j.1462-2920.2012.02784.x. [Epub ahead of print].

267

268 **Rodrigues, D.F., Goris, J., Vishnivetskaya, T., Gilichinsky, D., Thomashow, M.F. &**  
269 **Tiedje, J.M. (2006).** Characterization of *Exiguobacterium* isolates from the Siberian  
270 permafrost. Description of *Exiguobacterium sibiricum* sp. nov. *Extremophiles*. **10**, 285-  
271 94.

272

273 **Rozas, J. & Rozas, R. (1999).** DnaSP version 3: an integrated program for molecular  
274 population genetics and molecular evolution analysis. *Bioinformatics* **15**, 174-5.

275

276 **Souza, V., Espinosa-Asuar, L., Escalante, A.E., Eguiarte, L.E., Farmer, J., Forney,**  
277 **L., Lloret, L., Rodríguez-Martínez, J.M., Soberón, X. & other authors. (2006).** An  
278 endangered oasis of aquatic microbial biodiversity in the Chihuahuan desert. *Proc Natl*  
279 *Acad Sci U S A*. **103**, 6565-70.

280

281 **Yumoto, I., Hishinuma-Narisawa, M., Hirota, K., Shingyo, T., Takebe, F., Nodasaka,**  
282 **Y., Matsuyama, H. & Hara, I. (2004).** *Exiguobacterium oxidotolerans* sp. nov. a novel  
283 alkaliphile exhibiting high catalase activity. *Int J Syst Evol Microbiol*. **54**, 2013-2017

**TABLES**

	1	2	3	4	5	6
Oxidase	+	W	-	+	-	-
DNA G+C %	50,80	51,40	ND	53,2-55,8	50,4	53,2

**Utilisation of different carbon sources**

Mannose	-	+	-	-	+	-
Raffinose	+	+	-	-	-	-
Xylose	+	-	+	-	-	-
Glucose	-	+	+	+	+	+
Arabinose	+	+	-	-	-	+
Dulcitol	+	+	-	-	-	-
Fructose	-	-	+	+	+	+
Sorbitol	+	+	-	-	-	-
Salicine	-	-	+	+	+	-
Mannitol	-	-	+	+	+	-
Rhamnose	+	-	-	-	-	-
Trehalose	-	+	+	+	+	+
Malic acid	+	-	ND	ND	ND	-
Sucrose	-	+	+	+	+	+
Maltose	-	-	+	+	+	+
Starch	+	-	ND	ND	ND	ND
Lactose	+	-	-	-	-	+
Glycerol	-	+	+	W	-	+

**Acid production from**

Mannose	+	+	-	-	+	-
Raffinose	-	-	-	-	-	ND
Xylose	-	+	+	-	-	ND
Glucose	+	+	+	+	+	ND
Dulcitol	-	-	-	-	-	ND
Sorbitol	-	-	-	-	-	+
Salicine	+	+	+	+	+	-
Rhamnose	-	-	-	-	-	ND
Malic acid	-	-	ND	ND	ND	ND
Maltose	+	+	+	+	+	ND
Starch	-	-	ND	ND	ND	ND
Lactose	-	+	-	-	-	-
Glycerol	+	+	W	+	-	ND

**Growth with different NaCl concentrations Molarity(%)**

0.2M (0.11%)	+	+	ND	ND	+	+
0.85M (5%)	+	+	ND	ND	+	+
1.7M (10%)	+	+	ND	ND	+	+
2.5M (15%)	-	-	ND	ND	-	-
4M (23%)	-	-	ND	ND	-	-
5.2M (30%)	-	-	ND	ND	-	-

**Growth with different pH**

pH 4.5	-	-	ND	ND	-	-
pH 7	+	+	ND	ND	+	+
pH 9	+	+	ND	ND	+	+

**Table 1**

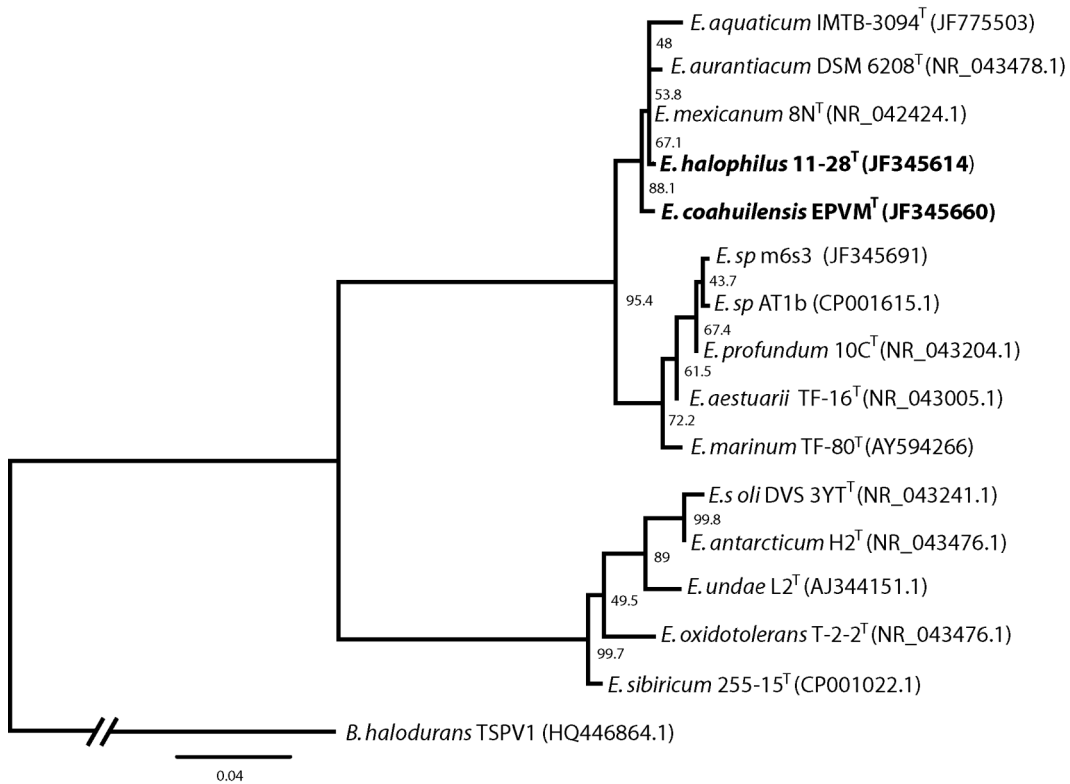
Differential characteristics between strains 11-28<sup>T</sup> and EPVM<sup>T</sup> and the most related type strains of members of *Exiguobacterium*. Strains: 1, strain 11-28<sup>T</sup>; 2, strain EPVM<sup>T</sup>; 3, *E. mexicanum* DSM 16483<sup>T</sup> (López-Cortés *et al.*, 2006); 4, *E. aurantiacum* DSM 6208<sup>T</sup> (Collins *et al.*, 1983; Fruhling, 2002); 5, *E. profundum* DSM 17289<sup>T</sup> (Crapart *et al.*, 2007); 6, *E. aquaticum* JCM 17977<sup>T</sup> (Raichand *et al.*, 2011). All strains were catalase-positive. All strains were positive for acid production from fructose, mannitol, trehalose and sucrose. All strains were negative for acid production from arabinose. W, Weak; ND, no data available.

	1	2	3	4	5	6
<b>Straight chain</b>						
C <sub>12:0</sub>	-	-	8,30	2,00	-	ND
C <sub>14:0</sub>	-	-	6,10	3,00	0,40	ND
C <sub>16:0</sub>	1,01	0,60	<b>32,80</b>	<b>27,00</b>	3,10	5,1
C <sub>17:0</sub>	-	-	ND	ND	ND	1,00
C <sub>18:0</sub>	-	-	7,00	5,00	ND	1,50
<b>Branched</b>						
iso-C <sub>11:0</sub>	0,50	0,59	1,50	2,00	-	ND
iso-C <sub>12:0</sub>	6,08	6,62	2,10	3,00	1,90	3,1
iso-C <sub>13:0</sub>	<b>11,13</b>	<b>10,81</b>	<b>11,20</b>	<b>18,00</b>	<b>13,30</b>	<b>11,9</b>
anteiso-C <sub>13:0</sub>	<b>25,22</b>	<b>27,34</b>	<b>8,90</b>	<b>12,00</b>	<b>16,10</b>	<b>19</b>
iso-C <sub>14:0</sub>	0,94	0,95	-	-	1,40	0,8
iso-C <sub>15:0</sub>	<b>14,06</b>	<b>14,74</b>	1,70	4,00	<b>17,80</b>	9,8
anteiso-C <sub>15:0</sub>	2,14	2,11	-	-	3,50	2
iso-C <sub>16:0</sub>	2,00	0,92	-	-	4,40	2,7
iso-C <sub>17:0</sub>	6,04	3,54	-	-	<b>15,40</b>	<b>16,1</b>
anteiso-C <sub>17:0</sub>	2,88	2,24	-	6,00	6,30	4,5
iso-C <sub>18:0</sub>	-	-	ND	ND	0,60	ND
<b>Unsaturated</b>						
C <sub>16:1-7c</sub> alcohol	5,55	6,10	6,50	-	3,70	ND
C <sub>16:1-11c</sub>	4,05	3,87	<b>10,30</b>	<b>10,00</b>	2,90	ND
iso-C <sub>17:1-10c</sub>	<b>14,66</b>	<b>15,22</b>	ND	ND	7,30	<b>12,7</b>
isol/antel B C <sub>17:1</sub>	3,74	3,76	2,80	ND	1,80	2,2
C <sub>18:1-9c</sub>	-	-	-	2,00	-	ND

**Table 2**

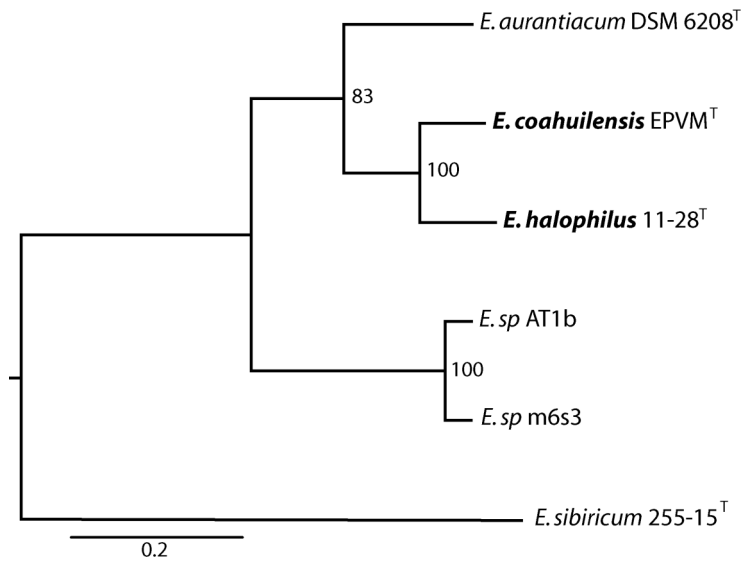
Fatty acid composition (%) of strains 11-28<sup>T</sup> and EPVM<sup>T</sup> and the most related type strains of members of *Exiguobacterium*. Strains: 1, strain 11-28<sup>T</sup>; 2, strain EPVM<sup>T</sup>; 3, *E. mexicanum* DSM 16483<sup>T</sup> (López-Cortés *et al.*, 2006); 4, *E. aurantiacum* DSM 6208<sup>T</sup> (Collins *et al.*, 1983; Fruhling, 2002); 5, *E. profundum* DSM 17289<sup>T</sup> (Crapart *et al.*, 2007); 6, *E. aquaticum* JCM 17977<sup>T</sup> (Raichand *et al.*, 2011).

## FIGURES



**Figure 1**

Phylogenetic reconstruction based on the 16S rRNA gene. A model test analysis (Posada, 2008) was used to obtain a specific substitution model for this gene: TIM1+I+ $\gamma$ . The sequence alignment of 1302 bp was used to obtain a phylogenetic reconstruction by a Maximum Likelihood (ML) approach using PhyML v.3.0 with 1000 bootstrap iterations.



**Figure 2**

Phylogenetic reconstruction based on the concatenated sequences of *citC*, *hsp70*, *recA* and *rpoB* genes. Jmodel test (Posada, 2008) was used to obtain a specific DNA substitution model for the concatenated alignment: GTR+I+ $\gamma$ . Phylogenetic relationships were reconstructed using PhyML v.3.0 with 1000 bootstrap iterations. Two reference *Exiguobacterium* type strains and two additional isolates were included in the phylogeny.

	<b>Species</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>
<b>1</b>	<i>E. halophilus</i> E11-28 <sup>T</sup>	100														
<b>2</b>	<i>E. coahuilensis</i> EPVM <sup>T</sup>	99.7	100													
<b>3</b>	<i>E. mexicanum</i> 8N <sup>T</sup>	99.9	99.7	100												
<b>4</b>	<i>E. aurantiacum</i> DSM 6208 <sup>T</sup>	99.7	99.6	99.8	100											
<b>5</b>	<i>E. aquaticum</i> IMTB-3094 <sup>T</sup>	99.5	99.3	99.6	99.4	100										
<b>6</b>	<i>E. sp</i> m6s3	98.7	98.9	98.7	98.5	98.5	100									
<b>7</b>	<i>E. sp</i> AT1b	98.9	99.1	98.9	98.7	98.5	99.8	100								
<b>8</b>	<i>E. profundum</i> 10C <sup>T</sup>	98.8	98.8	98.9	98.7	98.6	99.6	99.6	100							
<b>9</b>	<i>E. aestuarii</i> TF-16 <sup>T</sup>	98.7	98.9	98.7	98.5	98.3	99.8	99.8	99.5	100						
<b>10</b>	<i>E. marinum</i> TF-80 <sup>T</sup>	98.7	98.7	98.8	98.6	98.5	99.4	99.2	99.4	99.2	100					
<b>11</b>	<i>E. soli</i> DVS 3YT <sup>T</sup>	94.0	94.1	94.0	93.9	93.6	93.9	93.9	93.7	93.8	93.9	100				
<b>12</b>	<i>E. antarcticum</i> H2 <sup>T</sup>	94.3	94.4	94.3	94.2	93.9	94.1	94.2	94.0	94.1	94.2	99.6	100			
<b>13</b>	<i>E. undae</i> L2 <sup>T</sup>	94.2	94.3	94.2	94.1	93.9	94.1	94.2	94.0	94.1	94.2	98.8	99.1	100		
<b>14</b>	<i>E. oxidotolerans</i> T-2-2 <sup>T</sup>	94.6	94.7	94.6	94.5	94.2	94.5	94.6	94.4	94.5	94.6	98.5	98.8	98.5	100	
<b>15</b>	<i>E. sibiricum</i> 255-15 <sup>T</sup>	94.9	94.9	94.9	94.7	94.6	94.9	94.8	94.8	94.7	95.0	98.4	98.7	98.9	99.0	100
<b>16</b>	<i>B. halodurans</i> TSPV1	89.8	89.8	89.9	90.1	89.8	90.0	90.1	90.1	90.0	90.2	89.7	90.0	89.7	90.2	90.1

**Table S1.** Similarity percentage between *Exiguobacterium* species based on the *16srRNA* gene sequences. The similarity matrix was performed using DNAsp 5.10.01 (Rozas & Rozas, 1999).

## 7. DISCUSIÓN Y CONCLUSIONES

Para poder entender los mecanismos involucrados en la especiación bacteriana es necesario abordar el fenómeno desde diferentes ángulos. Por un lado, es esencial ampliar los análisis genéticos de poblaciones procariontes ambientales, ya que la mayor parte de la información multilocus y de genomas completos se enfoca en el análisis de bacterias patógenas. Además del estudio de la diversidad genética de las poblaciones naturales, es importante profundizar en los fenómenos de adaptación local. Para esto, es fundamental determinar el papel del ambiente y la geografía en la evolución de los linajes procariontes.

Tomando en cuenta las consideraciones anteriores, este trabajo tuvo como motivación, ampliar los estudios evolutivos en poblaciones bacterianas naturales en el valle de CCC. En primer lugar, se describieron tres linajes dentro del género *Exiguobacterium*, que muestran pocas (pero importantes) evidencias de recombinación homóloga. En segundo lugar, se abordó el papel de las barreras geográficas y ecológicas en la divergencia genética entre poblaciones bacterianas filogenéticamente cercanas.

### **7.1 Los linajes de *Exiguobacterium* en CCC son cercanos a especies de origen marino.**

El análisis filogenético realizado en este proyecto nos permitió determinar que los linajes de *Exiguobacterium* presentes en CCC son filogenéticamente cercanos a las especies marinas del género. Como se mencionó anteriormente, este hallazgo es congruente con trabajos previos en donde la mayor parte de los linajes bacterianos en el valle tienen una afiliación con linajes marinos (Souza *et al.*, 2006). Esta observación, es particularmente interesante en el caso de CCC, ya que este sitio era parte del suelo marino hace más de 150 millones de años. ¿Porqué después de tanto tiempo, los linajes bacterianos son más



cercanos a especies marinas y no a especies de agua continental?. El análisis de salinidad de los principales sistemas acuáticos, mostró que las proporciones iónicas en el agua y sedimento de CCC son muy parecidas a las proporciones del agua de mar (Rebollar *et al.*, 2012). Además de la salinidad, CCC presenta los niveles más bajos de fósforo reportados para aguas continentales, asemejándose a la mayor parte de los ambientes oligotróficos marinos (Elser *et al.*, 2005). Es posible que estas condiciones ambientales, semejantes a los ambientes marinos, permitan el desarrollo y la diversificación de linajes bacterianos con adaptaciones semejantes a las marinas. Tal es el caso de *B. coahuilensis*, el cual sustituye los fosfolípidos de membrana por sulfolípidos (Alcaráz *et al.*, 2008). Esta y otras adaptaciones a la oligotrofia han sido reportadas solamente en bacterias del fitoplancton marino (Van Mooy *et al.*, 2009). Los mecanismos adaptativos que han favorecido la existencia y evolución de estos linajes bacterianos en CCC, se podrán abordar en un futuro cercano a partir del análisis comparativo de genomas bacterianos originarios de este valle (ver **Capítulo 8: Perspectivas**).

Otro de los aspectos interesantes con respecto a la afiliación marina de *Exiguobacterium* en CCC es lo encontrado en el análisis AdaptML (**Capítulo 4**). En este análisis se obtuvieron cinco hábitats con condiciones ambientales contrastantes. Estos hábitats se localizan en distintos puntos de la filogenia de *Exiguobacterium*, y por lo tanto es posible inferir la historia del surgimiento de cada hábitat. Bajo este escenario habrá hábitats ancestrales que dieron lugar a los hábitats más recientes. Estos últimos corresponden a aislados localizados en las ramas más recientes de la filogenia. Interesantemente, el hábitat que se localiza en los nodos basales de la filogenia es aquel que coincide con las condiciones marinas (agua con salinidad equivalente al agua de mar). Bajo este escenario, F2 es el linaje que ha mantenido las condiciones del hábitat ancestral “marino”, mientras que F1 y F3 muestran grupos genéticos recientes asociados a nuevos hábitats con condiciones ambientales diferentes (**Capítulo 4**,

**Figura 3).** Con base en lo anterior, es posible sugerir que la divergencia genética observada al interior de los filogrupos de *Exiguobacterium* se encuentra estrechamente asociada a la exploración de nuevos nichos ecológicos.

## **7.2 Los linajes de *Exiguobacterium* en CCC indican historias evolutivas diferentes**

Uno de los principales objetivos de este trabajo fue describir la diversidad genética, así como el papel de la recombinación en la estructura poblacional de *Exiguobacterium*. A partir de los análisis filogenéticos y de genética de poblaciones clásica, fue posible detectar diferencias claras en la diversidad nucleotídica entre los tres linajes principales (F1, F2, F3). En particular, la prueba de Tajima indicó que F1 se encuentra en un proceso de expansión demográfica, mientras que en la filogenia, F2 y F3 están constituidos por ramas más profundas y divergentes entre si. Lo anterior sugiere que cada uno de estos grupos es un linaje genético independiente con historias evolutivas independientes sujetos a procesos demográficos contrastantes.

Con base en todo lo anterior, se realizaron diferentes pruebas fenotípicas para describir cada uno de estos linajes como especies distintas (unidades evolutivas independientes). En particular, nos enfocamos en describir dos cepas provenientes de F1 y F2 respectivamente denominados 11-28 y EPVM (**Capítulo 6**). A partir de estos análisis se concluyó que las cepas 11-28 y EPVM son efectivamente especies distintas. Sin embargo, con el fin de profundizar en los perfiles genotípicos y fenotípicos de estas dos cepas, es necesario realizar análisis adicionales que describan a mayor profundidad las diferencias entre estos dos linajes. Con respecto al genotipo, será necesario realizar análisis de BOX-PCR (“BOX repetitive elements PCR”) que describan los rearrreglos genómicos en cada una de las cepas, así como en las cepas más cercanas

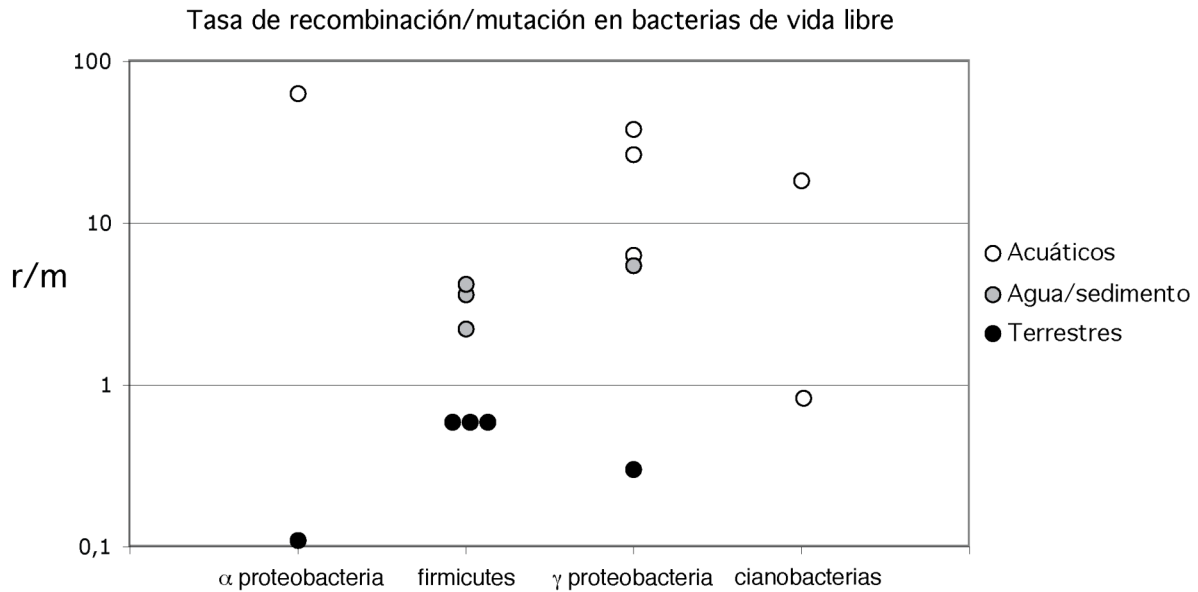
filogenéticamente a la especie tipo. Con respecto a los análisis fenotípicos, será necesario describir el perfil de proteínas, así como la presencia de lípidos polares y quinonas en cada una de las especies tipo (Kampfer *et al.*, 2003; Tindall *et al.*, 2010).

Una gran cantidad de estudios recientes, sugieren un escenario en donde la recombinación homóloga y la transferencia horizontal de genes tienen un papel determinante en movilizar información dentro y entre las poblaciones procariontes. Estos mecanismos de flujo genético, son en gran parte los responsables de la diversidad genética que se ha encontrado en la naturaleza y que se reflejan en la gran variedad de ambientes habitados por procariontes. Asimismo, la recombinación homóloga, según las evidencias más recientes, funciona como una fuerza cohesiva dentro de las especies (Shapiro *et al.*, 2012).

En este contexto y con el fin de profundizar en los mecanismos involucrados en la divergencia genética entre poblaciones, se realizaron diversas pruebas para evaluar la recombinación homóloga entre y dentro de los linajes de *Exiguobacterium*. Los análisis de recombinación homóloga obtenidos en este trabajo, indican que la frecuencia de la recombinación ha sido baja y no ha sido suficiente para romper el desequilibrio de ligamiento entre los alelos. Sin embargo, resultados adicionales con ClonalFrame nos permitieron determinar que una proporción importante de los polimorfismos asociados a cada uno de los filogrupos están explicados por recombinación y no por mutación.

Todo lo anterior, permite concluir que la recombinación homóloga, a pesar de ser un mecanismo poco frecuente, es uno de los mecanismos importantes en la divergencia genética de estos linajes bacterianos y por ende en la historia evolutiva de *Exiguobacterium*. Los trabajos recientes sobre el papel de la recombinación en poblaciones bacterianas y arqueanas muestra una gran variabilidad con respecto al estilo

de vida y la tasa de recombinación/mutación ( $r/m$ ), sin embargo, algunos patrones ecológicos han podido distinguirse (Vos y Didelot, 2009). Sin considerar a los organismos patógenos, comensales y endosimbiontes, es interesante observar que los organismos de vida libre acuáticos (terrestres y marinos) tienen tasas de  $r/m$  mucho más altas que los organismos terrestres (Figura 15).



**Figura 15**

Valores de  $r/m$  en diferentes linajes bacterianos con distintos estilos de vida. El eje X indica los valores de  $r/m$  en escala logarítmica. El eje Y indica el grupo filogenético al que pertenecen cada una de las especies analizada. (Datos tomados de Vos y Didelot, 2009, Keymer y Boehm 2011 y Rebollar *et al.*, 2012).

Lo anterior sugiere, que independientemente del linaje filogenético, los organismos que se desarrollan en ambientes contrastantes (terrestre vs acuático) tienden a tener tasas de  $r/m$  muy diferentes. El trabajo realizado en *Exiguobacterium* contribuye al conocimiento de los organismos que se desarrollan en ambientes acuáticos pero también en los sedimentos y tapetes por debajo del agua. Junto con datos recientes de vibrios marinos que viven en agua y sedimento (Keymer y Boehr, 2011), los tres linajes de *Exiguobacterium* tienen tasas de  $r/m$  intermedias entre los organismos acuáticos y los terrestres (ver Figura 15, círculos grises en el grupo de los firmicutes). Para determinar

si esta tendencia indica un patrón asociado al estilo de vida es necesario ampliar los datos de bacterias ambientales y profundizar en el papel de la recombinación al interior de estas poblaciones.

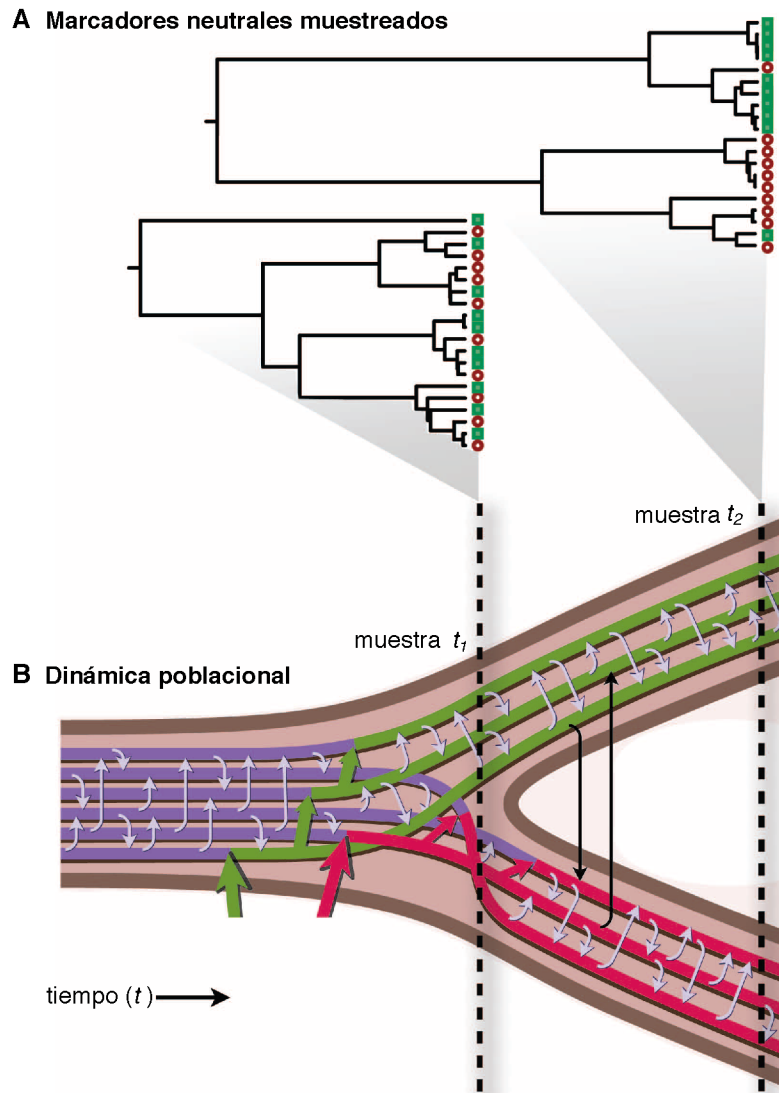
### **7.3 Los linajes de *Exiguobacterium* en CCC muestran un patrón de diferenciación ecológica asociada a los ambientes de agua y sedimento.**

Tomando en cuenta los trabajos previos a este trabajo, es claro que el aislamiento geográfico y las condiciones ambientales han permitido la estructuración de las poblaciones microbianas y han fomentado la evolución divergente de numerosos grupos procariontes (**Capítulo 1**). En particular, los pocos trabajos que existen a la fecha, sugieren que las condiciones ecológicas locales han aislado las poblaciones bacterianas mediante la exploración de diferentes nichos ecológicos y la presión selectiva de los diferentes ambientes. El objetivo principal de este proyecto fue el evaluar los parámetros (geográficos o ambientales) involucrados en la estructuración de las poblaciones de *Exiguobacterium* en el valle de CCC.

Con base en lo obtenido en este proyecto, fue posible definir la existencia de grupos genéticos al interior de cada linaje que están asociados a dos ambientes contrastantes: agua y sedimento. Este resultado contribuye al conocimiento de los factores ecológicos involucrados en la divergencia de las poblaciones bacterianas. Así mismo, confirma la observación de que a escalas locales, los factores ambientales (bióticos y abióticos) influyen en la estructura poblacional, mientras que el aislamiento geográfico parece influir sólo a escalas de distancia mayores (miles de kilómetros o entre continentes).

Volviendo a los linajes de *Exiguobacterium* en CCC, es claro que los individuos que conforman cada linaje tienen un porcentaje muy alto de similitud al comparar las secuencias de los cinco genes analizados. A pesar de la alta similitud en los genes analizados fue posible detectar una sub-estructura al interior de cada linaje. Una de las preguntas que se desprenden de estos resultados, es sobre el significado de la existencia de estos cluster al interior de cada linaje (F1, F2 y F3). ¿Estamos realmente observando eventos de divergencia determinados por las condiciones ecológicas de dos ambientes contrastantes? ¿La presencia de estos pequeños grupos representa el inicio de la diferenciación genética de estos grupos?.

El trabajo reciente, realizado por el grupo de Eric Alm, ha dado luz al respecto de estas pregunta en un modelo de vibrios marinos. El análisis multilocus de 1000 aislados de vibrios mostró eventos de diferenciación de nicho asociado a partículas planctónicas de diferentes tamaños (Hunt *et al.*, 2008). Más adelante se llevó a cabo la secuenciación de 20 genomas que pertenecen a los dos nichos principales y que tienen un 99% de similitud con el *16srRNA*. A partir de esta última aproximación se comprobó que la estructura de estos grupos está asociado a un proceso de diferenciación ecológica reciente, en donde existen genes únicos para cada nicho así como altas tasas de recombinación entre los individuos de un mismo nicho (Shapiro *et al.*, 2012). Estos resultados comprueban que incluso al interior de linajes filogenéticamente cercanos existen procesos de diferenciación ecológica aislados genéticamente entre sí, y que pueden llevar a la eventual divergencia evolutiva (Figura 16).



**Figura 16.**

Diferenciación ecológica de poblaciones bacterianas recombinantes a dos tiempos evolutivos distintos. **(a)** Genealogía de marcadores neutrales mostrando el patrón de diferenciación genética en dos tiempos evolutivos. **(b)** Modelo de diferenciación ecológica. Las flechas grises representan eventos de recombinación al interior de cada grupo en proceso de divergencia. Flechas rojas y verdes representan eventos de adquisición de genes relevantes en el proceso de adaptación ecológica. (Modificada de Shapiro *et al.*, 2012).

El modelo de diferenciación ecológica planteada por Shapiro y colaboradores concuerda perfectamente con la interpretación de los datos obtenidos en *Exiguobacterium*. Sin embargo para comprobar esta propuesta al interior de los linajes de *Exiguobacterium* en CCC, es necesario realizar análisis genómicos que nos permitan

detectar posibles firmas genéticas asociadas a la adaptación a diferentes nichos ecológicos (agua o sedimento). En el **capítulo 8** de esta tesis se detallarán las perspectivas de este proyecto, dentro de las cuáles se realizará un proyecto de genómica comparada entre bacterias de sedimento y agua en el valle de CCC.

Como se mencionó al inicio de esta sección, es esencial profundizar en el estudio de poblaciones naturales bacterianas si queremos realmente entender cómo es que evolucionan los linajes procariontes. Las comunidades bacterianas de CCC representan una oportunidad única para abordar preguntas ecológicas y evolutivas en procariontes. Por un lado la particularidad de este sitio nos permite entender los procesos de adaptación a ambientes restrictivos y extremos desde el punto de vista de la salinidad y la disponibilidad de recursos. Por otro lado, el estudio de *Exiguobacterium* en el valle, nos permitió comenzar a describir la biología de organismos poco conocidos y que probablemente jueguen un papel importante en la dinámica ecológica de las comunidades bacterianas. Finalmente, la abundancia y diversidad de linajes bacterianos en los sistemas acuáticos nos permitirá abordar preguntas a cerca de los mecanismos globales que subyacen a la especiación bacteriana y nos permitan explicar el porqué de la gran diversidad procariontes en nuestro planeta.



## 8. PERSPECTIVAS

El trabajo realizado en esta tesis de doctorado es en realidad el inicio para comenzar a describir a profundidad la evolución de grupos bacterianos abundantes en CCC, como es el caso del género *Exiguobacterium*. En este capítulo de la tesis se detallarán las principales líneas de trabajo que se desprenden de este proyecto y que ya están siendo abordadas en el laboratorio.

### 8.1 Caracterización genotípica y fenotípica de los linajes de *Exiguobacterium* en CCC.

En esta tesis se describió la genealogía general de los grupos de *Exiguobacterium* (Capítulo 4) y se realizaron diversas pruebas fenotípicas y genéticas para describir cada uno de los linajes (Capítulo 5). Sin embargo, consideramos importante profundizar en la caracterización de cada uno de estos linajes con dos objetivos principales: el primero, definir la plasticidad fenotípica y genotípica de F1, F2 y F3 y correlacionar estos patrones con el origen de cada uno de los aislados. El segundo objetivo es describir formalmente a estos grupos como especies diferentes desde el punto de vista genético y fenotípico.

El primer objetivo será abordado a partir de la descripción de niveles de diversidad distintos. Por un lado se llevará a cabo la genotipificación de cada uno de los aislados por medio de BOX-PCR. Esta aproximación permite describir la plasticidad genómica y comparar los patrones con la filogenia obtenida previamente (Rebollar *et al.*, 2012). Asimismo, será posible contestar preguntas como ¿Los linajes de sedimento son más o menos diversos que los de agua? ¿Existen diferencias entre los patrones de cepas provenientes de ambientes distintos? o ¿Los patrones genómicos correlacionan con los

linajes filogenéticos?. Estos patrones genómicos están dados por la dinámica de los elementos repetitivos a través del genoma y por lo tanto tienen en muchos casos, mayor resolución que los análisis de genes constitutivos conservados (Whitaker, 2006). Incluso, los análisis de genotipificación han permitido discernir patrones genéticos asociados al aislamiento geográfico y/o a condiciones ambientales particulares (Cho y Tiedje, 2000; Sikorski y Nevo 2005). Además de evaluar los rearrreglos genómicos, se determinará la amplitud de nicho de cada uno de los aislados de *Exiguobacterium*. Para esto se realizarán curvas de crecimiento utilizando un gradiente de varios parámetros como el pH, la temperatura y la salinidad. Esta aproximación nos permitirá definir el nicho, en términos de las condiciones abióticas en las que cada uno de estos aislados se desarrolla. Esta aproximación nos permitirá determinar si la amplitud de nicho está relacionada con el ambiente de donde provienen, con la plasticidad genotípica o con la afiliación filogenética de cada uno de los aislados. En particular, será interesante evaluar si los aislados de sedimento tienen un nicho distinto a los de agua. Por otra parte podremos evaluar si los aislados de ambientes hipersalinos están adaptados a estos ambientes o si se pueden desarrollar también en ambientes salinos y viceversa.

Los análisis previos utilizando toda la muestra de *Exiguobacterium*, serán esenciales para concluir el segundo objetivo, que es la descripción de estos linajes como especies nuevas dentro del género. Para completar el análisis fenotípico de estos linajes, será necesario realizar perfiles de proteínas, así como la cuantificación de lípidos polares y quinonas, según los requisitos establecidos por los taxónomos de procariontes (Kampfer *et al.*, 2003; Tindall *et al.*, 2010).

## 8.2 Genómica comparada de *Exiguobacterium*.

Con base en los resultados obtenidos en esta tesis, se desprende la pregunta de si los linajes asociados a sedimento o agua, realmente representan grupos genéticos en proceso de diferenciación ecológica. Las curvas de crecimiento propuestos en la **sección 8.1**, nos ayudarán a determinar si los individuos de sedimento tienen condiciones de crecimiento óptimas diferentes a los individuos de agua. La aproximación complementaria a la caracterización experimental del nicho, es justamente profundizar en los rasgos genéticos particulares de los aislados de sedimento y agua.

Con base en lo anterior se secuenciaron los genomas de dos cepas de *Exiguobacterium*: uno proveniente de agua de RM y otro proveniente de sedimento de LH (11-28 y EPVM respectivamente). Como se describió en el **capítulo 4**, estas cepas provienen de dos ambientes contrastantes en términos de la composición iónica y se encuentran en proceso de ser descritas como especies nuevas dentro del género. El objetivo principal de la comparación de estos dos genomas es determinar si existen genes únicos a cada cepa que pudieran dar información acerca de la adaptación de estos organismos a sus respectivos ambientes.

El genoma de la cepa 11-28 ya se encuentra ensamblado y anotado, y por lo tanto se están comenzando a analizar las vías metabólicas presentes en este organismo. Por otra lado, el genoma de la cepa EPVM se encuentra en proceso de ser ensamblado para su posterior anotación.

De manera complementaria a la descripción de los dos genomas de *Exiguobacterium* en CCC, será importante caracterizar la abundancia de nutrientes, así como la abundancia bacteriana en el agua y el sedimento de donde provienen cada una de las cepas. De esta

manera, tendremos más información acerca de las condiciones ecológicas en las que se desarrollan cada una de las cepas estudiadas.

Por otra parte, se llevó a cabo la secuenciación de otra cepa de *Exiguobacterium* (N39) proveniente de una salina andina localizada a 4,400 mts sobre el nivel del mar y con altas concentraciones de sal (Ordoñez *et al.*, 2009). Esta cepa es filogenéticamente cercana a las dos cepas secuenciadas de CCC, y por lo tanto consideramos interesante incluirla como parte del análisis comparativo. Las tres cepas (11-28, EPVM y N39) se desarrollan en ambientes salinos sin embargo provienen de lugares y condiciones contrastantes. El genoma de la cepa N39 ya se encuentra ensamblado y anotado. Actualmente se está evaluando la calidad de la anotación para así comenzar con la descripción de las vías metabólicas presentes en este organismo.

Con el fin de analizar los genomas de *Exiguobacterium* en el contexto de otras especies dentro del género, se llevará a cabo la comparación con los genomas disponibles a la fecha. Como se mencionó en los antecedentes, existen dos genomas secuenciados del género *Exiguobacterium*, uno de los cuáles es psicrófilo (*E. sibiricum* 255-15; <http://genome.ornl.gov/microbial/exig/>) y el otro es termófilo aislado de las lagunas de Yellowstone (*E. sp AT1b*; Vishnivetskaya *et al.*, 2011).

De estas cinco cepas (255-15, AT1b, N39, 11-28 y EPVM), 255-15 se localiza en el *cluster I* de la filogenia del género, mientras que el resto se localizan en el *cluster II* (ver Figura 9 **Capítulo 2**). La comparación de estos cinco genomas nos permitirá evaluar no sólo el contenido genético entre las dos cepas de CCC, sino que además nos permitirá evaluar si existen rasgos únicos asociados a los diferentes estilos de vida.

Asimismo, la comparación de los cinco genomas de *Exiguobacterium*, nos permitirá por un lado definir el “core” genómico de este género. Por otro lado, nos permitirá determinar los genes únicos presentes en los aislados de CCC y posiblemente logremos determinar rasgos genéticos asociados al estilo de vida salino. Además del análisis de genes compartidos versus genes únicos, será interesante evaluar las vías metabólicas presentes o predominantes en cada uno de estos organismos con el objetivo de describir el posible papel de estos aislados en sus contextos ecológicos particulares.

Es importante mencionar que la comparación genómica entre organismos filogenéticamente cercanos ha arrojado resultados muy interesantes con respecto a las adaptaciones de los microorganismos a ambientes particulares (Rocap *et al.*, 2003, Peña *et al.*, 2010). En particular, para el caso de las bacterias de CCC, será importante profundizar en la presencia de estrategias de adaptación a la salinidad tales como bombas de protones o la síntesis de solutos compatibles (Grant, 2004). Por otra parte, la síntesis de pigmentos como estrategia de tolerancia a la radiación solar en ecosistemas desérticos y de desecación ha mostrado ser importante en diversos grupos bacterianos (Flores *et al.*, 2009). Será interesante determinar si la cepa de agua (11-28) posee estrategias particulares de protección contra la radiación UV en contraste con la cepa de sedimento (EPVM).

Por otro lado, la presencia de oxígeno en ecosistemas microbianos se ha sugerido como un factor determinante para la estructura espacial de las comunidades. En particular, los ambientes de columna de agua son considerados aeróbicos mientras que los sedimentos constituyen uno de los ambientes con gradientes de oxígeno más pronunciados (Fenchel y Finlay, 2008). Con base en lo anterior será importante comparar las vías metabólicas asociadas a la aeróbiosis y anaeróbiosis como parte de la adaptación de estas cepas a hábitats contrastantes (sedimento y columna de agua).

Finalmente, la densidad microbiana en sedimentos y tapetes microbianos es en general mucho más alta que en ambientes de columna de agua. Por lo anterior, se ha sugerido que las interacciones bióticas en tapetes y sedimentos juegan un papel fundamental en la adaptación de las poblaciones microbianas. En específico, las interacciones de competencia promueven la producción de metabolitos, antibióticos, toxinas, así como moléculas esenciales en la producción de biopelículas. Por otra parte, la probabilidad de intercambio genético (por transformación, transducción o conjugación) aumenta en ambientes de alta densidad celular.

Por todo lo anterior, esperaríamos que la cepa de *Exiguobacterium* aislada de la columna de agua hipersalina (11-28) posea estrategias de adaptación a la alta salinidad, baja actividad de agua y resistencia a la radiación. Así mismo, la cepa de *Exiguobacterium* aislada del tapete microbiano (EPVM) podría presentar adaptaciones a los gradientes pronunciados de oxígeno, así como diversas estrategias moleculares para contener con los altos números celulares y la competencia intra e interespecífica característica de estos ambientes.

## 9. REFERENCIAS

- Abby S y V Daubin, 2007. Comparative genomics and the evolution of prokaryotes, *Trends in Microbiol*, 15(3):135-141.
- Acinas SG *et al.*, 2004. Fine-scale phylogenetic architecture of a complex bacterial community. *Nature*, 430(6999): 551-4.
- Alcaráz LD, *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 USA*, 105(15): 5803-8.
- Allen EE y JF Banfield, 2005. Community genomics in microbial ecology and evolution. *Nat Rev Micro*, 3(6): 489-98.
- Allen EE, GW Tyson, RJ Whitaker, JC Detter, PM Richardson y JF Banfield, 2007. Genome dynamics in a natural archaeal population, *Proc Natl Acad Sci USA*, 104(6): 1883-1888.
- Ast JC y PV Dunlap, 2005. Phylogenetic resolution and habitat specificity of members of the *Photobacterium phosphoreum* species group, *Environ Microbiol*, 7(10): 1641-54.
- Bielawski JP, KA Dunn, G Sabehi y O Bèjà, 2005. Darwinian adaptation of proteorhodopsin to different light intensities in the marine environment. *Proc Natl Acad Sci USA*, 101(41): 14824-9.
- Bisharat N, DI Cohen, MC Maiden, DW Crook, T Peto y RM Harding, 2007. The evolution of genetic structure in the marine pathogen, *Vibrio vulnificus*. *Infect Genet Evol*, 7(6):685-93.
- Bonilla-Rosso G *et al.*, 2012. Comparative metagenomics of two microbial mats at Cuatro Ciénegas Basin II: Community Structure and Composition in Oligotrophic Environments. *Astrobiology*, en prensa.
- Breitbart *et al.*, 2009. Metagenomic and stable isotopic analyses of modern freshwater microbialites in Cuatro Ciénegas, Mexico. *Environ Microbiol*. 11(1):16-34.
- Cadillo-Quiroz *et al.*, 2012. Patterns of Gene Flow Define Species of Thermophilic Archaea. *Plos Biol*, 10(2): e1001265.

Castillo A, LE Eguiarte y V Souza, 2005. A genomic population genetics analysis of the pathogenic enterocyte effacement island in *Escherichia coli*: the search for the unit of selection, *Proc Natl Acad Sci USA*, 102(5): 1542-7.

Cerritos *et al.*, 2008. *Bacillus coahuilensis* sp. nov., a moderately halophilic species from a desiccation lagoon in the Cuatro Ciénegas Valley in Coahuila, Mexico. *IJSEM*, 58(4): 919-923.

Cerritos R, 2008. Análisis del concepto biológico, cohesivo, ecológico y filogenético de especie en bacterias halófilas del valle de Cuatro Ciénegas, Coahuila. Tesis de Doctorado en Ciencias Biomédicas, Instituto de Ecología, UNAM.

Cerritos *et al.*, 2010. Diversity of culturable thermo-resistant aquatic bacteria along an environmental gradient in Cuatro Ciénegas, Coahuila, México. *Antonie Van Leeuwenhoek*, DOI 10.1007/s10482-010-9490-9.

Chanal *et al.*, 2006. The desert of Tataouine: an extreme environment that hosts a wide diversity of microorganisms and radiotolerant bacteria. *Environ Microbiol*, 8(3): 514-525.

Cho J y JM Tiedje, 2000. Biogeography and degree of endemism of fluorescent *Pseudomonas* strains in soil, *Appl and Environ Microbiol*, 66(12): 5448-5456.

Ciccarelli FD, T Doerks, C von Mering, CJ Creevey, B Snel y Bork P, 2006. Toward Automatic Reconstruction of a Highly Resolved Tree of Life. *Science*. 311 (5765):1283-1287.

Cohan FM, 2002. What are bacterial species?. *Annu. Rev. Microbiol.* 56(1): 457-487.

Cohan FM y EB Perry, 2007. A Systematics for Discovering the Fundamental Units of Bacterial Diversity. *Current Biology*: 17(10): R373-R386.

Coleman ML *et al.*, 2006. Genomic islands and the ecology and evolution of *Prochlorococcus*, *Science*. 311:1768-1770.

Coscollá M y F González-Candelas, 2007. Population structure and recombination in environmental isolates of *Legionella pneumophila*. *Environ Microbiol*. 9(3): 643-656.

Crapart S, *et al.*, 2007. *Exiguobacterium profundum* sp. nov., a moderately thermophilic, lactic acid-producing bacterium isolated from a deep-sea hydrothermal vent. *IJSEM*. 57(2): 287-92.



- DeLong EF y N Pace, 2001. Environmental diversity of bacteria and archaea. *Syst Biol.* 50(4): 470-8.
- Desnues *et al.*, 2008. Biodiversity and biogeography of phages in modern stromatolites and thrombolites. *Nature.* 452(7185): 340-343
- De Wit R y T Bouvier, 2006. 'Everything is everywhere, but, the environment selects'; what did Baas Becking and Beijerinck really say?. *Environ Microbiol.* 8(4): 755-758.
- Doolittle WF y RT Papke, 2006. Genomics and the bacterial species problem. *Genome Biol.* 7(9)116.
- Dykhuizen DE, 1998. Santa Rosalia revisited: Why are there so many species of bacteria?. *Antonie van Leeuwenhoek.* 73: 25-33.
- Elser *et al.*, 2005. Effects of phosphorus enrichment and grazing snails on modern stromatolitic microbial communities. *Freshwater Biol.* 50(11): 1808-1825.
- Escalante AE. 2008. Ecología Evolutiva de Procariontes en Cuatro Ciénegas Coahuila, México. Tesis de Doctorado en Ciencias Biomédicas, Instituto de Ecología, UNAM.
- Escalante AE, LE Eguiarte, L Espinosa-Asuar, LJ Forney, AM Noguez y V Souza, 2008. Diversity of aquatic prokaryotic communities in the Cuatro Cienegas basin. *FEMS Microbiol Ecol*, 65(1):50-60.
- Escalante AE *et al.*, 2009. *Pseudomonas cuatrocienegasensis* sp. nov., isolated from an evaporating lagoon in the Cuatro Cienegas valley in Coahuila, Mexico. *IJSEM.* 59(6): 1416-1420.
- Falush D *et al.*, 2003, Traces of Human Migrations in *Helicobacter pylori* Populations. *Science.* 299: 1582-1585.
- Farrow JAE, S Wallbanks y MD Collins, 1994. Phylogenetic interrelationships of round spore-forming bacilli containing cell walls based on lysine and the non-spore-forming genera *Caryophanon*, *Exiguobacterium*, *Kurthia* and *Planococcus*. *Int J Syst Bacteriol.* 44: 74-82.
- Feil EJ y BG Spratt, 2001. Recombination and the population structures of bacterial pathogens, *Annu Rev Microbiol*, 55:561-90.
- Feil *et al.*, 2003. How clonal is *Staphylococcus aureus*?. *J Bacteriol.* 185(11): 3307-16.

Fenchel T y B Finlay, 2008. Oxygen and the Spatial Structure of Microbial Communities. *Biol Rev.* 83: 553–569.

Finlay BJ, 2002. Global Dispersal of Free-Living Microbial Eukaryote Species. *Science.* 296(5570): 1061-1063.

Flores MR, OF Ordoñez, MJ Maldonado y ME Farías, 2009. Isolation of UV-B resistant bacteria from two high altitude Andean lakes (4,400 m) with saline and non saline conditions. *J Gen Appl Microbiol.* 55(6): 447-58.

Forney LJ, X Zhou y CJ Brown, 2004. Molecular microbial ecology: land of the one-eyed king. *Curr Opin in Microbiol.* 7:210–220.

Foti M, S Ma, D Yu.Sororkin, JLW Rademaker, JG Kuenen y G Muyzer, 2006. Genetic diversity and biogeography of haloalkaliphilic sulphuroxidizing bacteria belonging to the genus Thioalkalivibrio. *FEMS Microbiol Ecol.* 56: 95–101.

Fraser C, EJ Alm, MF Polz, BG Spratt y WP, 2009. Handle the bacterial species challenge: making sense of genetic and ecological diversity. *Science.* 323(5915): 741-6.

Grant WD, 2004. Life at low water activity. *Phil Trans of the Royal Society.* 359(1448): 1249-1267.

Gogarten JP, WF Doolittle y JF Lawrence, 2002. Prokaryotic evolution in light of gene transfer. *Mol Biol and Evol.* 19(12): 2226-38.

Hanage WP, C Fraser y BG Spratt, 2005, Fuzzy species among recombinogenic bacteria. *BMC Biol.* doi:10.1186/1741-7007-3-6.

Hanage WP, C Fraser y BG Spratt, 2006. Sequences, sequence clusters and bacterial species. *Philos Trans R Soc Lond, B, Biol Sci.* 361(1475): 1917-1927.

Hanage WP, C Fraser y BG Spratt, 2006. The impact of homologous recombination on the generation of diversity in bacteria. *J Theor Biol.* 239(2): 210-219.

Hedrick PW, 2005. Genetics of populations, 3<sup>rd</sup> Edition, Jones & Bartlett Publishers, MA EUA.

Hewson I y JA Fuhrman, 2004. Richness and diversity of bacterioplankton species along an estuarine gradient in Moreton Bay, Australia. *Appl Environ Microbiol.* 70(6): 3425-33.

Hudson RR, M Slatkint y WP Maddison, 1992. Estimation of Levels of Gene Flow From DNA Sequence Data. *Genetics*. 132: 583-589.

Hughes JB *et al.*, 2006, Microbial biogeography: putting microorganisms on the map. *Nat Rev*. 4: 102-112.

Hunt DE, LA David, D Gevers, SP Preheim, EJ Alm y MF Polz. 2008. Resource Partitioning and Sympatric Differentiation Among Closely Related Bacterioplankton. *Science*. 320(5879): 1081-1085

Istock C.A, KE Duncan, N Ferguson y X Zhou, 1992. Sexuality in a natural population of bacteria *Bacillus subtilis* challenges the clonal paradigm. *Mol Ecol*. 1(2):95-103.

Jain R, MC Rivera, JE Moore y JA Lake, 2002, Horizontal gene transfer in microbial genome evolution. *Theor Popul Biol*. 61(4):489-95.

Jolley KA, DJ Wilson, P Kriz, G McVean y MCJ Maiden, 2005. The influence of mutation, recombination, population history, and selection on patterns of genetic diversity in *Neisseria meningitidis*. *Mol Biol Evol*. 22(3): 562-9.

Jones BW, JE López, J Huttenburg y MK Nishiguchi, 2006. Population structure between environmentally transmitted vibrios and bobtail squids using nested clade analysis. *Mol Ecol*. 15:4317-4329.

Jørgensen Barker B y A Boetius, 2007. Feast and famine — microbial life in the deep-sea bed. *Nat Rev Micro*. 5(10): 770-781.

Kampfer P, S Buczolits, A Albrecht, HJ Busse y E Stackebrandt, 2003. Towards a standardized format for the description of a novel species (of an established genus): *Ochrobactrum gallinifaecis* sp. nov. *IJSEM*. 53(3): 893-896.

Keymer DP y AB Boehm. 2011. Recombination Shapes the Structure of an Environmental *Vibrio cholerae* Population. *Appl Environ Microbiol*. 77(2): 537-544.

Kim IG, MH Lee, SY Jung, JJ Song, TK Oh, JH Yoon-Kim, 2005. *Exiguobacterium aestuarii* sp. nov. and *Exiguobacterium marinum* sp. nov., isolated from a tidal flat of the Yellow Sea in Korea. *IJSEM*. 55(2): 885-889.

Konstantinidis KT *et al.*, 2009. Comparative systems biology across an evolutionary gradient within the *Shewanella* genus. *Proc Natl Acad Sci USA*. 106 (37): 15909-14.

Lynch M y TJ Crease, 1990. The analysis of population survey data on DNA sequence variation. *Mol Biol Evol.* 7(4): 377-94.

López-Cortés A, P Schumann, R Pukall y E Stackebrandt, 2006. *Exiguobacterium mexicanum* sp. nov. and *Exiguobacterium artemiae* sp. nov. isolated from the brine shrimp *Artemia franciscana*. *Syst Appl Microbiol.* 29: 183-190.

Maiden M, 2006. Multilocus Sequence Typing of Bacteria, *Annu. Rev. Microbiol.* 60: 561–88.

Maynard-Smith J, NH Smith, M O'Rourke, BG Spratt, 1993. How clonal are bacteria?. *Proc Natl Acad Sci USA.* 90(10): 4384-8.

Maynard Smith J, EJ Feil y NH Smith, 2000, Population structure and evolutionary dynamics of pathogenic bacteria, *Bioessays*, 22:1115-1122.

Maughan H y G Van der Auwera, 2011. Bacillus taxonomy in the genomic era finds phenotypes to be essential though often misleading. *Infect Genet Evol.* 11(5): 789-797.

McInerney JO, JA Cotton y D Pisan, 2008. The prokaryotic tree of life: past, present...and future?. *Trends Ecol Evol.* 23(5):276-81.

Medini D *et al.*, 2008. Microbiology in the post-genomic era. *Nat Rev Microbiol.* 6:419-430.

Miller SR, MD Purugganan y SE Curtis, 2006. Molecular Population Genetics and Phenotypic Diversification of Two Populations of the Thermophilic Cyanobacterium *Mastigocladus laminosus*, *Appl and Environ Microbiol*, 72(4): 2793–2800.

Miller SR, RW Castenholz y D Pedersen, 2007. Phylogeography of the thermophilic cyanobacterium *Mastigocladus laminosus*. *Appl Environ Microbiol.* 73(15): 4751-4759.

Nesbo C, M Dlutek y WF Doolittle, 2006. Recombination in *Thermotoga*: Implications for Species Concepts and Biogeography. *Genetics.* 172:759-769.

Oakley BB, F Carbonero, CJ van der Gast, RJ Hawkins y KJ Purdy, 2010. Evolutionary divergence and biogeography of sympatric niche-differentiated bacterial populations. *ISME J.* 1-10.

- Ochman H, JG Lawrence y EA Groisman, 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature*. 405(6784): 299-304.
- Ordoñez OF, MR Flores, JR Dib, A Paz, ME Farías, 2009. Extremophile Culture Collection from Andean Lakes: Extreme Pristine Environments that Host a Wide Diversity of Microorganisms with Tolerance to UV Radiation. *Microb Ecol*. 58(3): 461-473.
- Oren A, 2004. Prokaryote diversity and taxonomy: current status and future challenges. *Philos Trans R Soc Lond, B, Biol Sci* (2004) vol. 359 (1444) pp. 623-38
- Papke RT, NB Ramsing, MM Bateson y DM Ward, 2003. Geographical isolation in hot spring cyanobacteria. *Environ Microbiol*. 5(8): 650-659.
- Papke RT, JE Koenig, F Rodriguez-Valera y WF Doolittle, 2004. Frequent recombination in a saltern population of Halorubrum. *Science*. 306:1928-1929.
- Papke RT, O Zhaxybayeva, EJ Feil, K Sommerfield, D Muise y WF Doolittle, 2007. Searching for species in haloarchaea, *Proc Natl Acad Sci U S A*. 104(35): 14092-14097.
- Peimbert M *et al.*, 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*. *En prensa*.
- Peña A *et al.*, 2010. Fine-scale evolution: genomic, phenotypic and ecological differentiation in two coexisting *Salinibacter ruber* strains. *ISME J*. doi:10.1038/ismej.2010.6.
- Perez-Losada M, RP Viscidi, JC Demma, J Zenilman y KA Crandall, 2005. Population Genetics of *Neisseria gonorrhoeae* in a High-Prevalence Community Using a Hypervariable Outer Membrane porB and 13 Slowly Evolving Housekeeping Genes. *Mol Biol Evol*. 22(9): 1887-1902.
- Pikuta EV, RB Hoover y J Tang, 2007. Microbial Extremophiles at the Limits of Life. *Critical Revs. in Microbiol*. 33(3): 183-209.
- Raymond B, KL Wyres, SK Sheppard, RJ Ellis y MB Bonsall, 2010. Environmental Factors Determining the Epidemiology and Population Genetic Structure of the *Bacillus cereus* Group in the Field. *PLoS Pathog*. 6(5): e1000905.

Rebollar EA *et al.*, 2012. Water-sediment niche differentiation in ancient marine lineages of *Exiguobacterium* endemic to the Cuatro Ciénegas Basin. *Environ Microbiol*. doi:10.1111/j.1462-2920.2012.02784.x

Rius N, M.C Fusté, C Guasp, J Lalucat y J.G Lorén, 2001, Clonal population Structure of *Pseudomonas stutzeri*, a species with exceptional genetic diversity, *J Bacteriol*, 183(2): 736-744.

Rocap G *et al.*, 2003. Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature*. 424(6952): 1042-1047.

Rodrigues DF, J Goris, TA Vishnivetskaya, D Gilichinsky, MF Thomashow y JM Tiedje, 2006. Characterization of *Exiguobacterium* isolates from the Siberian permafrost. Description of *Exiguobacterium sibiricum* sp. nov. *Extremophiles*. 10(4): 285-94.

Rodrigues DF *et al.*, 2009. Biogeography of two cold-adapted genera: Psychrobacter and *Exiguobacterium*. *ISME J*. 3(6):658-65

Rodriguez-Verdugo A, 2008. Variación estacional en la diversidad de *Pseudomonas* asociadas a un sistema acuático fluctuante. Tesis de Licenciatura en Biología, Facultad de Ciencias, UNAM.

Rozas J y R Rozas, 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*. 15(2): 174-5.

Sarkar SF y DS Guttman, 2004. Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. *Appl and Environ Microbiol*. 70(4): 1999-2012.

Sapp J, 2005. The prokaryote-eukaryote dichotomy: meanings and mythology. *Microbiol Mol Biol Rev*. 69(2): 292-305.

Sikorski J y E Nevo, 2005. Adaptation and incipient sympatric speciation of *Bacillus simplex* under microclimatic contrast at "Evolution Canyons" I and II, Israel. *Proc Natl Acad Sci U S A*. 102:15924-15929.

Sikorski J y E Nevo, 2007. Patterns of thermal adaptation of *Bacillus simplex* to the microclimatically contrasting slopes of 'Evolution Canyons' I and II, Israel. *Environ Microbiol*. 9(3): 716-726.

- Silva C, P Vinuesa, LE Eguiarte, V Souza, E Martínez-Romero, 2005. Evolutionary genetics and biogeographic structure of *Rhizobium gallicum* sensu lato, a widely distributed bacterial symbiont of diverse legumes. *Mol. Ecol.* 14: 4033-4050.
- Shapiro BJ, LA David, J Friedman y EJ Alm, 2009. Looking for Darwin's footprints in the microbial world. *Trends in Microbiol.* 17(5): 196-204.
- Shapiro *et al.*, 2012. Population Genomics of Early Events in the Ecological Differentiation of Bacteria. *Science.* 336(6077): 48-51.
- Sokurenko EV, R Gomulkiexicz y DE Dykhuizen, 2006. Source–sink dynamics of virulence evolution. *Nat Rev.* 4: 548-555.
- Sorokin *et al.*, 2006. Multiple-locus sequence typing analysis of *Bacillus cereus* and *Bacillus thuringiensis* reveals separate clustering and a distinct population structure of psychrotrophic strains. *Appl Environ Microbiol.* 72(2): 1569-78.
- Souza V, TT Nguyen, RR Hudson, D Piñero y RE Lenski, 1992. Hierarchical analysis of linkage disequilibrium in *Rhizobium* populations: evidence for sex?. *Proc Natl Acad Sci USA.* 89(17): 8389-93.
- Souza V *et al.*, 2006, An endangered oasis of aquatic microbial biodiversity in the Chihuahuan desert. *Proc Natl Acad Sci U S A.* 103(17): 6565– 6570.
- Souza V, LE Eguiarte, J Siefert y JJ Elser, 2008. Microbial endemism: does phosphorous limitations enhance speciation. *Nat Rev Microbiol.* 6(7): 559-564.
- Spiers AJ, A Buckling y PB Rainey, 2000. The causes of *Pseudomonas* diversity. *Microbiol.* 146(10): 2345-50.
- Tindall BJ, R Rosselló-Mora, HJ Busse, W Ludwig y P Kampfer, 2010. Notes on the characterization of prokaryote strains for taxonomic purposes. *IJSEM.* 60(1): 249-266.
- Van Mooy BAS *et al.*, 2009. Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature.* 458(7234): 69-72.
- Vinuesa P, C Silva, D Werner, E Martinez-Romero, 2008. Multilocus Sequence Analysis for Assessment of the Biogeography and Evolutionary Genetics of Four *Bradyrhizobium* Species That Nodulate Soybeans on the Asiatic Continent. *Appl Environ Microbiol.* 74(22): 6987-6996.

- Vishnivetskaya TA, S Kathariou, JM Tiedje, 2009. The *Exiguobacterium* genus: biodiversity and biogeography. *Extremophiles*. 13(3): 541-555.
- Vishnivetskaya TA *et al.*, 2011 Complete Genome Sequence of the Thermophilic *Exiguobacterium* sp. AT1b. *J of Bacteriol*. 193(11): 2880-1.
- Vos M y G Velicer, 2008. Isolation by Distance in the Spore-Forming Soil Bacterium *Myxococcus xanthus*. *Curr Biol*. 18(5): 386-391.
- Vos M y X Didelot. 2009. A comparison of homologous recombination rates in bacteria and archaea. *ISME J*. 3(2):199-208.
- Vos. A, 2010. species concept for bacteria based on adaptive divergence. *Trends in Microbiol*. 19(1): 1-7.
- Ward BB y GD O´Mullan, 2002. Worldwide distribution of *Nitrosococcus oceani*, a marine ammonia-oxidizing  $\gamma$ -Proteobacterium, detected by PCR and sequencing of 16srRNA and amoA genes. *Appl and Environ Microbiol*. 68(8): 4153-4157.
- Whitaker RJ, DW Grogan y JW Taylor, 2003. Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science*. 301: 976-978.
- Whitaker RJ, DW Grogan y JW Taylor, 2005. Recombination shapes the natural population structure of the hyperthermophilic archaeon *Sulfolobus islandicus*. *Mol Biol Evol*. 22(12): 2354-2361.
- Whitaker RJ y JF Banfield, 2006. Population genomics in natural microbial communities, *TRENDS in Ecol and Evol*. 21(9): 508-516.
- Whitaker RJ, 2006. Allopatric origins of microbial species, *Phil. Trans. R. Soc. B*. 361:1975-1984.
- Whitman WB, DC Coleman, WJ Wiebe, 1998. Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA*. 95(12): 6578-83.
- Wise MG, LJ Shimkets, JV McArthur, 1995. Genetic structure of a lotic population of Burkholderia (*Pseudomonas*) cepacia. *Appl Environ Microbiol*. 61(5):1791-8.
- Woese C, O Kandler y ML Wheelis, 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA*. 87(12): 4576-9.



Woese C, 1987. Bacterial evolution, *Microbiol Rev.* 51(2):221-271.

Wolaver BD, JM Sharp, JM Rodriguez y JC Ibarra Flores, 2008. Delineation of Regional Arid Karstic Aquifers: An Integrative Data Approach. *Ground Water.* 46(3): 396-413.

Xu J, 2006. Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. *Mol Ecol.* 15: 1713-1731.

Zwick *et al.*, 2011. Genetic variation and linkage disequilibrium in *Bacillus anthracis*. *Sci. Rep.* 1:1-8.