



CENTRO DE INVESTIGACION Y DE ESTUDIOS  
AVANZADOS  
DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD ZACATENCO

Departamento de Biotecnología y Bioingeniería

**« Factores ambientales que definen la composición de  
las comunidades bacterianas en un suelo haloalcalino  
extremo altamente heterogéneo»**

*Que presenta:*

Martha Adriana Martínez Olivas

*para obtener el grado de*

*Doctor en Ciencias*

*en la especialidad de Biotecnología*

*Director:*

Dr. Luc Julien Jerome Dendooven

*Co-Director:*

Dra. Yendi Ebenezer Navarro Noya

Ciudad de México

6 de enero de 2019



*«El orgullo se preocupa de quién está en lo correcto, la humildad de lo que es correcto.»*

Ezra T. Benson

*«During our teaching in universities at home and abroad, we have been repeating a key message to graduate students: While it is important to learn about the methods developed by previous generations of scientists, do not let yourself be silenced by their aura. If you think you have a good idea, work on it, develop it, listen to criticisms, and publish it, thus contributing to the advancement of the field. Do not let people tell you that everything is known, or that you are too young or not good enough to contribute to this field — or any other field of science.»*

Numerical Ecology by Pierre Legendre and Louis Legendre



## *Resumen*

En un gran número de estudios se ha mostrado que la distribución de microorganismos en el suelo no es aleatoria, si no que su abundancia cambia a lo largo de gradientes ambientales (formando patrones espaciales). En este estudio se examinó la variabilidad espacial de ciertas características en un suelo salino y alcalino extremo y se determinó cuáles de ellas ejercen control sobre las comunidades bacterianas y de arqueas. La estructura de comunidades y las características del suelo fueron determinadas en 13 puntos a lo largo de un transecto de 211 m en la cama del exlago de Texcoco. Se utilizaron técnicas geoestadísticas y de análisis multivariados para definir cuáles factores influyen más sobre la distribución de comunidades. Se presentó a lo largo del transecto una alta variabilidad en la conductividad electrolítica, contenido de agua y C orgánico. Euryarchaeota dominó entre las arqueas presentes, con excepción de sitios donde la EC fue baja y cubiertos de vegetación. Proteobacteria, Actinobacteria y Gemmatimonadetes fueron los phyla bacterianos dominantes con poca variabilidad. Los Phyla menos dominantes fueron más heterogéneos (ej. Firmicutes). El contenido de agua y la EC resultaron los conductores principales de la estructura de comunidades Arquea. La presencia de Thaumarchaeota depende de EC y contenido de agua bajos, mostrando patrones periódicos a lo largo del transecto. La diversidad bacteriana fue menos controlada por características del suelo medidas que la de Arquea, siendo éstas mejor adaptadas a ambientes extremos. Las herramientas geoestadísticas pueden ser de mucha ayuda si se intentan definir las dinámicas e interacciones ecológicas entre Arquea, Bacteria y su entorno, teniendo en cuenta que tanto la escala como las estrategias de muestreo son importantes al intentar definir los patrones de distribución espacial.

## *Abstract*

In a large number of studies have been shown that the distribution of microorganisms in the soil is not random, but that their abundance changes along environmental gradients (spatial patterns). This study examined the spatial variability of certain characteristics in an extreme saline and alkaline soil and how they controlled the bacterial and archaeal communities, thus determining the main drivers of the diversity and structure of the communities. The archaeal and bacterial community structure and soil characteristics were determined at 13 points along a 211 m transect in the former lake Texcoco. Geostatistical techniques were used to describe spatial distribution patterns of the microbial communities, as well as multivariate analysis techniques to evaluate how the studied physicochemical properties defined the prokaryotic community structure. High variability in the electrolytic conductivity (EC) values, water content (H) and organic C content were presented along the transect. Euryarchaeota dominates Archaea except in sites where EC and H were low, covered vegetation. Proteobacteria, Bacteroidetes and Actinobacteria were the dominant bacterial phyla with low variability. Minor phyla, such as Firmicutes, were more heterogeneous. Multivariate analysis showed that soil H and EC were the main drivers of distribution. As it can be seen from geostatistical analysis the presence of Thaumarchaeota depends on low EC and H, showing periodic patterns along the transect. Bacterial alpha diversity was less controlled by soil characteristics than Archaea, being Archaea more adapted to extreme conditions. Geostatistical tools can be very helpful if tried to define dynamics and ecological interactions between Archaea, Bacteria and their environment, taking into account that both, the scale and sampling strategies are of important consideration to identify patterns of spatial distribution.

# *Agradecimientos*

A las personas que creyeron en mi, cuando yo no lo hacía.

A los obstáculos por que me permiten seguir retándome y conociéndome.

A mis amigos, los que de lejos me muestran siempre su apoyo y cariño, y a aquellos que tengo cerca para seguir molestando. Como ya lo reconocí una vez, sus nombres no estarán en este texto, pero si en mi pensamiento, sin duda, sin ustedes no hubiera sido posible nada de esto.

A los que estuvieron conmigo al inicio de éste camino y que hoy siguen el suyo.

A aquellos que me han dado una oportunidad, aún en mis peores momentos.

A mis padres y mi familia, que desde lejos me muestran su afecto, cariño y apoyo.

A Omar por su apoyo, por compartir conmigo su entusiasmo por la ciencia, por animarme y comprenderme.

A CONACyT por el apoyo brindado mediante la beca.

A los Directores Dr. Dendooven y Dra. Navarro por el espacio para la realización del proyecto.

A mi comité por su apoyo en la revisión del proyecto.

A Freddy y a Norma por su valiosa ayuda a lo largo de todo el análisis de datos.

A mis compañeros del laboratorio de ecología molecular de suelos y de limnología por su apoyo.

Y hoy, como siempre, a la música, que siempre me permite el enfoque que la mayor parte de las veces me es difícil encontrar.





# Índice general

<b>Resumen</b>	<b>V</b>
<b>Abstract</b>	<b>VI</b>
<b>Agradecimientos</b>	<b>VII</b>
<b>1. Introducción</b>	<b>1</b>
<b>2. Antecedentes</b>	<b>3</b>
2.1. El suelo como ecosistema . . . . .	3
2.2. Ecología espacial . . . . .	4
2.3. Biogeografía . . . . .	5
2.4. Geoestadística . . . . .	7
2.4.1. Autocorrelación . . . . .	8
2.4.1.1. Índice de Moran . . . . .	9
2.4.1.2. Construcción del correlograma . . . . .	9
2.4.2. Semivariograma . . . . .	11
2.4.3. Predicción . . . . .	11
2.4.3.1. Modelos de predicción estadística: Kriging . . . . .	12
2.4.3.2. Selección del modelo predictivo . . . . .	15
2.4.3.3. Métodos de validación de la interpolación . . . . .	16
2.5. Herramientas moleculares para el estudio ecológico del suelo . . . . .	17
2.5.1. Librerías 16S rRNA por pirosecuenciación . . . . .	18
2.6. Métodos para Examinar la Diversidad de Comunidades Microbianas . . . . .	19
2.6.1. Métodos de medición a nivel de especie . . . . .	19
2.6.1.1. Métodos de diversidad alfa . . . . .	20
2.6.1.2. Métodos de diversidad Beta . . . . .	22
<b>3. Justificación y Objetivos</b>	<b>24</b>
3.1. Justificación . . . . .	24
3.2. Objetivos . . . . .	25
3.2.1. Objetivo General . . . . .	25
3.2.2. Objetivos Particulares . . . . .	25

<b>4. Materiales y Métodos</b>	<b>26</b>
4.1. Sitio experimental y muestreo . . . . .	26
4.1.1. Muestreo . . . . .	27
4.2. Caracterización fisicoquímica del suelo . . . . .	28
4.2.1. Determinación de pH . . . . .	28
4.2.2. Determinación de conductividad electrolítica (EC) . . . . .	29
4.2.3. Determinación del contenido de agua . . . . .	29
4.2.4. Determinación de la clase textural . . . . .	29
4.2.5. Determinación de Carbono Total, inorgánico y orgánico . . . . .	30
4.3. Extracción de DNA metagenómico del suelo . . . . .	32
4.4. Amplificación y Preparación de Librería de Genes rRNA 16S . . . . .	33
4.4.1. Bacterias . . . . .	33
4.4.2. Arqueas . . . . .	34
4.4.3. Secuenciamiento . . . . .	34
4.5. Determinación de estructura y la diversidad de comunidades . . . . .	34
4.6. Asignación manual de secuencias . . . . .	35
4.7. Análisis filogenético . . . . .	35
4.8. Análisis estadístico y geoestadístico . . . . .	36
4.8.1. Estadística descriptiva . . . . .	36
4.8.2. Análisis geoestadísticos . . . . .	36
4.8.2.1. Interpolación . . . . .	37
4.8.2.2. Validación de la interpolación . . . . .	37
4.8.3. Determinación de conductores fisicoquímicos . . . . .	37
4.9. Accesibilidad de datos . . . . .	38
<b>5. Resultados y Discusión</b>	<b>39</b>
5.1. Características fisicoquímicas del suelo . . . . .	39
5.1.1. Análisis descriptivos . . . . .	41
5.1.2. Correlaciones entre características del suelo . . . . .	43
5.2. Análisis geoestadísticos de parámetros fisicoquímicos del suelo . . . . .	44
5.2.1. Exploratorios geoestadísticos . . . . .	44
5.2.2. Descripción de la estructura espacial de características del suelo . . . . .	45
5.2.2.1. Autocorrelación . . . . .	45
5.2.2.2. Semivariogramas . . . . .	46
5.3. Estructura de comunidades microbianas . . . . .	49
5.3.1. Estructura de las comunidades de Arquea . . . . .	49
5.3.2. Estructura de las comunidades de Bacteria . . . . .	52
5.3.3. Relación Arquea/Bacteria . . . . .	57
5.3.4. Asignación manual de secuencias . . . . .	61
5.4. Mapas de Distribución . . . . .	65
5.4.1. Modelos predictivos de características del suelo . . . . .	65
5.4.1.1. Validación . . . . .	67
5.4.2. Mapas de Distribución de comunidades Arquea . . . . .	67
5.4.3. Mapas de Distribución de comunidades bacterianas . . . . .	71

5.5. Factores que conducen la distribución de comunidades . . . . .	74
5.5.1. Correlaciones: comunidades y características . . . . .	76
5.5.1.1. Arqueas y sus conductores . . . . .	76
5.5.1.2. Bacterias y sus conductores . . . . .	79
5.6. Patrones de distribución espacial . . . . .	86
<b>6. Conclusión</b>	<b>88</b>
<b>A. Físicoquímicos y exploratorios</b>	<b>91</b>
A.1. Estadística descriptiva . . . . .	91
A.1.1. Box-plot físicoquímicos . . . . .	91
A.1.2. Pruebas de normalidad y tendencia . . . . .	91
A.2. Geoexploratorios . . . . .	91
<b>B. Estructura Espacial</b>	<b>97</b>
B.1. Autocorrelogramas . . . . .	97
B.2. Semivariogramas . . . . .	97
<b>C. Estructura de comunidades</b>	<b>101</b>
C.1. Curvas de rarefacción . . . . .	101
C.2. Gráficas de niveles Taxonómicos . . . . .	101
C.2.1. Asignación qíime . . . . .	101
C.2.2. Asignación manual . . . . .	101
C.3. Asignación manual . . . . .	110
<b>D. Correlaciones y gráficos de ordenación</b>	<b>121</b>
D.1. PCA . . . . .	121
D.2. Beta diversidad PCoA . . . . .	121
D.3. Clustering . . . . .	127
<b>Bibliografía</b>	<b>129</b>

# Índice de figuras

2.1. Correlogramas Moran . . . . .	10
2.2. Selección de Modelo . . . . .	15
4.1. Sitio de Muestreo . . . . .	27
4.2. Triángulo de Texturas . . . . .	31
4.3. Determinación de Carbono . . . . .	32
5.1. Textura en Texcoco . . . . .	42
5.2. Semivariogramas Ajustados para EC y H . . . . .	47
5.3. Abundancia porcentual de Géneros Arquea . . . . .	50
5.4. Abundancia porcentual de phyla bacterianos . . . . .	53
5.5. Heatmap de abundancia de géneros bacterianos . . . . .	54
5.6. Abundancia porcentual de especies asignadas manualmente . . . . .	62
5.7. Abundancia porcentual de Especies asignadas manualmente . . . . .	63
5.8. Abundancia porcentual de Taxas asignados manualmente . . . . .	64
5.9. Mapas FQ1 . . . . .	66
5.10. Mapas de $\alpha$ Diversidad Arquea . . . . .	69
5.11. Mapas phyla Arquea . . . . .	70
5.12. Mapas de $\alpha$ Diversidad Bacteria . . . . .	72
5.13. Mapas phyla Bacteria . . . . .	73
5.14. Mapas Bacteria . . . . .	75
5.15. Correlaciones Arqueas . . . . .	77
5.16. Correlaciones Arqueas . . . . .	78
5.17. CAP Arquea . . . . .	79
5.18. Correlaciones Bacteria . . . . .	81
5.19. CAPs Bacteria . . . . .	82
5.20. Correlaciones Bacteria . . . . .	83
5.21. Correlación asignación manual Bacteria . . . . .	84
A.1. Boxplot . . . . .	92
A.2. Q-Q plot . . . . .	93
A.3. Q-Q plot Textura . . . . .	94
A.4. Geoexploratorios . . . . .	95
A.5. Geoexploratorios . . . . .	96

B.1. Autocorrelogramas I de Moran . . . . .	98
B.2. Semivariogramas . . . . .	99
B.3. Semivariogramas . . . . .	100
C.1. Curvas de Rarefacción . . . . .	102
C.2. Resumen de Taxas Arquea . . . . .	103
C.3. Familias Arquea . . . . .	104
C.4. Resumen de Taxas Bacteria . . . . .	105
C.5. Familias Bacteria . . . . .	106
C.6. Resumen de Taxas Arquea por asignación manual . . . . .	107
C.7. Resumen de Taxas por asignación manual para Bacteria . . . . .	108
C.8. Asignación manual de Familias Bacteria . . . . .	109
D.1. Análisis de Componentes principales (PCA) de las características del suelo . . . . .	122
D.2. PCA para Arqueas . . . . .	123
D.3. PCA para Bacteria . . . . .	124
D.4. Beta Diversidad de Arquea y Bacteria . . . . .	125
D.5. Beta Diversidad de Arquea y Bacteria . . . . .	126
D.6. Clustering de sitios . . . . .	128

# Índice de tablas

2.1. Modificada de Martiny et al. 2006. Los estudios están clasificados por la escala geográfica sobre la cual se tomaron las muestras, reportada como la distancia más pequeña entre los puntos de muestreo. Si los autores informaron que el patrón en la composición microbiana se correlacionó con una característica del medio ambiente, se informa en la columna "correlación con", aunque esta relación no haya sido probada estadísticamente. . . . .	6
5.1. Características Fisicoquímicas a lo largo del transecto de muestreo en la cama del exlago de Texcoco. Descripción del suelo para cada sitio de muestreo establecido a distancias irregulares a lo largo de un transecto al sureste en el área del exlago de Texcoco y análisis estadísticos descriptivos para cada parámetro medido en conjunto. Cada sitio de muestreo se refiere con la clave "Tx"seguido por un número consecutivo indicando progresión en separación espacial. La Fecha de muestreo fue establecida el 22-08-2014 . . . . .	40
5.2. <b>Clasificación Textural.</b> Tabla de sitios (con las correspondientes transformaciones de coordenadas) distribución de tamaños de partícula y clasificación textural de acuerdo a la clasificación USDA (Figs. 4.2 y 5.1) . . . . .	41
5.3. <b>Variabes ambientales con las más altas correlaciones.</b> Correlación entre variables determinada para evitar redundancia en el cálculo de la interpolación. . . . .	43
5.4. <b>Multicolinealidad y Factor de inflacion de varianza (VIF).</b> Propiedades del suelo que son aproximadamente combinación lineal de otras. . . . .	44
5.5. <b>Autocorrelación.</b> Distancias a las cuales los parámetros fisicoquímicos medidos muestran autocorrelación significativa según el índice de Moran calculado usando 1000 permutaciones ( $p<0.05$ ) . . . . .	46
5.6. <b>Parámetros de Ajuste de Semivariograma.</b> La tabla muestra los parámetros característicos del modelo de ajuste del semivariograma, sill, nugget y rango. Así mismo una relación calculada nugget/sill que muestra la presencia o ausencia de autocorrelación. . . . .	47

5.7. <b>Diversidad de arqueas.</b> Índices de diversidad y profundidad de secuenciación de Arquea a cada sitio de muestreo (Tx001-Tx013) a lo largo de un transecto corriendo al sureste en suelo del exlago de Texcoco. . . . .	51
5.8. <b>Diversidad de Bacteria.</b> Índices de diversidad y profundidad de secuenciación de Bacteria para cada sitio de muestreo (Tx001-Tx013) a lo largo de un transecto corriendo al sureste en suelo del exlago de Texcoco. . . . .	56
5.9. <b>Relación Arquea/Bacteria.</b> Tabla de relaciones Arquea/Bacteria para los índices de diversidad y niveles taxonómicos en cada sitio a lo largo del transecto de muestreo . . . . .	58
5.10. <b>Validación de la interpolación.</b> Índices calculados para los mapas generados por kriging ordinario en dos polígonos de distintas áreas (3000 pixeles de A:2.82 m <sup>2</sup> y B:1.29 m <sup>2</sup> cada uno) para las características del suelo y ciertos phyla de Arquea y Bacteria. . . . .	68
5.11. <b>Conjunto de variables con la mejor correlación para las comunidades microbianas.</b> Distancias Euclidianas de las variables escaladas que tienen la máxima correlación $r$ con las disimilitudes en las comunidades (Función BIOENV del paquete Vegan en R). La tabla muestra los coeficientes Spearman ( $r$ ) para la correlación entre las comunidades bacterianas y arquea del suelo y las propiedades fisicoquímicas del mismo. . . . .	85
C.1. Asignación manual de secuencias representativas para Arquea . . . . .	111
C.2. Asignación manual de secuencias representativas para Bacteria . . . . .	116

# Abreviaturas

<b>EC</b>	Conductividad Electrolítica
<b>H</b>	Humedad o Contenido de agua
<b>Ci</b>	Contenido de Carbono inorgánico
<b>Co</b>	Contenido de Carbono orgánico
<b>CT</b>	Contenido de Carbono Total
<b>ASE</b>	Error Estándar promedio ( <i>Average Standar Error</i> )
<b>RMSE</b>	Raíz cuadrada del Error ( <i>Root Mean Squared Error</i> )
<b>RMSEr</b>	Raíz cuadrada del Error normalizado ( <i>Root Mean Squared Error relative</i> )
<b>DNA</b>	Ácido DesoxirriboNucléico
<b>RNA</b>	Ácido Ribonucléico
<b>PCR</b>	Reacción en Cadena de la Polimerasa
<b>pb</b>	pares de bases
<b>OTU</b>	Unidades Taxonómicas Operacionales
<b>BIOM</b>	Matrices de Observación Biológica
<b>PCoA</b>	Análisis de Coordenadas Principales
<b>CAP</b>	Análisis Canónico de coordenadas Principales
<b>PCA</b>	Análisis de Componentes Principales
<b>AOA</b>	Arqueas Oxidantes de Amonio
<b>AOB</b>	Bacterias Oxidantes de Amonio
<b>I</b>	Índice (p.ej. en Índice de Moran (I de Moran))
<b>KO</b>	Kriging Ordinario
<b>LOO</b>	Leave One Out
<b>rRNA</b>	Ácido Ribo nucléico ribosomal
<b>N</b>	Norte
<b>E</b>	Este
<b>O</b>	Oeste
<b>Lat</b>	Latitud
<b>Long</b>	Longitud
<b>SMN</b>	Sistema Meteorológico Nacional
<b>GPS</b>	Sistema de Posicionamiento Global
<b>ss</b>	suelo seco



# Lista de Símbolos

$M$	concentración	M (mol L <sup>-1</sup> )
$P$	peso	g
$\rho$	coeficiente correlación Spearman	No aplica



*A mis padres Martha y Simón, a mis abuelos Mamá Pina, Manuel, Carolina y Benito; hay una parte de ustedes en cada esfuerzo que realizo. Gracias por enseñarme paciencia y un poco de su gran fortaleza.*



# Capítulo 1

## Introducción

*«In questions of science, the authority of a thousand is not worth the humble reasoning of a single individual.»*

---

*Galileo Galilei*

En numerosos estudios se ha demostrado que la distribución de microorganismos en el suelo no es aleatoria, sino que su abundancia y actividad cambian a lo largo de gradientes ambientales (patrones espaciales). En algunos estudios se ha observado una correlación entre las comunidades con el pH, la conductividad electrolítica (EC) y la salinidad (Andrew et al. [2012](#); Inskeep et al. [2013](#); Crits-Christoph et al. [2013](#)).

El estudio de patrones espaciales microbianos, trae consigo importantes implicaciones tales como la predicción de arreglos físicos (espaciales) de microorganismos, la comprensión de los procesos que dan lugar a la diversidad y procesos evolutivos, además del impacto que tienen las actividades humanas sobre la biodiversidad (y por consecuencia en los ciclos biogeoquímicos) a diferentes escalas, e incluso puede llevar al desarrollo de estrategias para el manejo de ecosistemas (Philippot et al. [2009](#); Bru et al. [2011](#)).

El suelo drenado del lago de Texcoco tiene características que lo hacen un ambiente único y altamente heterogéneo, con valores de pH mayores de 10 y de EC que alcanzan y sobrepasan los  $150 \text{ dS m}^{-1}$  (Dendooven et al. 2010). Se han llevado a cabo estudios para determinar la estructura de las comunidades, tanto bacterianas como de Arqueas y de hongos, que se pueden encontrar en la zona (Valenzuela-Encinas et al. 2012; Ruiz-Romero et al. 2013), pero no se han relacionado a características propias del suelo en escala "*landscape*".

Dada la alta heterogeneidad del suelo con características salinas y alcalinas extremas, se plantea que la distribución de comunidades bacterianas presentes en la zona del exlago de Texcoco, se encontrará directamente relacionada con la variabilidad de parámetros del suelo, específicamente pH y conductividad electrolítica (EC), encontrándose zonas de transición de dichas comunidades.

# Capítulo 2

## Antecedentes

*«Everything about microscopic life is terribly upsetting. How can things so small be so important?»*

---

*Issac Asimov*

### 2.1. El suelo como ecosistema

El suelo comprende el ambiente más heterogéneo de la tierra, tanto espacial como temporalmente y gracias a su estructura, un ensamble organo-mineral, proporciona un entorno para los organismos vivos. Los componentes orgánicos, minerales y microorganismos afectan profundamente las propiedades físicas, químicas y biológicas y los procesos de los sistemas terrestres. Los procesos físicos, químicos y biológicos en el ambiente del suelo no son independientes, si no que son procesos interactivos. De esta forma, se pueden definir los suelos como reactores biogeoquímicos complejos, reservorios de organismos y el principal compartimento de los ecosistemas terrestres bajo la influencia de actividades humanas (Huang, Yuncong y Sumner 2012). Los suelos se encuentran estructurados a diferentes escalas espaciales, la estructura espacial puede presentarse tanto de forma horizontal como vertical, siendo el perfil vertical el que más ampliamente se conoce, ya que se presenta en varias capas, que

difieren en el contenido de materia orgánica, contenido mineral y textura. Por otro lado, el perfil horizontal no se encuentra tan claro y puede ir desde una escala muy pequeña (microhabitats) hasta una escala tan amplia como un continente (Wallenius 2011). A una escala global, el clima constituye un factor de suma importancia para la formación de los suelos, mientras que, conforme la escala disminuye la estructura del suelo determina los microambientes que puede presentar, así, se puede decir que la composición y estructura de los organismos en el suelo manifiesta tanto la organización espacial como el papel que presentarán en los procesos de descomposición (Lavelle et al. 2003).

## 2.2. Ecología espacial

Entender los procesos que llevan a la generación y mantenimiento de patrones espaciales en la biodiversidad es una meta fundamental de la investigación ecológica (Bissett et al. 2010). Sobre una larga escala espacial, la ecología "*landscape*" se enfoca en los patrones espaciales y los procesos relacionados a ellos; considerando el desarrollo de la heterogeneidad espacial y las interacciones espaciales y temporales a través de áreas heterogéneas. Este tipo de estudios han sido ampliamente utilizados en ecología vegetal y han sido extendidos hasta incorporar animales terrestres y acuáticos en ellos y en la actualidad, con el desarrollo de las herramientas "ómicas", se ha logrado incorporar también a los microorganismos. El cómo, cuándo y dónde obtener las muestras de las pruebas de campo que serán utilizadas en análisis ecológicos, son parte fundamental de la estrategia de muestreo, y es necesario tomarla en cuenta tanto para considerar los gastos que se generarán de los análisis, como para obtener datos cuya interpretación no se dificulte (Jongman, Ter Braak y Van Tongeren 2007). Los datos obtenidos en ecología son principalmente multivariados,



es decir cada unidad de muestreo es caracterizada por algunos atributos, lo que ocasiona que los datos sean muy complejos, estén sujetos a redundancia y ruido y que existan relaciones internas entre ellos (Jongman, Ter Braak y Van Tongeren 2007). En los ecosistemas, la heterogeneidad espacial es funcional, es decir, el funcionamiento de los ecosistemas depende de ello, puesto que puede representar una contribución importante en la coexistencia de especies en comunidades y en la persistencia de las poblaciones (Ettema y Wardle 2002).

### 2.3. Biogeografía

La biogeografía es la ciencia que intenta describir y explicar los patrones espaciales que presenta la diversidad biológica. Aunque el concepto de biogeografía existe desde hace ya tiempo, es en la actualidad donde surgen mejores herramientas para el estudio espacial de la distribución biológica, especialmente al nivel de los microorganismos. A la fecha, los estudios sobre la diversidad microbiana del suelo son cada vez mayores. Se ha incrementado la cantidad de estudios que muestran que los organismos de vida libre, detectados mediante herramientas “ómicas” diversas, varían en abundancia, distribución y diversidad a varias escalas espaciales y taxonómicas, algunos de estos estudios se muestran en la Tabla 2.1.

Algunos estudios como el de Sullivan, McBride y Thies 2013, demuestran la intrincada relación que existe entre la química del suelo y las comunidades bacterianas, este estudio encontró que la variabilidad microbiana en un suelo contaminado con metales pesados era conducida principalmente por el pH así como también por los metales cobre y zinc presentes y el azufre total. En el año 2011 Bru y colaboradores determinaron que del 43 al 85 % de la variación espacial en abundancia de comunidades podía explicarse por la medición de parámetros ambientales, siendo el pH el

TABLA 2.1: Modificada de Martiny et al. 2006. Los estudios están clasificados por la escala geográfica sobre la cual se tomaron las muestras, reportada como la distancia más pequeña entre los puntos de muestreo. Si los autores informaron que el patrón en la composición microbiana se correlacionó con una característica del medio ambiente, se informa en la columna “correlación con”, aunque esta relación no haya sido probada estadísticamente.

Organismos	Escala (Km)	Hábitat	Herramienta	Correlación con:	Referencia
<b>Pseudomonadales</b>	20,000	suelo	Aislamiento BOX-PCR	Distancia lineal	Cho y Tiedje, 2000*
<b>Bacteria 3-CBD</b>	20,000	suelo	ARDRA aislamiento		Fulthorpe et al., 1998*
<b>Crenarchaeota</b>	200	suelo	PCR-SSCP de 16S	A pequeña escala con distancia	Sliwinsky y Goodman, 2004*
<b>Crenarchaeota</b>	200	suelo	PCR-SSCP de 16S	Rizósfera vs Cantidad de suelo	Sliwinsky y Goodman, 2004*
<b>Bacteria, Arquea y Eucariotas</b>	3	Salar	DGGE, TRFLP, ARISA	Salinidad	Casamayor et al., 2002*
<b>Bacteria</b>	5	Vertiente	DGGE de 16S	Salinidad	Crump et al., 2004*
<b>Bacteria Gram-Neg.</b>	0.8	Suelo	única fuente de Carbono	Latitud	Staddon et al., 2004*
<b>Microorganismos</b>	0.1	Suelo agrícola	AFLP		Franklin y Mills, 2002
<b>Bacteria</b>	0.01	Suelo	RFLPs de 16S		Noguez et al., 2005*
<b>Bacteria</b>	0.0035	Laguna	ARISA	Distancia	Lear et al., 2014
<b>Bacteria</b>	0.0005	Suelo	FAMES	Propiedades físicoquímicas y temporales	Regan et. al., 2014

3-CDB: Degradadora de 3-clorobenzoato; AFLP: polimorfismo en longitud de fragmentos amplificados; ARDRA: Análisis de restricción de DNA Ribosomal; DGGE: Electroforesis de Gradiente desnaturizante; PCR-SSCP: Reacción en Cadena de Polimerasa-Polimorfismo conformacional de cadena sencilla; RAPD: Amplificación aleatoria de DNA polimórfico; RFLP: Polimorfismo de longitud de fragmentos de restricción; ARISA: Análisis automatizado de los espaciadores intergénicos ribosomales; FAMES: Metilación de ácidos grasos. \* Las referencias marcadas pueden ser tomadas de Martiny et al. 2006

principal operador, por otro lado Wessèn y colaboradores en ese mismo año detectaron que el tamaño y estructura de las comunidades de arqueas y bacterianas oxidantes de amonio (AOA y AOB respectivamente) no se encontraron aleatoriamente distribuidas en el sitio, si no que presentaron patrones espaciales a escala hectárea, que pudieron ser modelados. Los patrones no evidenciaron los sistemas de cultivo, pero la comunidad de AOB estuvo débilmente relacionada a diferencias en el pH y humectación del suelo, mientras las AOA a diferencias en el pH y contenido de arcilla. Determinaron que el tamaño de las comunidades se relacionó de manera distinta a las propiedades del suelo: carbono orgánico y nitrógeno total con la abundancia de AOA (correlación positiva), mientras el pH y contenido de arcilla se encontraron negativamente relacionados con la abundancia de AOA. En contraste con los patrones espaciales, determinaron que las comunidades AOB y AOA pueden ocupar diferentes nichos en agro-ecosistemas.

Considerando los estudios que se han realizado sobre biogeografía de microorganismos, se puede decir que tanto bacterias como arqueas se pueden presentar formando patrones que dependerán de distintos factores e incluso de la temporada en la que sean examinados.

## 2.4. Geoestadística

La geoestadística es una manera de describir la continuidad espacial de cualquier fenómeno natural, implica un conjunto de herramientas estadísticas que incorporan las coordenadas espaciales a variables de medición observadas en el procesamiento de los datos, consiguiendo la descripción y el modelado de patrones, incluyendo el error y predicción de datos (Goovaerts 1998). Los análisis geoestadísticos se centran principalmente en el desarrollo de funciones de estructura tales como el semivariograma y el correlograma, que son utilizadas posteriormente para interpolar en

el espacio el valor de variables en sitios no muestreados. Con la aplicación de esta herramienta se puede conocer el grado de dependencia espacial (ver 2.4.1 Autocorrelación) y el grado de covarianza que existe entre las variables de interés, el patrón espacial de dichas variables, la escala a la cuál dichos patrones se repiten, la mejor representación gráfica para las variables y el grado de incertidumbre una vez que se ha llevado a cabo la interpolación, todo ello tomando en cuenta la escala espacial que se elige (Gallardo y Maestre 2008). Hasta el momento, varios estudios han demostrado que la aplicación de modelos geoestadísticos a la ecología es una herramienta adecuada para reflejar la estructura espacial de suelos ya sea agrícolas o afectados por la ganadería (Franklin y Mills 2003; Philippot et al. 2009; Wessén et al. 2011) u otras actividades humanas e incluso en suelos extremos (Pasternak et al. 2013). Es necesario considerar algunas funciones de estructura tales como el correlograma y el semivariograma, que llevaran a la predicción:

### 2.4.1. Autocorrelación

La autocorrelación se puede definir como el grado de dependencia espacial de una variable, lo cual se describe concretamente por la primera ley de la geografía, la cual dice: "Todo está relacionado, pero los objetos cercanos están aún más relacionados que los objetos distantes"; así, la autocorrelación espacial positiva indica que lo que sucede en un sitio depende de lo que sucede en su cercanía, y esta aseveración constituye las bases de las técnicas de interpolación. Este término es muy importante ya que si la dependencia espacial es mínima, un mapa (objeto que se pretende lograr en Geoestadística) no mostrará nada relevante. La Autocorrelación espacial es una propiedad importante encontrada en la mayoría de los datos ecológicos y puede presentarse en todas las escalas espaciales. El método general para describir autocorrelación de una variable es calcular algún índice de covarianza para una serie de

distancias y construir correlogramas que ilustran la autocorrelación a cada distancia o pares de distancias.

#### 2.4.1.1. Índice de Moran

En casos univariados cuantitativos, la correlación espacial puede ser estimada utilizando el Índice de Moran (Legendre y Legendre 2012), representado por la siguiente ecuación:

$$I(d) = \frac{\frac{1}{W} \sum_{h=1}^n \sum_{i=1}^n w_{hi} (y_h - \tilde{y})(y_i - \tilde{y})}{\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y})^2} \quad \text{para } h \neq i \quad (2.1)$$

donde  $y_h$  y  $y_i$  son los valores observados en los sitios  $h$  e  $i$ ,  $d$  es la clase de distancia considerada, para la cual antes de realizar los cálculos de coeficiente, es necesario realizar el cálculo de la matriz de distancias  $D = [d_{hi}]$

El cálculo del coeficiente I de Moran, está relacionado con el coeficiente de correlación de Pearson, su numerador corresponde a una covarianza entre los valores que se encuentran en todos los pares de puntos y el denominador es un estimador de máxima similitud de la varianza, a diferencia de el coeficiente  $\rho$  de Pearson, I de Moran sólo involucra una variable y además sólo considera los valores correspondientes a determinadas clases de distancias. Usualmente los valores de I de Moran se encuentran en un intervalo de +1 a -1, donde los valores positivos representan una correlación positiva y viceversa, aunque es posible obtener valores mayores o menores a ese intervalo (Legendre y Legendre 2012).

#### 2.4.1.2. Construcción del correlograma

El correlograma es un gráfico donde los valores de autocorrelación se grafican contra las clases discretizadas de distancias entre los sitios de muestreo. Cuando se

calcula el correlograma espacial se debe asumir que toda el área de estudio presenta la misma estructura espacial. A cada coeficiente de autocorrelación se le realiza una prueba de significancia que puede ser mediante el uso de permutaciones. La hipótesis nula de esta prueba es que el coeficiente no es significativamente diferente de cero (donde los valores cercanos a cero indican la falta de autocorrelación) (Legendre y Legendre 2012).

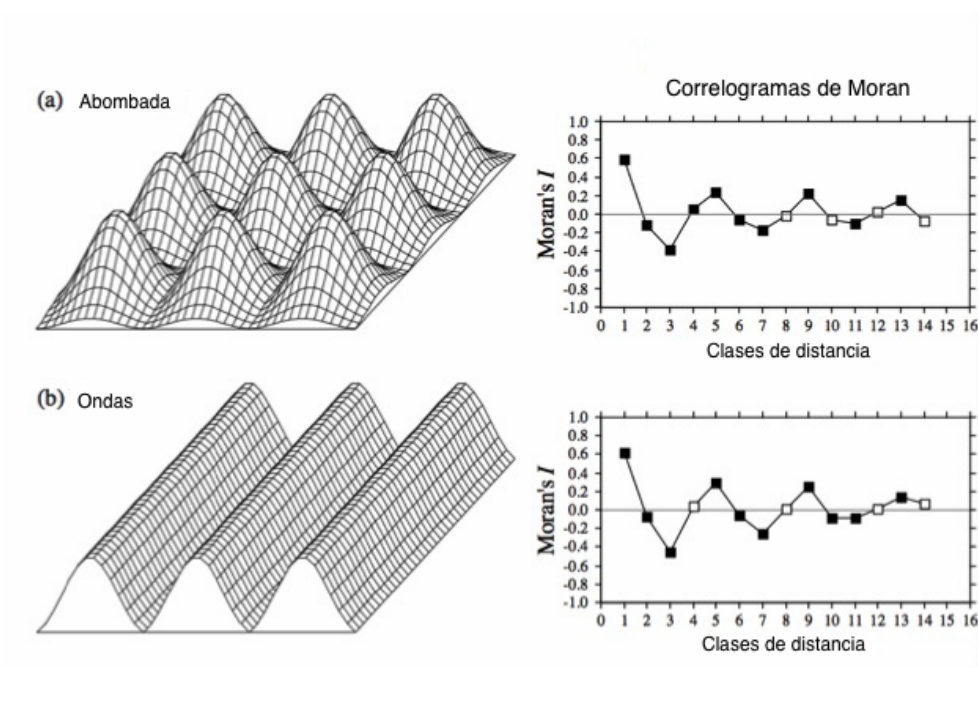


FIGURA 2.1: **Correlogramas de Moran.** Ejemplos de correlogramas de Moran asociados con la correspondiente estructura espacial. Obtenida y modificada de Legendre y Legendre 2012.

Los autocorrelogramas generalmente reflejan la estructura en un sitio dependiendo del comportamiento de las autocorrelaciones en la gráfica, un ejemplo de ello podemos observarlo en la Figura 2.1.

### 2.4.2. Semivariograma

Es la herramienta inicial de la que se vale la geoestadística, y a diferencia del correlograma, describe los patrones en términos de desigualdad entre observaciones en función de la distancia. Está definida por la ecuación descrita a continuación Eq.2.2:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(u_{\alpha}) - z(u_{\alpha} + h)]^2 \quad (2.2)$$

En aplicaciones predictivas, el semivariograma empírico es ajustado a un modelo para describir los valores característicos de los semivariogramas: el sill, nugget y rango. El sill constituye el valor ordenado donde el variograma se estabiliza, el rango es la distancia a la cual esto ocurre y el nugget corresponde a la semivarianza a una distancia cero, un valor diferente a cero puede indicar que existen estructuras a una escala más pequeña de la considerada.

### 2.4.3. Predicción

El objetivo de la mayoría de los estudios geoestadísticos en la ciencia del suelo es estimar las propiedades del mismo en puntos que no han sido muestreados y realizar un mapa de ellos. Los modelos o algoritmos de predicción espacial, tal como lo explica Hengl en 2009, pueden ser clasificados de acuerdo a la cantidad de análisis estadístico, es decir la cantidad de conocimiento experto que se incluya en el análisis:

- (a) **Modelos determinísticos o mecánicos:** En este tipo de modelos se usan parámetros que son arbitrarios o empíricos. No hay ningún tipo de error estimado y no hay ninguna suposición acerca de la variabilidad, en este tipo de modelos entran la interpolación inversa de la distancia y regresión sobre coordenadas, entre otros.

- (b) **Modelos Estadísticos Lineales o de Probabilidad:** En este tipo de modelos, los parámetros son estimados de forma objetiva y siguiendo una teoría de probabilidad, además de que el conjunto de datos de entrada debe cumplir antes con ciertas Hipótesis estadísticas. Las predicciones se acompañan de un estimado del error de predicción. Dentro de este grupo de modelos entran: el Kriging, la correlación ambiental, los modelos Bayesianos y los modelos Híbridos (regresión-kriging).
- (c) **Sistemas Expertos:** Finalmente se encuentran este tipo de modelos que pueden ser subjetivos completamente y por ello a veces resultan irreproducibles al estar principalmente basados en los datos. Modelos de este tipo son los algoritmos de "*Machine Learning*" y redes neuronales, que últimamente están siendo más utilizados.

Existen otras formas de clasificación de algoritmos de predicción que se basan tanto en el tratamiento que se le da a los datos como en el tipo de muestreo y el efecto sobre las predicciones (Hengl 2009). Para efectos prácticos sólo se abordan los métodos más comunes utilizados en ciencia del suelo.

#### 2.4.3.1. Modelos de predicción estadística: Kriging

Existe una familia de algoritmos de regresión por mínimos cuadrados que se denomina Kriging y que es comúnmente utilizado para la interpolación en geoestadística y por consecuencia para la creación de mapas de patrones (Goovaerts 1998; Hengl 2009).

Además de proporcionar una varianza de error, kriging tiene varias ventajas sobre otras técnicas de interpolación (1) Considera anisotropía para el patrón de variabilidad espacial a través del uso del modelo de semivariograma y (2) permite la



estimación de un atributo blanco sobre un soporte distinto al soporte de los datos. (Goovaerts 1998).

Dentro de esta familia de algoritmos para la forma lineal del estimador, existen tres métodos más conocidos: el *Kriging Universal, Simple* y el *Kriging Ordinario*. El **kriging universal** supone que puede llegar a haber una tendencia de invalidación de datos con un número de predictores  $p > 0$ , es decir hay una tendencia que se supone conocida al igual que la función de covarianza (usualmente hay que suponer tanto el polinomio que expresa la tendencia como el variograma). En el caso del **Kriging simple**, se requieren conocer los valores esperados y la función de la covarianza de la función aleatoria, asumiendo una media constante en todo el dominio, omitiendo un término que se asocia con la determinación de error. Y finalmente el método usualmente utilizado aplicado al suelo: El **Kriging Ordinario** (Goovaerts 1998; Hengl 2009; Bivand, Pebesma y Gómez-Rubio 2013a).

**Kriging ordinario (KO)** El Kriging ordinario es una versión estándar del Kriging, cuyo valor esperado se asume estacionario o constante alrededor de cada localidad por punto de estimación, pero se desconoce. Las predicciones están basadas en el modelo Eq. 2.3:

$$Z(s) = \mu + \varepsilon'(s) \quad (2.3)$$

Donde  $\mu$  es la función estacionaria constante, es decir la media global, y  $\varepsilon'(s)$  es la parte estocástica de la variación correlacionada. Las predicciones en el kriging ordinario se llevan a cabo siguiendo la ecuación Eq. 2.4

$$\hat{Z}_{KO}(s_0) = \sum_n^{i=1} w_i(s_0) * z(s_i) = \lambda_0^T * z \quad (2.4)$$

que considera  $\lambda_0$  como el vector de ponderaciones kriging ( $w_i$ ) y a  $z$  como el vector de  $n$  observaciones en sitios primarios. Para el cálculo del vector de ponderaciones, y que éste refleje verdaderamente la condición de autocorrelación espacial se utiliza el modelo ajustado del semivariograma (Eq. 2.2), el cual se utiliza para derivar las semivarianzas alrededor de los sitios y resolver las ponderaciones multiplicando la matriz de covarianzas de los sitios por el vector de covarianza en el sitio nuevo (Eq. 2.5).

$$\lambda_0 = \mathbf{C}^{-1} * \mathbf{c}_0 \quad (2.5)$$

Una técnica de predicción estadística, tal es el caso de KO, produce además de los valores en los nuevos sitios, medidas asociadas a la incertidumbre de la predicción: la varianza estimada del error de predicción. La varianza del KO (Eq.2.6) es definida por el promedio ponderado de las covarianzas del nuevo sitio ( $s_0$ ) a todos los puntos de calibración (sitios de valores conocidos) más un multiplicador que explicaría la razón de cambio de la solución (conocido como *multiplicador de Lagrange* ( $\varphi$ )) (Hengl 2009; Bivand, Pebesma y Gómez-Rubio 2013a).

$$\hat{\sigma}_{KO}^2(\mathbf{s}_0) = (C_0 + C_1) - \mathbf{c}_0^T * \lambda_0 = C_0 + C_1 - \sum_n^{i=1} w_i(\mathbf{s}_0) * C(\mathbf{s}_0, \mathbf{s}_i) + \varphi \quad (2.6)$$

El sistema de ecuaciones se torna complicado de resolver, pero la mayoría de las herramientas geoestadísticas con las que se cuenta actualmente ya hacen dicho cálculo y lo entregan en un mapa y, adicional al mapa de interpolación, el mapa de las variaciones del error.

Es posible utilizar correlaciones con factores ambientales como predictores auxiliares para realizar las interpolaciones, en este tipo de regresión pueden entrar el análisis Fuzzy (continuo) o modelos de regresión lineal generalizada. Sin embargo

la aplicación de estos modelos requieren una gran cantidad de datos para brindar resultados adecuados.

### 2.4.3.2. Selección del modelo predictivo

Es posible realizar una selección previa del modelo que vayamos a aplicar siguiendo un diagrama de decisión, que fue implementando en ciertos programas geoestadísticos tal como el paquete gstat (Pebesma 2004) del entorno R y que podemos consultar de manera clara en la Figura 2.2 obtenida de Hengl 2009

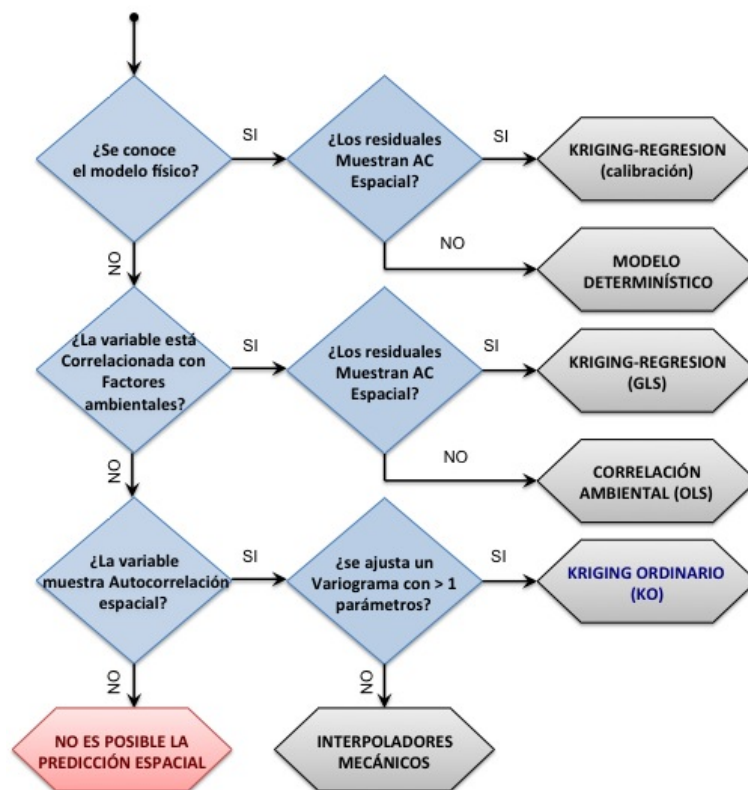


FIGURA 2.2: Diagrama de decisión del modelo de predicción espacial. Obtenida y modificada de: Hengl 2009 Pag.32

### 2.4.3.3. Métodos de validación de la interpolación

Para validar un mapa, es decir asegurar que sea un mapa de calidad, la mejor manera es usar la comparación de valores reales en ciertos puntos, marcados como puntos de validación con los valores estimados por el modelo. Es costoso implementar un muestreo adicional exclusivamente para la validación de la precisión, por lo cual existen métodos que se utilizan comúnmente a fin de salvar dichos obstáculos, entre los cuales el más utilizado es la **validación cruzada**, que divide los puntos de un mismo conjunto de datos en puntos de calibración y validación para posteriormente realizar el análisis (Hengl 2009; Bivand, Pebesma y Gómez-Rubio 2013a). Existen ciertos tipos de validación cruzada:

1. **Validación cruzada por *k-fold***: el conjunto de datos es separado en  $k$  partes iguales y cada una es usada para la validación.
2. **Validación cruzada *Leave-one-out (LOO)***: Cada punto en el conjunto es usado para validar, es decir se realiza la predicción usando todos los puntos excepto uno, continúa con el siguiente y así sucesivamente hasta completar todo el conjunto, asegurando así cada punto contra el conjunto completo de datos.
3. **Validación cruzada por *Jackknifing***: este método es similar a LOO, sin embargo se dirige a definir el sesgo del análisis estadístico y no a las predicciones.

El error de predicción apunta a la *precisión de la predicción* y de acuerdo a dicho error se aplican ciertas medidas o índices para definirla (Hengl 2009; Chabala, Muloowa y Lungu 2017):

- a. El error estándar promedio (*ASE*), que es calculado utilizando la siguiente ecuación:

$$ASE = \sqrt{\frac{1}{N} \sum_{j=1}^N \left[ Z'(s_j) - \left( \sum_{j=1}^N Z'(s_j) / N \right) \right]^2} \quad (2.7)$$

Siendo  $Z'(s_j)$  los valores estimados y  $N$  el número total de puntos de validación.

- b. La raíz cuadrada del error (*RMSE*) que al igual que *ASE* es una medida relevante para medir la precisión y que se calcula de la siguiente forma:

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N [Z(s_j) - Z'(s_j)]^2} \quad (2.8)$$

Siendo, al igual que en *ASE*,  $Z'(s_j)$  los valores estimados,  $N$  el número total de puntos de validación y  $Z(s_j)$  las observaciones en los puntos de validación. Y,

- c. El error *RMSE* puede ser además normalizado ( $RMSE_r$ ) por la variación total:

$$RMSE_r = \frac{RMSE}{s_z} \quad (2.9)$$

donde  $s_z$  es la variación total (desviación estándar).

## 2.5. Herramientas moleculares para el estudio ecológico del suelo

La ecología molecular microbiana tiene su base en la extracción y caracterización de ácidos nucleicos y otros componentes subcelulares. Esta área de la ecología de suelo ha sido rápidamente desarrollada de unos 10 años a la fecha y continúa explorándose.

La extracción de moléculas como el ácido desoxirribonucleico (DNA) y ribonucleico (ARN) de células presentes en las muestras de suelo y su análisis ya sea directo o con el uso de las técnicas de amplificación como la reacción en cadena de la polimerasa (PCR) permitió empezar a detectar y caracterizar una amplia diversidad de microorganismos (Thies 2007).

El objetivo de la mayoría de los análisis moleculares de comunidades es describir la diversidad de poblaciones mediante el cálculo de riqueza, diversidad y uniformidad de Taxas, sin embargo debido al sesgo que ocasiona la extracción de DNA o RNA de las muestras es difícil asegurar la verdadera abundancia, lo cual se puede sobrellevar utilizando por ejemplo diferentes técnicas de extracción (Thies 2007; Islam et al. 2012). Entre las técnicas más comúnmente utilizadas para estudiar diversidad se encuentran: análisis de fosfolípidos de membrana (PLFA), polimorfismo de restricción en la longitud de fragmentos (RFLP), librerías rRNA 16S y pruebas de isótopos entre otras, y actualmente el uso de técnicas de secuenciación ha representado un progreso significativo en este tipo de estudios, su utilidad recae principalmente en determinar secuencias de genes de miembros importantes dentro de las comunidades del suelo, para desarrollar *primers* cada vez más específicos y pruebas que respondan a importantes preguntas ecológicas.

### 2.5.1. Librerías 16S rRNA por pirosecuenciación

Dentro de las técnicas desarrolladas para estudiar el gen 16S rRNA se encuentra la pirosecuenciación, éste método de secuenciación de DNA se basa en la síntesis de dicha molécula, donde intervienen en cascada cuatro enzimas para lograr la obtención de una señal luminométrica que refleja la cantidad de nucleotidos incorporados a un *primer* hibridado al DNA molde (Magurran y McGill 2011).

## 2.6. Métodos para Examinar la Diversidad de Comunidades Microbianas

Cuando los estudios de diversidad de comunidades se llevan a cabo a nivel taxonómico, en el caso de los microorganismos, se conducen sobre las Unidades Taxonómicas Operacionales (OTUs). Debido a la gran cantidad de índices propuestos para cuantificar la biodiversidad en su naturaleza multidimensional, se hace un poco complicado tomar una decisión certera de qué índice es más apropiado utilizar a pesar de las fuertes correlaciones existentes entre dichos índices puesto que muestran aspectos varios del mismo fenómeno. Sin embargo en análisis ecológicos utilizar diferentes medidas de diversidad lleva a un entendimiento más completo de como los cambios en especies raras y abundantes conducen interacciones y de componentes complejos que influyen la diversidad (Morris et al. 2014).

### 2.6.1. Métodos de medición a nivel de especie

Las comunidades ecológicas no son un ente aislado, se encuentran interactuando continuamente con su entorno, y los procesos que se llevan a cabo en dicho entorno no son independientes. Así diferentes comunidades ecológicas serán diferentes también, en el número y tipos de especies que contienen. Para entender las razones por las que esto sucede dependiendo del entorno se pueden medir tres niveles de diversidad: la **Alfa ( $\alpha$ ) diversidad** o diversidad en un sitio, la **Beta ( $\beta$ ) diversidad** diversidad entre sitios y **Gamma ( $\gamma$ ) diversidad** o diversidad a gran escala (Moreno 2001; Magurran y McGill 2011).

### 2.6.1.1. Métodos de diversidad alfa

La diversidad alfa ( $\alpha$ ) es la diversidad dentro de una comunidad o intrínseca y que tiene que ver con la riqueza y abundancia de especies dentro de un sitio o hábitat. Representa la diversidad a lo largo de todas las subunidades (o escalas) relevantes (hábitat). El cálculo de la mayoría de los índices para medir  $\alpha$  diversidad, por tanto, abarca los valores de riqueza y abundancia relativa, pero clasificándose matemáticamente con base a la cuantificación o estructura (dominancia y equidad) de especies en la comunidad (Moreno 2001; Morris et al. 2014).

**2.6.1.1.1. Métodos por cuantificación** La **riqueza específica** es la métrica más simple para representar diversidad e implica sólo el número total de especies o atributos presentes sin que sea de gran impacto su importancia. Es importante otra medida de cuantificación, **la rarefacción** que permite comparar números de especies entre comunidades de tamaños de muestra diferentes (Moreno 2001; Morris et al. 2014).

**2.6.1.1.2. Métodos por estructura de comunidades** En estos métodos la abundancia proporcional de especies puede ser incorporada y abarca métodos paramétricos no paramétricos e índices de abundancia proporcional, estos últimos pueden considerar dominancia o equidad. Los más comunes son:

**a) Chao1:** el índice Chao1 es un método no paramétrico en el que se determina el número total de especies con base en el número de clases raras (OTUs) encontrados en una muestra.

$$Chao1 = S_{obs} + \frac{(f_1)^2}{2f_2} \quad (2.10)$$



dónde  $S_{obs}$  es el número de especies en una muestra y  $f_1$  y  $f_2$  es el número de especies representadas solamente por un único individuo (*Singletons*) y el número de especies representadas por dos (*Dobletons*) respectivamente (Moreno 2001; Magurran y McGill 2011).

- b) Simpson:** El índice Simpson se encuentra dentro de los índices de dominancia, es decir, es sensible a especies abundantes y representa la probabilidad de que dos individuos elegidos aleatoriamente pertenezcan a diferentes especies y plantea que un sistema es más diverso cuanto menos dominancia de especies haya. Se puede representar con la ecuación:

$$Simpson(D1) = 1 - \sum P_i^2 \quad (2.11)$$

Donde  $P_i$  corresponde a la abundancia proporcional de la especie  $i$  con respecto al número total de individuos en la muestra, este índice se encuentra sumamente relacionado con el siguiente (Moreno 2001; Morris et al. 2014).

- c) Simpson recíproco:** El simpson recíproco es el inverso del índice original de Simpson, dicha transformación sirve para hacerlo intuitivo, es decir que incrementa con la diversidad (Magurran y McGill 2011; Morris et al. 2014), su ecuación es:

$$Simpson\ recíproco(D2) = \frac{1}{\sum P_i^2} \quad (2.12)$$

- d) Simpson E: (Uniformidad),** representa el grado al cual los individuos se reparten entre especies se calcula a partir de Simpson recíproco (Eq. 2.13) y es matemáticamente independiente. Un simpson E menor es indicativo de que una o pocas

especies dominan en la muestra (Magurran y McGill 2011; Morris et al. 2014).

$$Simpson\ Evenness(E) = \frac{D2}{S} \quad (2.13)$$

Siendo  $D2$  la dominancia por Simpson recíproco y  $S$  la riqueza.

e) **Shannon:** En el caso de los índices de equidad se encuentra el índice Shannon, que representa la incertidumbre de identidad de un individuo desconocido, asumiendo que todas las especies están representadas en la muestra. Toma en cuenta tanto abundancia como similitud de especies presentes por lo cual es sensible tanto a especies raras como a especies abundantes. Se calcula a partir de Eq. 2.14

$$Shannon = - \sum_{i=1}^z P_i * \ln(P_i) \quad (2.14)$$

De igual forma que para Simpson ( $D1$ ),  $P_i$  corresponde a la abundancia proporcional de la especie  $i$  con respecto al número total de individuos (Moreno 2001; Magurran y McGill 2011; Morris et al. 2014).

f) **PD (Diversidad Filogenética):** Es la medición de la diversidad dentro de una sola comunidad considerando únicamente presencia/ausencia de Taxas. Simplemente se calcula como la suma de la longitud de las ramas de un árbol filogenético el cual es cubierto o representado en una muestra dada (Cadotte et al. 2010).

#### 2.6.1.2. Métodos de diversidad Beta

La meta principal de la mayoría de los estudios de comunidades es comparar la composición de diferentes comunidades y que difieren en algún parámetro o tratamiento (por ejemplo diferencia en Salinidad). La diversidad beta ( $\beta$ ) se evalúa con

base en índices o coeficientes de similitud, de disimilitud o de distancia entre muestras a partir de datos cuantitativos o cualitativos (Moreno 2001; Magurran y McGill 2011). Dentro de estas métricas tenemos los métodos de ordenación y clasificación y cualitativos (Jaccard) o cuantitativos (Sorenson).

**2.6.1.2.1. UniFrac** La métrica UniFrac (Unique Fraction Metric) incluye las distancias filogenéticas en un árbol filogenético entre dos comunidades, el cálculo incluye información sobre el parentesco de miembros en una comunidad (Lozupone et al. 2007). Dentro de esta métrica podemos aplicar *weighted UniFrac* que divide el largo total de la longitud de las ramas entre las diferentes comunidades y la abundancia relativa de los OTUs entre muestras, mientras que *Unweighted UniFrac* considera solo la riqueza.

## Capítulo 3

# Justificación y Objetivos

*«Bacteria are the dark matter of the biological world with 4 million mostly unknown species in a ton of soil»*

---

*Edward O. Wilson*

### 3.1. Justificación

El suelo del exlago de Texcoco tiene características que lo hacen un ambiente único y altamente heterogéneo, con valores de pH mayores de 10 y conductividades electrolíticas que alcanzan y sobrepasan los  $150 \text{ dS m}^{-1}$ , se han conducido estudios para conocer las comunidades bacterianas y de arqueas que se pueden encontrar en la zona, pero no se han relacionado en escala "landscape" a las características propias del suelo que puedan llevar a un modelo predictivo y posteriormente a la determinación de la contribución biogeoquímica de las comunidades en el área y/o a una estrategia de manejo del ecosistema

## **3.2. Objetivos**

### **3.2.1. Objetivo General**

«Describir patrones espaciales de comunidades microbianas en suelo haloalcalino altamente heterogéneo»

### **3.2.2. Objetivos Particulares**

Objetivo 1. Determinar el grado de autocorrelación espacial de los parámetros fisicoquímicos medidos en el suelo.

Objetivo 2. Determinar la estructura espacial de los parámetros fisicoquímicos medidos a lo largo del transecto.

Objetivo 3. Describir la distribución de comunidades microbianas en suelos haloalcalinos extremos.

Objetivo 4. Determinar características Fisicoquímicas y/o ambientales del suelo que conducen la variación de comunidades microbianas.

# Capítulo 4

## Materiales y Métodos

*«If I set out to prove something I am  
no real scientist– I have to learn to  
follow where the facts lead me– I have  
to learn to whip my prejudices...»*

---

*Spallazani*

### 4.1. Sitio experimental y muestreo

El sitio experimental se eligió de acuerdo a exámenes previos del área del exlago de Texcoco (19.30° N, 98.53° O) en los que se observaron valores altos de pH y conductividad electrolítica. El contenido de sal en el exlago de Texcoco es resultado del afloramiento de aguas subterráneas salobres, infiltración de parte del agua, evaporación y deposición de sal especialmente durante la temporada de lluvias. Algunas partes de la cama del lago han sido drenadas, lo que contribuye a las variaciones encontradas en el contenido de sal en el sitio (Dendooven et al. 2010). *Distichlis spicata* L., un pasto resistente a la sal, puede ser encontrado en parches en la cama del lago. Forma extensiones sobre el suelo cuando el contenido de sal es muy alto.

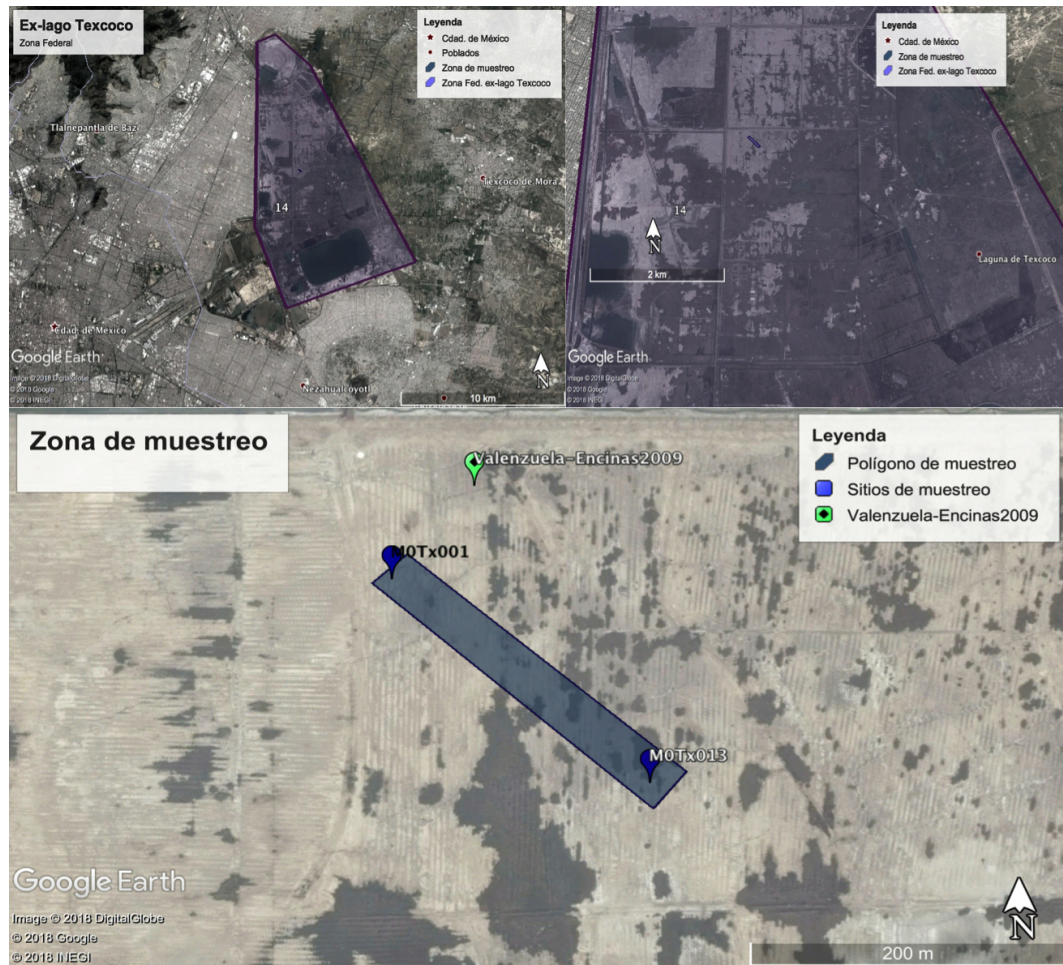


FIGURA 4.1: Localización geográfica del sitio de muestreo *Zona Federal del ex-lago de Texcoco*

### 4.1.1. Muestreo

La temperatura media anual en el sitio es de 15.9 °C, con heladas poco frecuentes y una precipitación pluvial media anual de 686 mm. El muestreo fue realizado el día 22 de Agosto de 2014 (la temperatura registrada en el área fue de 19 °C), después de un periodo de alta precipitación pluvial (precipitación promedio en el mes de agosto 189.2 mm, según datos del SMN <http://smn.cna.gob.mx/es/climatologia>).

Se realizó un muestreo en transecto, definido mediante observación directa de las características del suelo, fueron tomados 12 puntos a diferentes distancias de

un punto seleccionado aleatoriamente, la profundidad de muestreo fue de 5-20 cm aproximadamente. La distancia total aproximada de muestreo fue de 211 m. Para las lecturas mediante GPS (eTrex Vista® C Garmin, Olathe, Kansas, USA) en cada punto se usó el datum correspondiente a NAD27 México. Las muestras fueron almacenadas a 4 °C para su análisis posterior.

## 4.2. Caracterización fisicoquímica del suelo

Las muestras fueron cribadas mediante tamices de 5 y 2 mm mesh y secadas en invernadero durante 24 h. Fueron determinados los siguientes parámetros: pH, Conductividad electrolítica (EC), contenido de agua (H), Textura (Txt), Carbono Total (CT), Carbono inorgánico (Ci) y se calculó Carbono Orgánico (Co). Todos los métodos para caracterización fisicoquímica han sido estandarizados en el laboratorio de ecología molecular de suelos del CINVESTAV.

### 4.2.1. Determinación de pH

Para la determinación de pH fue necesario pesar 10 g de suelo y agregar agua destilada en una proporción de 1:2.5 (p:v), se mezcló la suspensión preparada durante 10min, después de los cuales se dejó reposar durante 10 min más, transcurrido el tiempo se tomó la lectura con potenciómetro (UltraBasic Benchtop UB-10, Denver instruments) previamente calibrado.



### 4.2.2. Determinación de conductividad electrolítica (EC)

Para la medición de la conductividad electrolítica se pesaron 200 g de suelo, al cual se agregó agua destilada hasta el punto de saturación, detectado con la obtención de una consistencia pastosa y brillante. Las muestras así preparadas, se refrigeran a 4 °C durante 24 h, tras las cuales se procede a centrifugar (centrífuga Sorvall ST-40 R) durante 45 min a 2739 xg, se recuperó el sobrenadante y se midió la EC utilizando un conductímetro (Hanna Instruments HI-2300).

### 4.2.3. Determinación del contenido de agua

Para la determinación del contenido de agua (H) se utilizó un método gravimétrico, en el cual fueron pesados 10 g de suelo en recipiente de aluminio previamente pesado, las muestras fueron colocadas a secar en horno a 105 °C durante 24 h, transcurrido el periodo de secado, las muestras fueron pesadas nuevamente y fue obtenido el peso del suelo anhidro mediante sustracción. Para la obtención del %H se aplicó la siguiente ecuación:

$$\%H = \frac{P_s - P_{sanhidro}(105\text{ }^\circ\text{C})}{P_s} * 100 \quad (4.1)$$

### 4.2.4. Determinación de la clase textural

La textura se determina en función de la velocidad de sedimentación de partículas, para lo cual, en el vaso de una batidora fueron agregados 50 g de suelo seco y tamizado con 200 ml de agua destilada y, como dispersante, 10 ml de Hexametáfosfato de Sodio (0.08 M), dicha mezcla se agitó durante 10 min y fue vaciada en una probeta de 1 l, la cual se aforó con agua destilada. Una vez aforada se sometió a 10 ciclos de agitación, y se colocó en una superficie plana y estable, transcurridos 40 seg, utilizando el Hidrómetro de Bouyoucos y un termómetro, fue tomada una primera

lectura. Se dejó reposar durante 2 h, tras las cuales se tomó una segunda lectura. Los porcentajes resultantes de la aplicación de fórmulas pertinentes para limos, arcillas y arena (Ecuaciones 4.2, 4.3 y 4.4), se utilizan para la determinación de la clase textural con base en el triángulo de textura (Fig. 4.2) para lo cual también se utilizó el paquete SoilTexture (Moeys 2016) dentro del ambiente R.

$$\%limo + \%arcilla = \frac{1era.lectura + (T1 - 20) * 0,36}{P_{ss}} * 100 \quad (4.2)$$

$$\%arcilla = \frac{2da.lectura + (T2 - 20) * 0,36}{P_{ss}} * 100 \quad (4.3)$$

$$\%arena = 100 - (\%limo + \%arcilla) \quad (4.4)$$

Dónde  $P_{ss}$  es el Peso del suelo seco,  $T1$  y  $T2$  son la primera y la segunda temperaturas tomadas, respectivamente.

#### 4.2.5. Determinación de Carbono Total, inorgánico y orgánico

Para la determinación de **Carbono Total (CT)** se pesaron alrededor de 200 mg de suelo seco y se colocaron al fondo de un tubo contenedor de 2.5x25, al cuál se adicionaron 10 ml de mezcla digestora (mezcla crómica). Se coloca una varilla de vidrio como soporte para un tubo de rosca interno que contiene 20 ml de NaOH 1 M, que fue colocado inmediatamente después de la mezcla digestora. Una vez dispuesto el tubo con NaOH se cubre el tubo contenedor con una tapa subbaseal procurando que selle perfectamente (Fig. 4.3A). Se calienta hasta 130 °C poco a poco y se mantiene a esta temperatura hasta que se dejen de observar burbujas en la mezcla digestora. Simultáneamente se colocan al menos tres blancos, que corresponden solamente al tubo con NaOH dentro del tubo contenedor. Una vez retirado el calentamiento, se

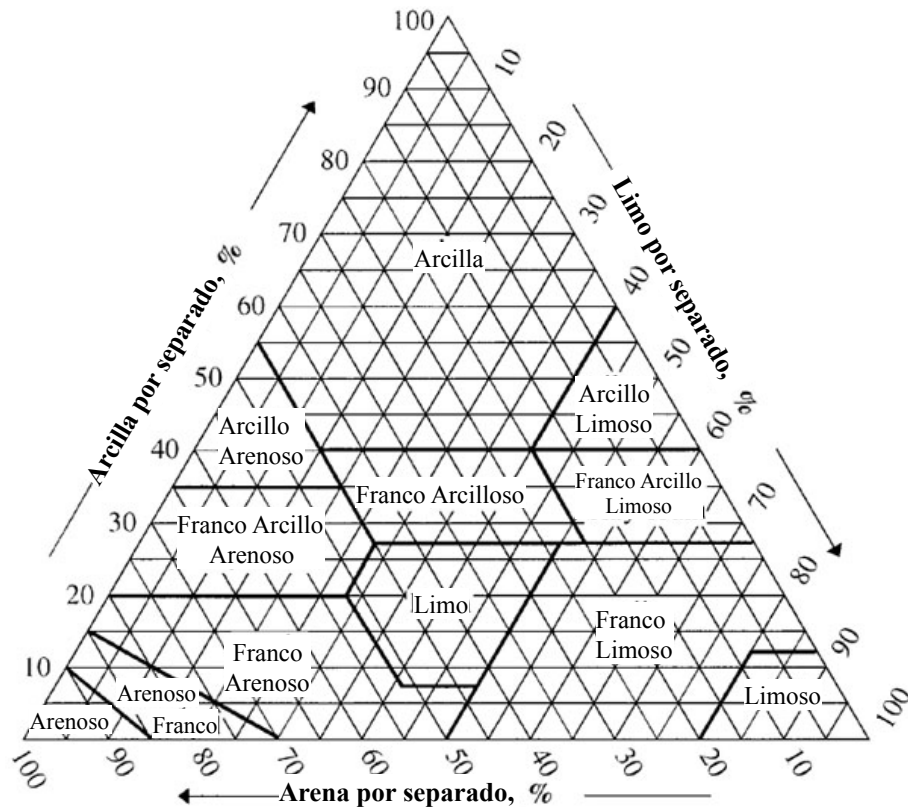


FIGURA 4.2: Triángulo de evaluación de la clase textural de acuerdo al porcentaje de componentes obtenido de: *U.S. Department of Agriculture*

dejan reposar por 24 h, transcurrido el tiempo de reposo se procede a titular la solución de NaOH utilizando HCl hasta llegar a pH 4.0, para la lo que se utilizó un titulador automático (Metrohm 716 DMS Titrino). Teniendo los volúmenes de titulación se aplica la Ecuación 4.5 (Amato 1983).

$$mgCKg^{-1}ss = \frac{(V_{HCl0.1N} - V_{blanco}) * N_{HCl} * F.Alicuota * 12}{P_{ss}} * 1000 \quad (4.5)$$

Para la determinación de **Carbono inorgánico (Ci)** fueron pesadas muestras de alrededor de 1 g de suelo seco y colocadas en contenedores medianos, el contenedor con la muestra fue colocado en un frasco de vidrio de 1 l junto con un recipiente ámbar pequeño conteniendo 20 ml de NaOH 1 M, el recipiente de vidrio fue cerrado

utilizando tapas con septas, procurando sellarlas perfectamente. A través de la septa se inyectaron 5 ml de HCl 5 N en el contenedor con la muestra de suelo (Fig. 4.3B) y permaneció en incubación durante 24 h a temperatura ambiente. Transcurrido el tiempo de incubación, se obtuvo la solución de NaOH, de la cual se tomó una alícuota de 5 ml y se procedió a titular, al igual que para la determinación de Ci. Con los volúmenes de titulación fue calculada la cantidad de Carbono inorgánico en la muestra Eq. 4.5 (Tinsley, Taylor y Moore 1951; Bundy y Bremner 1972).

Para obtener la cantidad de **Carbono orgánico (Co)** basta con calcularla por sustracción a partir del Carbono total e inorgánico.

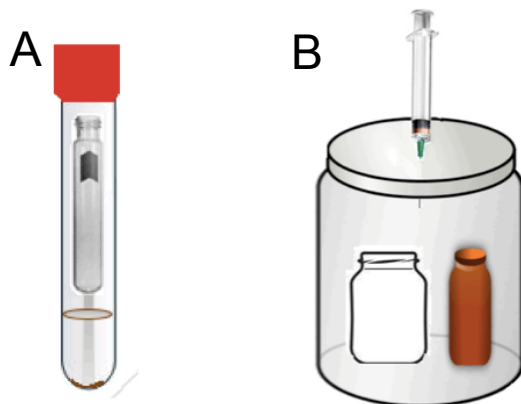


FIGURA 4.3: Modelo de la técnica de evaluación de (A) Carbono Total y (B) Carbono inorgánico

### 4.3. Extracción de DNA metagenómico del suelo

Las muestras de suelo, almacenadas a 4 °C, fueron sometidas a preincubación durante siete días a temperatura ambiente, en condiciones de humedad constante y aerobiosis. Se utilizaron 0.5 g de las muestras de suelo preincubadas y se siguió el método descrito por Ceja-Navarro et al., en el 2009, para la eliminación de materia orgánica. Se usó una solución de pirofosfato 0.15 M y se sometieron las muestras

a centrifugación (2739 xg, 10 min) hasta conseguir un sobrenadante transparente, y buffer de Fosfatos (pH 8.0) para retirar el exceso de pirofosfato utilizado. Para la extracción de DNA metagenómico fueron usados tres métodos diferentes, el primero propuesto por Valenzuela-Encinas et al., en el 2008, se basa en la aplicación de choque térmico, el segundo, propuesto por Sambrook y Russell en el 2001, implica la disrupción celular por medio de lisozima (10 mg ml<sup>-1</sup>) y el último, propuesto por Winston y Hoffman en 1987, aplica la lisis por medio de detergentes en conjunto con lisis mecánica.

## 4.4. Amplificación y Preparación de Librería de Genes rRNA 16S

### 4.4.1. Bacterias

Para Bacterias se eligió amplificar la región V1-V6 (900 pb) del gen 16S rRNA utilizando los oligonucleótidos 8-F (5'-AGA GTT TGA TCI TGG CTC A- 3') y 949-R (5'-CCG TCW ATT KCT TTG AGT T-3'), conteniendo los adaptadores A y B 454 FLX para pirosecuenciación y una secuencia única de 10 pb correspondientes a la etiqueta incorporada a dichos adaptadores (Navarro-Noya et al. 2013). El programa de amplificación fue establecido a 25 ciclos de 45 seg de desnaturalización a 94 °C, 45 seg de alineamiento a 50.2 °C y 1 min 30 seg de extensión a 72 °C usando un termociclador Touchgene gradient de Techne®. Cantidades iguales del producto de cinco reacciones fueron purificados usando el kit DNA Clean & Concentrator siguiendo las recomendaciones del fabricante (Zymo Research, Irvine, CA, USA), y posteriormente fueron cuantificadas usando el ensayo PicoGreen® dsDNA (Invitrogen, Carlsbad, Ca, USA) en el Fluoroespectrofotómetro NanoDrop® 3300 (Thermo Scientific NanoDrop).

#### 4.4.2. Arqueas

Para la amplificación a partir de Arqueas, la región V1-V3 del gen 16S rRNA fue establecida usando los oligonucleótidos etiquetados a 10-pb A-25-F (5'- CYG GTT GAT CCT GCC RG-3') y A-571-R (5'- GCT ACG GNY SCT TTA RGC-3') conteniendo los adaptadores correspondientes 454 FLX A y B. El programa de amplificación fue establecido a 30 ciclos con una desnaturalización de 40 seg a 94 °C, alineamiento de 45 seg a 51.9 °C y un tiempo de extensión de 1 min 30 seg a 72 °C. Los productos de amplificación fueron tratados como para los productos de amplificación bacteriana para su posterior secuenciamiento.

#### 4.4.3. Secuenciamiento

La secuenciación fue realizada por MacroGen Inc. (DNA Sequencing Service, Seoul, Korea) usando un sistema de pirosecuenciación Roche 454 GS-FLX Plus (Roche, Mannheim, Germany).

### 4.5. Determinación de estructura y la diversidad de comunidades

El análisis de los datos de pirosecuenciación fue realizado usando el pipeline del software QIIME version 1.8.0 (Caporaso et al. [2010b](#)). Como primer paso las secuencias de pobre y baja calidad fueron eliminadas del conjunto de datos, para bacterias la longitud establecida fue  $\leq 930$  pb, con un score de calidad de  $\leq 25$ , conteniendo  $\leq 6$  homopolímeros, y un mínimo de dos errores en etiquetas. Para arqueas la longitud se estableció a  $\leq 530$  pb, con un score de calidad de  $\leq 25$ , conteniendo  $\leq 6$  homopolímeros y con un mínimo de 2 errores en etiquetas de oligonucleótidos. Las OTUs fueron determinadas a 97% de similitud con un marco de

referencia abierto usado por Uclust (Edgar 2010) y contra la base de datos de Greengenes (<http://greengenes.lbl.gov>). Las secuencias representativas fueron alineadas con PyNAST a 75 % de identidad mínima (Caporaso et al. 2010a). Las secuencias quiméricas fueron identificadas y removidas del conjunto de datos utilizando ChimeraSlayer (Haas et al. 2011). La asignación taxonómica y las abundancias relativas fueron obtenidas por clasificación naïve Bayesian rRNA del Ribosomal Data Project (<http://rdp.cme.msu.edu/classifier/classifier.jsp>) (Wang et al. 2007) con un intervalo de confianza del 80 % y removiendo *singletons*. Las matrices de observación Biológica BIOM fueron construidas a partir de las asignaciones taxonómicas y los metadatos en las muestras. Las distribuciones de los niveles taxonómicos fueron calculadas usando las matrices BIOM.

## 4.6. Asignación manual de secuencias

Para la asignación taxonómica de OTUs, las secuencias representativas fueron manualmente corroboradas contra secuencias de referencia depositadas en las bases de datos de RDP (<http://rdp.cme.msu.edu>), GenBank (<http://blast.ncbi.nlm.nih.gov>) y EzBioCloud (<http://www.ezbiocloud.net/identify>)

## 4.7. Análisis filogenético

Un análisis de coordenadas principales (PCoA) usando ambas distancias UniFrac (Weighted y Unweighted) fue realizado para comparar las comunidades a través del transecto, además de otros índices de Beta diversidad (Bin-Jaccard, Bray-Curtis y Euclidean, Apéndice D.5).

## 4.8. Análisis estadístico y geoestadístico

Para describir patrones espaciales se plantea un esquema experimental que toma como punto central el uso de funciones Geoestadísticas, para lo cual es necesario conocer la estadística descriptiva de los datos.

### 4.8.1. Estadística descriptiva

Los estadísticos descriptivos fueron determinados para cada parámetro fisicoquímico en el transecto utilizando el entorno R (RStudio Team 2012; R Core Team 2016). Los siguientes análisis estadísticos fueron aplicados: mínimo, máximo, cuartiles, desviación estándar, varianza y pruebas de normalidad (Kolmogorov-Smirnoff, Shapiro-Wilks y Q-Q plots). El sesgo y Kurtosis fueron calculados y gráficas de box-plot y de tendencia fueron realizadas para cada factor.

### 4.8.2. Análisis geoestadísticos

Los análisis exploratorios geográficos fueron realizados utilizando los paquetes GeoR, gstat y sp dentro del entorno de R para determinar la situación espacial de cada factor fisicoquímico (valores en cada punto de muestreo) (Pebesma 2004; Pebesma y Bivand 2005; Bivand, Pebesma y Gómez-Rubio 2013b; Ribeiro y Diggle 2016). Los autocorrelogramas (Eq. 2.1) fueron determinados con el paquete ncf (Bjørnsstad 2016) usando 1000 permutaciones. Los variogramas omnidireccionales (Eq. 2.2), fueron calculados utilizando los paquetes geoR y gstat. Un semivariograma de nube (geoR), fue utilizado para identificar tendencias o valores atípicos y un semivariograma agrupado fue calculado para realizar el ajuste posterior. El modelado del semivariograma, el cual busca caracterizarlo describiendo cómo los parámetros variaron con la distancia, fue realizado utilizando el paquete gstat.



#### 4.8.2.1. Interpolación

El método de interpolación usado fue la forma estándar del kriging, kriging ordinario (Isaaks y Srivastava 1989; Goovaerts 1998; Hengl 2009) con la cual las predicciones fueron obtenidas (Eq.2.4). El kriging fue calculado utilizando el paquete gstat en dos polígonos, uno de 0.71 Ha ( $7141m^2$ ) compuesto de 3000 pixeles, cada uno de  $2.38 m^2$  alrededor del transecto de muestreo y otro de 0.39 Ha ( $3870m^2$ ) compuesto también por 3000 pixeles cada uno de  $1.29 m^2$ , usando el modelo de los semivariogramas ajustados para los factores fisicoquímicos seleccionados.

#### 4.8.2.2. Validación de la interpolación

La validación fue realizada utilizando el método de validación cruzada "leave-one out"(LOO) en el paquete gstat (Pebesma 2004; Gräler, Pebesma y Heuvelink 2016; Hengl 2009). La validación cruzada por LOO se aplicó para cada punto en el conjunto de datos, uno a la vez. Se calcularon el Error promedio estándar ( $ASE$ , Eq. 2.7), raíz cuadrada del error ( $RMSE$ , Eq. 2.8) y el error normalizado ( $RMSE_r$ , Eq. 2.9) por validación cruzada de los valores predichos contra los valores observados en cada sitio. La Interpolación por kriging es considerada buena cuando los valores de  $ASE$  y  $RMSE$ , que aseguran la variabilidad de los valores predichos, son similares. El  $RMSE$  normalizado describe la consistencia del modelo, explicando la variación de parámetros específicos. Un valor  $RMSE_r$  cercano a 40 % asegura una predicción adecuada (Hengl 2009; Chabala, Mulolwa y Lungu 2017).

#### 4.8.3. Determinación de conductores fisicoquímicos

Las correlaciones múltiples fueron realizadas utilizando el paquete ncf tanto para seleccionar factores como para determinar las correlaciones entre ellos. Un análisis

de componentes principales (PCA) fue realizado utilizando los paquetes FactoMineR y FactoExtra (Lê, Josse y Husson 2008; Kassambara y Mundt 2017) mientras que los heatmap fueron realizados utilizando el paquete ComplexHeatmap (Gu, R. y M. 2016).

Las relaciones entre los grupos bacterianos y las características fisicoquímicas fueron exploradas con un análisis canónico de coordenadas principales (CAP) en el paquete Vegan (Oksanen et al. 2017). El *Spearman Test* fue utilizado para calcular las correlaciones entre grupos taxonómicos y las características del suelo, usando los paquetes Hmisc y Corrplot (Harrell Jr y Dupont 2016; Taiyun y Simko 2016).

El *False discovery rate* (FDR), Selección de variables correlacionadas (corSelect) y la detección de multicolinealidad (multicol) fueron realizadas con el paquete Fuzzysim (Barbosa 2016) y la función BIOENV en el paquete vegan (Oksanen et al. 2017) fueron usadas primero para seleccionar las variables con la menor multicolinealidad, las que mostraron mejor relaciones directas con las comunidades y finalmente aquellas que en conjunto afectaron más los grupos taxonómicos.

## 4.9. Accesibilidad de datos

Las secuencias obtenidas en este estudio fueron enviadas al SRA en NCBI, asociadas al Bioproject ID PRJNA414475 bajo los siguientes números de acceso: [SAMN07840001] a [SAMN07840013] para Arquea y de [SAMN07840024] a [SAMN07840035] para Bacteria.

# Capítulo 5

## Resultados y Discusión

*«Science, my lad, is made up of mistakes, but they are mistakes which it is useful to make, because they lead little by little to the truth.»*

---

*Jules Verne*

### 5.1. Características fisicoquímicas del suelo

Los resultados obtenidos de la caracterización fisicoquímica a lo largo del transecto de muestreo muestran poca variabilidad en los valores de pH cuyo rango se encuentra entre 10.3 a 10.6, por otro lado se observa gran variabilidad en los valores de EC (de 7.7 a 179.8 dS m<sup>-1</sup>, y contenido de agua (H) entre 133 y 555 g Kg<sup>-1</sup>, se observa también variación ligera, en la cantidad de CT e inorgánico y por consecuencia también en la cantidad de Co cuyos valores oscilan de 4.5 - 26.3 g Kg<sup>-1</sup> de suelo seco (Tabla 5.1).

La distribución del tamaño de partícula mostró menos variación a lo largo del transecto de muestreo y la textura correspondiente fue de Franco Arcillo Arenosa a Franco Arcillosa, según la clasificación textural del Departamento de Agricultura de

TABLA 5.1: Características Fisicoquímicas a lo largo del transecto de muestreo en la cama del exlago de Texcoco. Descripción del suelo para cada sitio de muestreo establecido a distancias irregulares a lo largo de un transecto al sureste en el área del exlago de Texcoco y análisis estadísticos descriptivos para cada parámetro medido en conjunto. Cada sitio de muestreo se refiere con la clave "Tx"seguido por un número consecutivo indicando progresión en separación espacial. La Fecha de muestreo fue establecida el 22-08-2014

Sitio	Localizacion NAD25		pH	H <sup>a</sup> (%)	EC <sup>b</sup> (dS m <sup>-1</sup> )	Ci <sup>c</sup> (g kg <sup>-1</sup> ss <sup>e</sup> )	Co <sup>d</sup> (g kg <sup>-1</sup> ss)
	Lat	Long					
Tx-001	19d30'45 N	98d29'24.8W	10.4	55.52	66.5	6.8	17.4
Tx-002	19d30'44.9 N	98d29'24.8W	10.5	42.24	84.1	8.2	15
Tx-003	19d30'44.9 N	98d29'24.8W	10.6	43.56	45.1	5.3	10
Tx-004	19d30'44.8 N	98d29'24.7W	10.5	47.33	81.2	7.4	17
Tx-005	19d30'44.5 N	98d29'24.3W	10.5	30.17	142	6	9.2
Tx-006	19d30'44.1 N	98d29'23.8W	10.5	33.22	46.7	3.7	5.7
Tx-007	19d30'43.1 N	98d29'22.8W	10.5	49.5	69.3	7.6	11.3
Tx-008	19d30'42.3 N	98d29'21.4W	10.6	19.47	18	5.8	4.5
Tx-009	19d30'41.3 N	98d29'20.3W	10.5	43.76	129.8	11.4	12.5
Tx-010	19d30'41.3 N	98d29'20.3W	10.6	27.54	179.8	6.3	12.6
Tx-011	19d30'41.3 N	98d29'20.2W	10.4	47.18	113.3	8.1	15.4
Tx-012	19d30'40.9 N	98d29'19.7W	10.3	13.33	7.7	5.1	12.5
Tx-013	19d30'40.5 N	98d29'19.3W	10.4	50.67	145.9	10	26.3
<b>Min</b>			10.31	13.33	7.7	3.7	4.5
<b>Max</b>			10.62	55.52	179.8	11.4	26.3
<b>Var<sup>f</sup></b>			0.009	165.97	2736.5	31.1	43.56
<b>SD<sup>g</sup></b>			0.095	128.8	52.3	2.1	5.6

<sup>a</sup>H: % agua; <sup>b</sup>EC: Conductividad electrolítica; <sup>c</sup>ss:suelo seco; <sup>d</sup>Ci: Contenido de C inorgánico;

<sup>e</sup>Co: Contenido de C orgánico <sup>f</sup>var: varianza; <sup>g</sup>SD: Desviación estándar

TABLA 5.2: **Clasificación Textural.** Tabla de sitios (con las correspondientes transformaciones de coordenadas) distribución de tamaños de partícula y clasificación textural de acuerdo a la clasificación USDA (Figs. 4.2 y 5.1)

Sitio	Lat <sup>a</sup>	Long <sup>b</sup>	E <sup>c</sup>	N <sup>d</sup>	Arena (g Kg <sup>-1</sup> )	Limo (g Kg <sup>-1</sup> )	Arcilla (g Kg <sup>-1</sup> )	Clase textural
Tx 001	19.5125	-98.99022	501026.152	2157536.056	61.63	12.75	25.63	Franco arcillo arenosa
Tx 002	19.51247	-98.99022	501026.152	2157532.737	51.01	16.76	32.23	Franco arcillo arenosa
Tx 003	19.51247	-98.99022	501026.152	2157532.737	47.07	23.96	28.97	Franco arcillo arenosa
Tx 004	19.51244	-98.99019	501029.3	2157529.417	51.07	15.67	33.27	Franco arcillo arenosa
Tx 005	19.51236	-98.99008	501040.842	2157520.566	33.81	31.55	34.64	Franco Arcillosa
Tx 006	19.51225	-98.98994	501055.532	2157508.394	61.07	17.53	21.4	Franco arcillo arenosa
Tx 007	19.51197	-98.98967	501083.863	2157477.413	56.16	15.87	27.97	Franco arcillo arenosa
Tx 008	19.51175	-98.98928	501124.785	2157453.072	54.85	25.95	19.2	Franco arenosa
Tx 009	19.51147	-98.98897	501157.313	2157422.087	54.61	16.2	29.19	Franco arcillo arenosa
Tx 010	19.51147	-98.98897	501157.313	2157422.091	38.96	30.37	30.67	Franco Arcillosa
Tx 011	19.51147	-98.98894	501157.313	2157422.08	59.73	10.89	29.37	Franco arcillo arenosa
Tx 012	19.51136	-98.98881	501174.102	2157409.92	33.12	34.97	31.91	Franco Arcillosa
Tx 013	19.51125	-98.98869	501186.694	2157397.749	52.59	19.88	27.53	Franco arcillo arenosa

<sup>a</sup>Lat: Latitud (en formato decimal); <sup>b</sup>Long: Longitud (en formato decimal); <sup>c</sup>E: Easting; <sup>d</sup>N: Northing (UTM 14N)

Estados Unidos. Los datos para cada sitio pueden ser observados en la tabla 5.2 y visualmente corroboradas en la Fig. 5.1.

### 5.1.1. Análisis descriptivos

La estadística descriptiva realizada para cada parámetro fisicoquímico medido proporcionó conocimiento del conjunto de datos a través del transecto de muestreo sin involucrar el estado espacial de cada parámetro.

Los estadísticos obtenidos fueron los siguientes: mínimos, máximos, cuartiles, Desviación estándar, varianza y pruebas de normalidad (Kolmogorov-Smirnoff y Shapiro-Wilks, Q-Q plots), así como gráficos para descartar la presencia de tendencia en los datos, algunos de los datos estadísticos pueden ser observados en la tabla 5.1, información más detallada se encuentra en la sección de Apéndice (Ap. A). Las

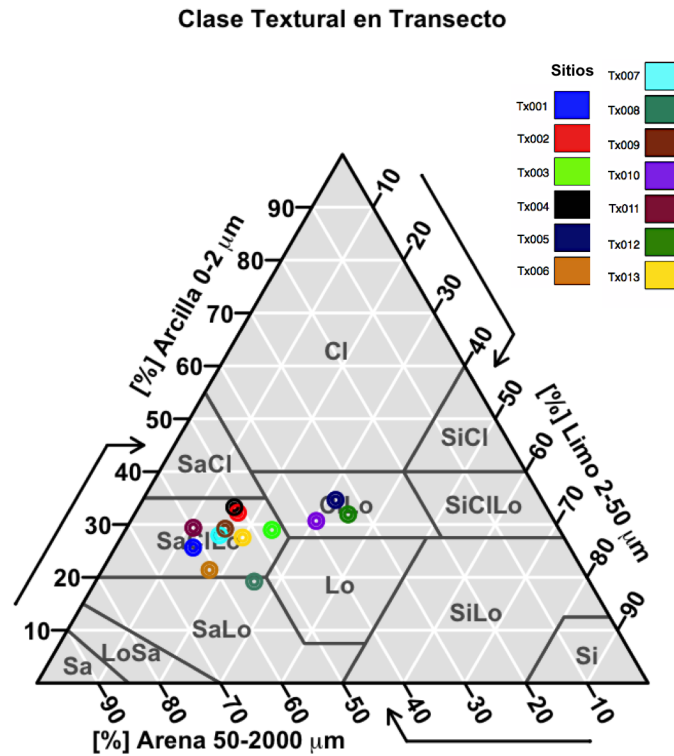


FIGURA 5.1: Observación de la clase textural en cada sitio del transecto según la clasificación USDA

pruebas de normalidad, especialmente Shapiro-wilk y los Q-Q plots, muestran que la distribución de cada parámetro medido a lo largo del transecto es aproximadamente normal, así mismo no se encontraron tendencias en ninguna de las variables determinadas.

Los gráficos Box-plot correspondientes a cada parámetro fisicoquímico y que facilitan la descripción estadística se encuentran en la sección Apéndice (Ap. A.1). Al realizar los análisis descriptivos se deja de lado la locación espacial por lo que se pierde cierta información, aún así resulta importante conocer los datos de esta manera para tomar decisiones geoestadísticas.

TABLA 5.3: **Variables ambientales con las más altas correlaciones.** Correlación entre variables determinada para evitar redundancia en el cálculo de la interpolación.

Var 1	Var 2	Corr <sup>a</sup>
Limo	%H	-0.833007
Limo	Arena	-0.892152
Co	CT	0.972836
Ci	CT	0.785613

<sup>a</sup>corr: de acuerdo a corr, una de las variables puede ser usada en los modelos en lugar de la otra para evitar redundancia.

### 5.1.2. Correlaciones entre características del suelo

El análisis de multicolinealidad (multicol) y de correlación (corSelect) entre características del suelo reveló variables que están muy correlacionadas con otras y podrían ser redundantes al momento de realizar la interpolación. Al ser combinación lineal de otras su comportamiento será similar o contrario, al momento de interpolar. Así, el limo muestra una correlación negativa y significativa con el contenido de agua y de arena y como era de esperarse el contenido de C total con el de C orgánico e inorgánico (Tabla 5.3).

El análisis de colinealidad muestra que las características del suelo que son independientes son el pH, la EC y el %H, mientras que el contenido de C puede ser explicado al 100 % por otras variables (Tabla 5.4).

TABLA 5.4: **Multicolinealidad y Factor de inflación de varianza (VIF).**  
Propiedades del suelo que son aproximadamente combinación lineal de otras.

Parámetro	coeficiente r	Tolerancia	VIF <sup>a</sup>	
<b>C total</b>	1	1.88E-11	5.33E+10	b
<b>Co<sup>d</sup></b>	1	3.01E-11	3.32E+10	b
<b>Ci<sup>e</sup></b>	1	2.15E-10	4.65E+09	b
<b>Arena</b>	0.999991	9.31E-06	1.07E+05	
<b>Limo</b>	0.999985	1.47E-05	6.80E+04	
<b>Arcilla</b>	0.999956	4.38E-05	2.28E+04	
<b>%H<sup>f</sup></b>	0.872065	1.28E-01	7.82E+00	c
<b>pH</b>	0.853892	1.46E-01	6.84E+00	c
<b>EC<sup>g</sup></b>	0.689002	3.11E-01	3.22E+00	c

<sup>a</sup>VIF :Factor de inflación de la varianza;

<sup>b</sup>puede ser explicada al 100 %

por otras variables; <sup>c</sup>variables independientes.

<sup>d</sup>Co: contenido de C orgánico; <sup>e</sup>Ci: contenido de C inorgánico;

<sup>f</sup>%H: contenido de agua; <sup>g</sup>EC: conductividad electrolítica

## 5.2. Análisis geoestadísticos de parámetros fisicoquímicos del suelo

### 5.2.1. Exploratorios geoestadísticos

Adicionalmente a la descripción estadística de las características del suelo a lo largo del transecto, se realizó la geoexploración (Ribeiro y Diggle 2016), la cual nos permite describir los datos espacialmente, observando particularmente la situación espacial tomando en cuenta la latitud y longitud de cada sitio de muestreo. Los gráficos geoexploratorios obtenidos se pueden observar en los Apéndices A.4 y A.5 El transecto de muestreo sigue una línea hacia el sureste con separaciones irregulares entre cada sitio muestreado. A lo largo del transecto se aprecian diferencias en los



valores observados, principalmente aquellos de EC y H, que demuestran la presencia de gran variabilidad independientemente de si la distancia entre dos puntos es pequeña, tomando en cuenta la varianza observada para cada variable a lo largo del transecto se puede observar cierto grado de heterogeneidad en la dirección de muestreo planteada para ciertos parámetros tal como los anteriores mencionados H y EC (valores de varianza 165.97 y 2736.5 respectivamente), mientras que otros demuestran un rango muy corto y un comportamiento casi homogéneo, tal como el pH y la clase textural (varianza de 0.009 y principalmente Franco arcillo arenosa).

## **5.2.2. Descripción de la estructura espacial de características del suelo**

### **5.2.2.1. Autocorrelación**

A través del análisis de autocorrelación, aplicando el índice de Moran se pudo identificar presencia de las distancias a las que cada parámetro presentó autocorrelación. Todas las características medidas a lo largo del transecto de muestreo presentaron autocorrelación positiva ( $p < 0,05$ ) a alguna distancia excepto el contenido de arena y sólo la EC presentó autocorrelación negativa a dos clases de distancia (Tabla 5.5).

La distribución de los valores en el dataset se encuentra espacialmente más agrupada cuando se presenta correlación positiva y más dispersa en el caso de correlación negativa (Tabla 5.5), lo que indica que los procesos espaciales presentes a lo largo del transecto no son aleatorios para la mayoría de los parámetros medidos. Según los autocorrelogramas, la estructura espacial presente es periódica o con presencia de parches, en donde cada parámetro llega a ser completamente distinto de como se presenta regularmente.

TABLA 5.5: **Autocorrelación.** Distancias a las cuales los parámetros físico-químicos medidos muestran autocorrelación significativa según el índice de Moran calculado usando 1000 permutaciones ( $p < 0.05$ )

Parámetro	Autocorrelación	Autocorrelación
	positiva	Negativa
pH	20 m	NS
EC <sup>a</sup>	30, 150 y 190 m	25 y 175 m
H <sup>b</sup>	75 m	NS
C orgánico <sup>c</sup>	211 m	NS
C inorgánico <sup>d</sup>	40 m	NS
Arcilla	90 m	NS
Limo	175 m	NS
Arena	NS <sup>e</sup>	NS

<sup>a</sup>EC: Conductividad electrolítica; <sup>b</sup>H: % Agua;

<sup>c</sup>Co: contenido de C orgánico; <sup>d</sup>Ci: contenido de C inorgánico;

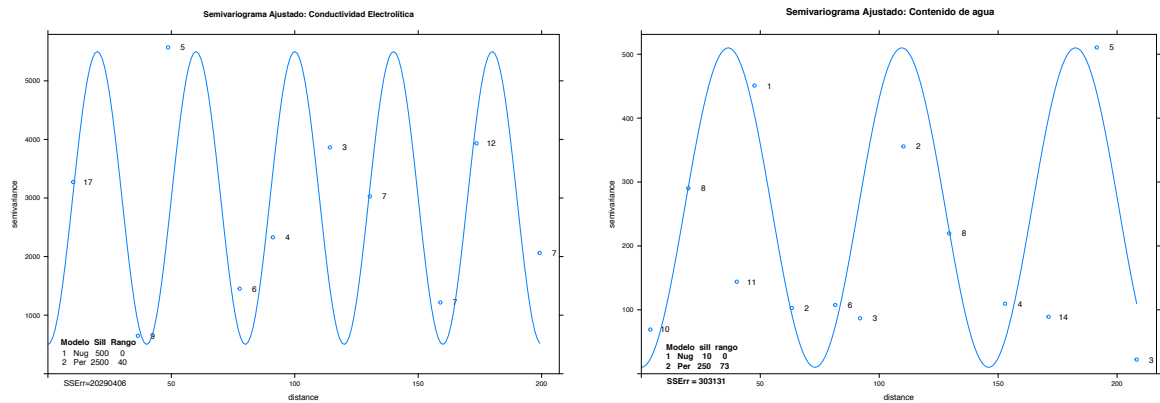
<sup>e</sup>NS: No hay autocorrelación significativa

### 5.2.2.2. Semivariogramas

Los semivariogramas omnidireccionales fueron calculados aplicando el estimador modulo (adaptando raíces cuartas del cuadrado de la diferencia, mediana y corrección del sesgo a fin de darle más robustez al análisis, según lo determinado por Cressie y Hawkins, 1993) y fueron obtenidos tanto el tipo nube, como el tipo agrupado, los semivariogramas para los parámetros EC y H se encuentran en la Figura 5.2a y 5.2b, los semivariogramas de los parámetros restantes pueden ser observados en el Apéndice B.3 y B.2.

Los semivariogramas de tipo nube proporcionan información acerca de la distribución de la semivarianza de los datos a cada par de distancias a lo largo de todo el transecto, permitiendo identificar valores atípicos dentro del conjunto de datos, mientras que el semivariograma agrupado, al ser ajustado proporciona los parámetros que servirán para definir la interpolación en el área de estudio.

El modelo que mejor ajustó a los datos de semivarianza en el presente estudio, fue periódico, al igual que en el análisis de correlación por autocorrelogramas, lo cual



(A) Semivariograma EC

(B) Semivariograma H

FIGURA 5.2: **Semivariogramas Ajustados** En la figura se observa el ajuste de los semivariogramas usando el modelo periódico para (A) EC conductividad Electrolítica ( $\text{dS m}^{-1}$ ) y (B) La cantidad de Agua (%).

TABLA 5.6: **Parámetros de Ajuste de Semivariograma.** La tabla muestra los parámetros característicos del modelo de ajuste del semivariograma, sill, nugget y rango. Así mismo una relación calculada nugget/sill que muestra la presencia o ausencia de autocorrelación.

Parámetro	Modelo	Sill	Rango (m)	Nugget	Relación Nugget/Sill
pH	Periódico	0.008	40	0.003	0.375
EC <sup>a</sup>	Periódico	2500	40	500	0.2
H <sup>b</sup>	Periódico	250	73	10	0.04
Ci <sup>c</sup>	Periódico	2.9	50	0.5	0.172
Co <sup>d</sup>	Periódico	28.6	55	0.8	0.028
Arcilla	Periódico	36	42.5	3	0.083
Arena	Periódico	75	42.5	20	0.267
Limo	Periódico	70	43	5	0.071

<sup>a</sup>EC: Conductividad electrolítica; <sup>b</sup>H: contenido de Agua (%);

<sup>c</sup>Ci: Contenido de C inorgánico; <sup>d</sup>Co: Contenido de C orgánico

indica de igual forma la presencia de una distribución parchada. De forma similar a los demás modelos, se obtienen los datos de ajuste, el sill, rango y nugget, para cada parámetro (los cuales pueden observarse en la Tabla 5.6) que servirán de modelo para realizar la interpolación correspondiente.

El contenido de  $C_o$  y  $C_i$ , el contenido de arcilla, limo, la EC y H mostraron por medio del semivariograma una alta autocorrelación espacial (Relación nugget/sill  $<0.25$ ), mientras que el pH y la arena mostraron una dependencia espacial media (nugget/sill entre 0.25 y 0.75) según lo descrito por Cambardella et al. 1994.

El pH en el suelo drenado del transecto muestreado fue extremadamente alcalino con poca variación espacial, mientras que el contenido de sal fue extremadamente variable. Tal como se menciona en la descripción del suelo, algunas partes han sido drenadas, la infiltración de una parte del agua en días lluviosos contribuye a dichas variaciones (Dendooven et al. 2010). De la misma forma, H fue altamente variable. La distribución del tamaño de partícula fue más constante aunque en estudios anteriores se habían encontrado grandes variaciones en el contenido de arcilla (Dendooven et al. 2010). El contenido de C orgánico también se observó altamente variable, encontrándose desde un sitio con pobre contenido orgánico a uno bastante rico. La descomposición de *Distichlis spicata* L. se ve obstaculizada en aquellos sitios donde el contenido de sal es bastante alto, así, la combinación de la variabilidad en la deposición del material de la planta y la mineralización generan en cierta medida la gran variabilidad del contenido de materia orgánica en éste suelo ocasionando de igual forma que la estructura espacial presente sea de forma periódica, encontrándose los parches antes referidos.

## 5.3. Estructura de comunidades microbianas

### 5.3.1. Estructura de las comunidades de Arquea

A partir de la secuenciación de la biblioteca genómica, se obtuvo un total de 11419 secuencias crudas de la region V1-V3 para Arqueas, con una longitud promedio de 455 pb. Después del filtro de calidad y la detección de quimeras, fueron obtenidas 10533 OTU<sub>97</sub> (Tabla 5.7).

El rango de diversidad reflejado por el índice Shannon oscila entre 2.10 y 5.87 correspondiendo el menor al sitio Tx007 y el mayor al sitio Tx008, mientras que el más alto índice Chao1 se encuentra en el sitio Tx011 (300) y el menor para el sitio Tx012 (30) seguido de Tx008 (Tabla 5.7). En índice de diversidad filogenética PD oscila entre 1.58 en el sitio Tx012 y 3.82 en el sitio Tx007. El sitio Tx007 fue más uniforme en el número de individuos que pertenecen a cada especie a diferencia de los sitios Tx008 y Tx012 en los cuales sólo pocas especies dominan de acuerdo al índice Simpson E (0.07-0.48). En general, las secuencias pertenecen a dos phyla arqueales, cuatro clases, cinco órdenes, cinco familias y nueve géneros (Tabla 5.7). El número de phylas arqueales osciló de 1 a 2 en cada sitio, clases de 2 a 4, órdenes de 2 a 4, familias de 1 a 4 y géneros de 3 a 6.

Los filotipos pertenecientes a Euryarchaeota (principalmente Halobacteriales) fueron los más abundantes, excepto en dos sitios (Tx008 y Tx012) en los que la abundancia relativa de Thaumarchaeota (principalmente Nitrososphaerales) fue mayor (Fig.5.3). Los miembros de la familia Halobacteriaceae fueron observados en el suelo de cada sitio y generalmente fueron los más abundantes, con una abundancia relativa de 15.8 % en el sitio Tx012 a 98.7 % en el sitio Tx002. La familia Nitrososphaeraceae (principalmente *Candidatus Nitrososphaera*) dominaron en el sitio Tx008 con una abundancia relativa de 78.4 %. Los filotipos se asignaron a ocho géneros, siendo

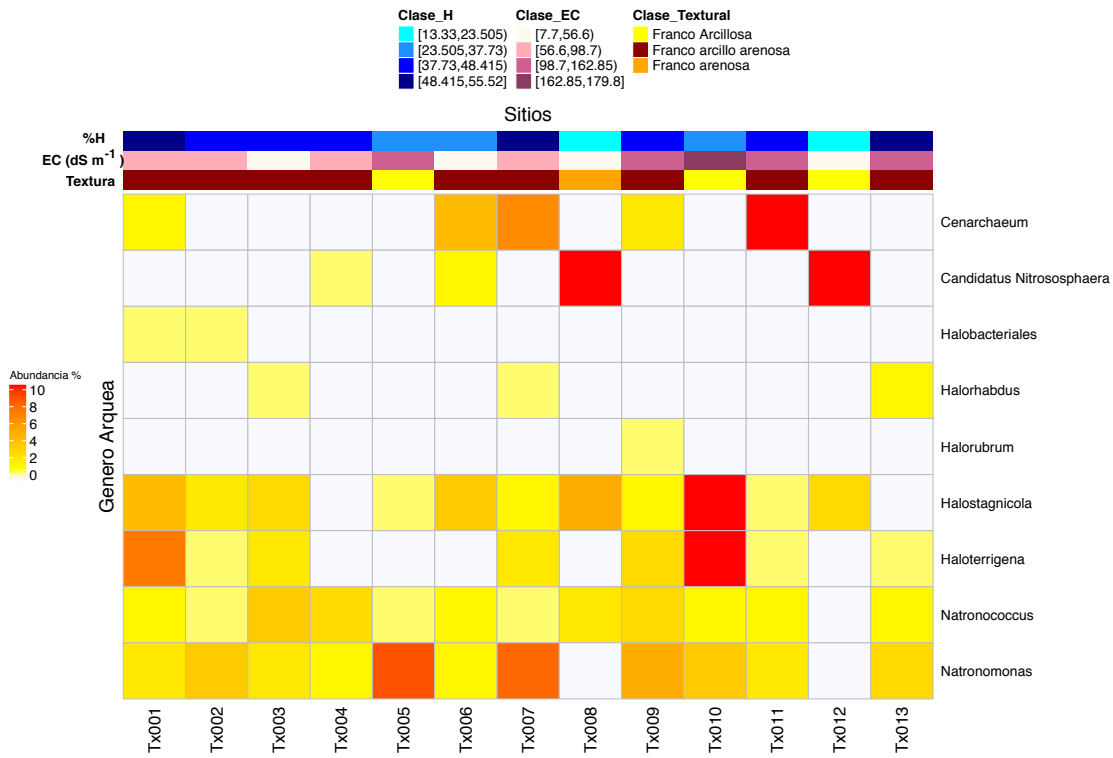


FIGURA 5.3: **Heatmap de abundancia Arqueas.** Heatmap correspondiente a la abundancia relativa porcentual de cada género de Arquea en cada sitio con anotaciones de valores de contenido de agua (%), conductividad electrolítica ( $\text{dS m}^{-1}$ ) y clasificación textural para cada sitio.

C. Nitrososphaera, *Crenarchaeum* y *Natronomonas* los más abundantes.

TABLA 5.7: **Diversidad de arqueas.** Índices de diversidad y profundidad de secuenciación de Arquea a cada sitio de muestreo (Tx001-Tx013) a lo largo de un transecto corriendo al sureste en suelo del exlago de Texcoco.

Sitio	Tx001	Tx002	Tx003	Tx004	Tx005	Tx006	Tx007	Tx008	Tx009	Tx010	Tx011	Tx012	Tx013	Total
<b>Secuencias<sup>a</sup></b>	802	1439	571	160	837	723	1365	1225	802	1112	826	630	927	11419
<b>OTU<sub>97</sub><sup>c</sup></b>	743	1292	518	152	775	685	1144	1187	750	1011	785	623	868	10533
<b>Phylum</b>	2	1	1	2	1	2	2	2	2	1	2	2	1	2 <sup>b</sup>
<b>Clase</b>	3	2	2	4	2	3	3	2	2	2	2	3	3	4
<b>Orden</b>	3	2	2	4	2	4	3	2	2	2	2	3	3	5
<b>Familia</b>	3	1	1	2	2	4	2	2	2	1	2	2	1	5
<b>Género</b>	6	5	5	3	3	5	6	3	6	4	5	2	4	9
<b>Parámetros de Diversidad</b>														
<b>Shannon</b>	5.24	5.15	5.01	5.61	5.33	4.79	5.87	2.1	5.66	4.72	5.31	2.87	5.6	
<b>Simpson</b>	0.95	0.94	0.94	0.97	0.94	0.91	0.97	0.49	0.97	0.91	0.94	0.7	0.96	
<b>Simpson E</b>	0.33	0.27	0.28	0.48	0.25	0.2	0.48	0.07	0.45	0.2	0.25	0.14	0.39	
<b>Simpson Reciprocal</b>	21.26	17.42	16.7	33.03	17.11	11.47	39.03	1.95	32.15	11.37	17.55	3.29	28.54	
<b>Chao1</b>	211.9	159.1	113.1	147.8	132.1	105.5	259.6	56.6	162.1	180	299.7	30	162.4	
<b>Diversidad Filogenética</b>	2.89	2.83	2.61	3.39	3.24	3.21	3.82	1.79	3.27	2.2	3.63	1.58	3.15	
<b>ACE<sup>d</sup></b>	239.55	218.11	165.55	175.18	164.38	145.56	355.46	86.2	182.63	216.98	257.6	31.56	228.47	

<sup>a</sup> Después de filtración de calidad; <sup>b</sup> asignado a ese nivel taxonómico; <sup>c</sup> número de OTUs encontrados después de ser normalizados a la muestra con menos secuencias; <sup>d</sup> ACE: Abundance-based Coverage Estimator

Euryarchaeota fue el phylum dominante a través del transecto de muestreo a excepción del suelo que presentó valores más bajos de EC (Tx008 y Tx012) donde el phylum Thaumarchaeota fue el que dominó. Estos resultados son consistentes con otros estudios de suelos y sedimentos alcalinos e hipersalinos, donde Euryarchaeota es el phylum dominante, siendo Halobacteriales el orden más abundante (Hollister et al. 2010; Ma y Gong 2013; Budakoglu et al. 2014; Vogt et al. 2017). Euryarchaeota es con frecuencia el phylum más abundante encontrado en el suelo del exlago de Texcoco (Valenzuela-Encinas et al. 2012; Navarro-Noya et al. 2015) aunque Thaumarchaeota se ha encontrado como el phylum predominante en estudios de suelos con menor contenido de EC, pH y arcilla (Wessén et al. 2011; Hatzenpichler 2012; Webster et al. 2015).

La diversidad y riqueza de Arquea en los sitios Tx008 y Tx012 fue la menor encontrada a lo largo del transecto, lo cual podría explicarse por la abundancia de Thaumarchaeota en esos sitios y que presentaron los menores valores de salinidad y quizá los más oligotróficos a lo largo del transecto. Resultados similares se obtuvieron en otros estudios en los que se encontró que las muestras dominadas por Thaumarchaeota mostraron una menor diversidad que en aquellas donde dominaban otros phyla de Arqueas (Auguet y Casamayor 2013; Navarro-Noya et al. 2015). Los sitios que presentaron mayor diversidad fueron Tx009 y Tx007, siendo en este último sitio donde se encontró mayor uniformidad de especies. Por otro lado, el sitio de mayor riqueza fue el Tx011. Aún hay muchas taxas no identificados en lugares con características extremas como los suelos y sedimentos hipersalinos.

### 5.3.2. Estructura de las comunidades de Bacteria

A partir de la secuenciación correspondiente a la librería de Bacteria, se obtuvo un total de 5786 secuencias crudas para la región V1-V6 amplificada, con una longitud



promedio de 930 pb. Después del filtro correspondiente y la detección quimérica, se obtuvieron 4999 OTU<sub>97</sub> (Tabla 5.8).

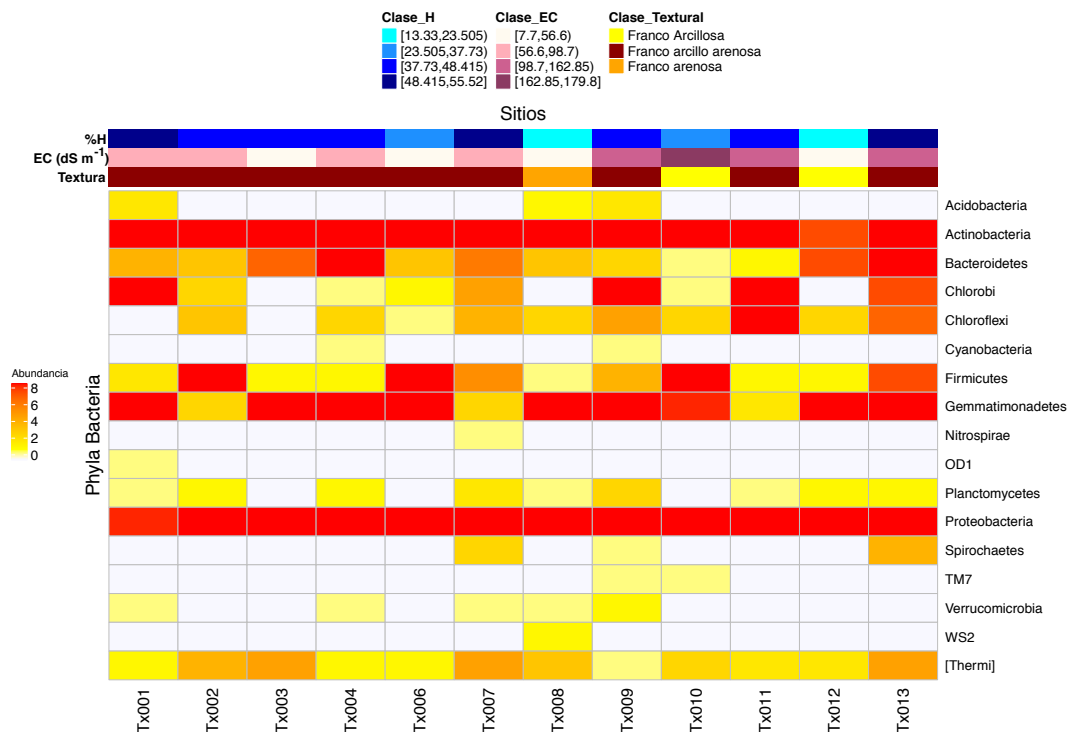


FIGURA 5.4: **Heatmap de abundancia bacteriana.** Heatmap correspondiente a la abundancia relativa porcentual de phyla bacterianos en cada sitio con anotaciones de valores de contenido de agua (%), conductividad electrolítica ( $\text{dS m}^{-1}$ ) y clasificación textural para cada sitio.

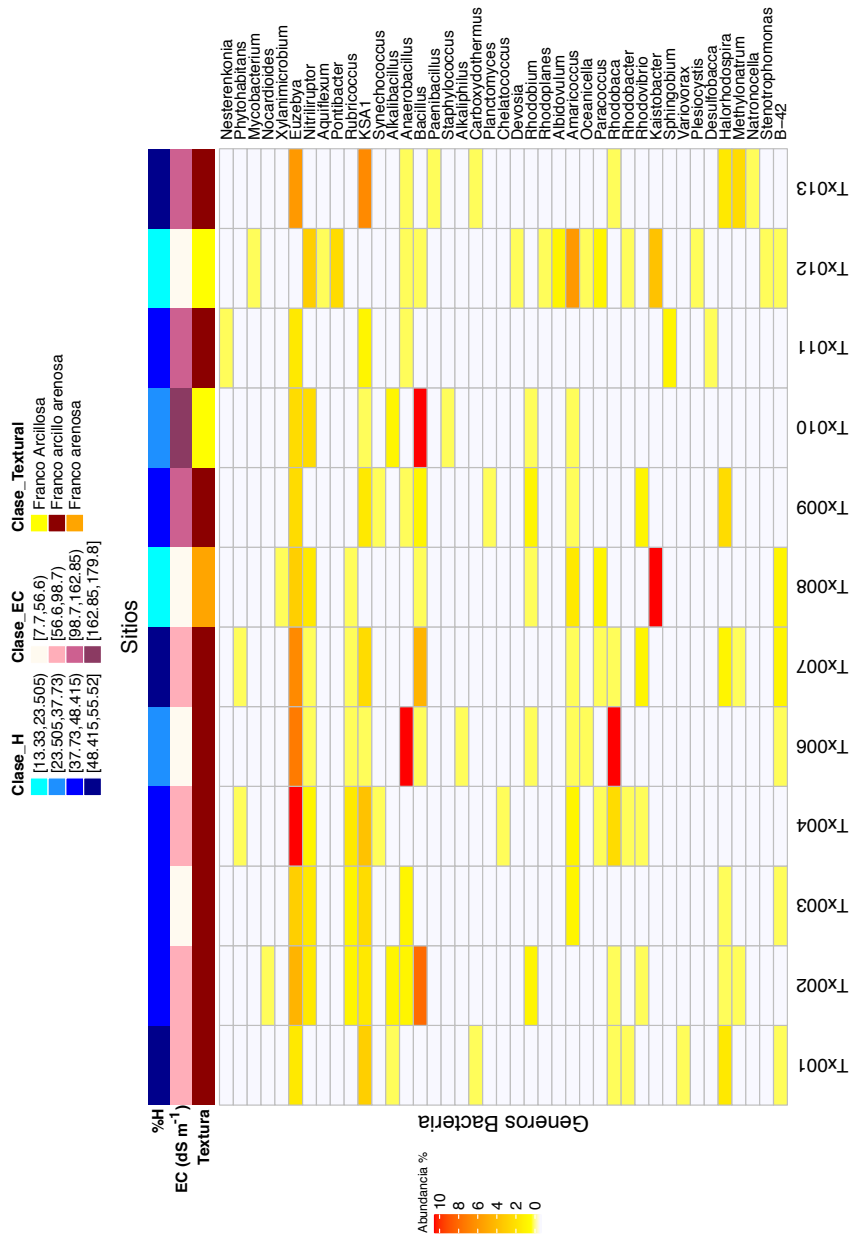


FIGURA 5.5: **Heatmap de abundancia bacteriana.** Heatmap correspondiente a la abundancia relativa porcentual de géneros bacterianos en cada sitio con anotaciones de valores de contenido de agua (%), Conductividad electrolítica (dS m<sup>-1</sup>) y Clasificación textural para cada sitio.

Los parámetros de  $\alpha$  diversidad indican que el sitio Tx007 presenta la mayor diversidad (mayor índice Shannon). El estimador de riqueza Chao1 osciló de 70 a 349.14, mientras que el índice PD se encuentra de 4.95 a 7.37 siendo el sitio Tx009 el que exhibió la mayor diversidad filogenética. En el caso de los sitios Tx001 y Tx010 sólo pocas especies bacterianas dominan mientras que el sitio Tx013 es el más uniforme en cuanto a número de especies, tal como se observa de los rangos del índice Simpson E (0.40-0.78)(Tabla 5.8). El número estimado de especies bacterianas fue mucho más alto en el sitio Tx002 que en el sitio Tx012, con base en los estimadores de riqueza Chao1 y ACE. Los índices Shannon y Simpson recíproco fueron más bajos en los sitios Tx001 y Tx010, el sitio Tx010 presenta el mayor valor de EC, por lo que puede ser observado un efecto sustancial de las condiciones Hipersalinas. En general, los filotipos pertenecen a 20 phylas bacterianos, 48 clases, 57 órdenes, 68 familias y 49 géneros. Los phyla más abundantes en las muestras fueron Actinobacteria seguido de Proteobacteria y Gemmatimonadetes (Fig. 5.4). Chlorobi y Firmicutes mostraron la mayor variabilidad en la abundancia relativa a lo largo del transecto. Miembros de *Euzzebya* (Actinobacteria) fueron los más abundantes, seguidos por KSA1 (Bacteroidetes) y *Bacillus* (Firmicutes)(Fig. 5.5)

TABLA 5.8: **Diversidad de Bacteria.** Índices de diversidad y profundidad de secuenciación de Bacteria para cada sitio de muestreo (Tx001-Tx013) a lo largo de un transecto corriendo al sureste en suelo del exlago de Texcoco.

Sitio	Tx001	Tx002	Tx003	Tx004	Tx006	Tx007	Tx008	Tx009	Tx010	Tx011	Tx012	Tx013	Total
<b>Secuencias<sup>a</sup></b>	668	676	436	323	231	794	304	700	297	245	280	167	5121
<b>OTU<sub>97</sub><sup>c</sup></b>	511	558	346	309	217	706	266	641	278	220	232	155	4439
<b>Phylum</b>	11	9	6	11	8	12	11	14	9	9	8	10	17 <sup>b</sup>
<b>Clase</b>	16	18	11	20	14	21	19	22	16	17	21	21	43
<b>Orden</b>	14	15	12	18	15	17	17	19	15	18	18	21	44
<b>Familia</b>	14	15	12	15	14	18	19	19	14	15	16	17	55
<b>Género</b>	9	13	8	12	12	13	10	10	8	6	17	9	41
<b>Índices de diversidad</b>													
<b>Shannon</b>	5.52	5.92	5.96	6.12	5.74	6.21	5.72	6.11	5.52	5.75	5.57	6.13	
<b>Simpson</b>	0.96	0.98	0.97	0.98	0.98	0.98	0.97	0.98	0.97	0.98	0.97	0.98	
<b>Simpson E</b>	0.4	0.54	0.48	0.66	0.64	0.71	0.58	0.71	0.47	0.61	0.6	0.78	
<b>Simpson Reciprocal</b>	27.08	42.34	38.77	54.82	42.81	61.27	39.16	57.23	29.99	41.67	35.11	61.27	
<b>Chao1</b>	323	349.14	207.43	178.06	130.08	197.24	93.25	168.35	132.33	129.5	70	126	
<b>Diversidad Filogenética</b>	6.36	6.15	5.88	6.24	4.97	7.11	6.1	7.37	5.16	4.95	5.4	6.86	
<b>ACE d</b>	222	337	223	227	133	224	128	182	138	144	74	131	

<sup>a</sup> Después de filtración de calidad; <sup>b</sup> Asignada a ese nivel taxonómico; <sup>c</sup> número de OTUs encontrados después de normalización a la muestra con el número más bajo de secuencias; <sup>d</sup> ACE: Abundance-based Coverage Estimator

Proteobacteria, Bacteroidetes y Actinobacteria son los taxones bacterianos dominantes en suelos salinos e hipersalinos (Ma y Gong 2013; Canfora et al. 2014; Liu et al. 2014; Santini, Warren y Kendr 2015; Kalwasinska et al. 2017) lo cual también se encuentra en este estudio. Chlorobi y Firmicutes que fueron identificados como phyla de abundancia menor en este estudio, con frecuencia ha sido encontrados como phyla menores en otros estudios de suelos salinos y alcalinos (Ma y Gong 2013), se ha relacionado frecuentemente con el contenido de C orgánico, sin embargo lo mismo se ha reportado en suelos arables (Constancias et al. 2015). Gemmatimonadetes, por su parte, ha sido reportado en otros sitios hipersalinos independientemente del grado de salinidad (Canfora et al. 2014).

Se observó una gran diferencia en el número de especies bacterianas entre las muestras Tx002 y la muestra Tx012 (Tabla 5.8). Un caso similar se observó en otro entorno hipersalino, una cal salina sodada en Polonia, donde se puede ver una gran variación en dos capas muestreadas (Kalwasinska et al. 2017).

### 5.3.3. Relación Arquea/Bacteria

La relación entre índices de  $\alpha$  diversidad de Arqueas/Bacterias fue calculada mostrando que para los sitios Tx008 (0.37 Shannon Arq/Bac) y Tx012 (0.43 Chao1 y ACE Arq/Bac) la diversidad y riqueza bacteriana, respectivamente, domina la diversidad y riqueza de arqueas, mientras que en el sitio Tx011, la riqueza de Arquea fue mayor que la riqueza de Bacteria (2.31 Chao1 y 1.70 ACE relación Arq/Bac). Los sitios Tx001, Tx004, Tx007 y Tx009 tienen una diversidad proporcional de arqueas y bacterias (Relación Arq/Bac Shannon 1) tal como se puede observar en la tabla 5.9.

TABLA 5.9: **Relación Arquea/Bacteria.** Tabla de relaciones Arquea/Bacteria para los índices de diversidad y niveles taxonómicos en cada sitio a lo largo del transecto de muestreo

Sitio	Tx001	Tx002	Tx003	Tx004	Tx006	Tx007	Tx008	Tx009	Tx010	Tx011	Tx012	Tx013
<b>Secuencias</b>	1.20	2.13	1.31	0.50	3.13	1.72	4.03	1.15	3.74	3.37	2.25	5.55
<b>OTU<sub>97</sub></b>	1.45	2.32	1.50	0.49	3.16	1.62	4.46	1.17	3.64	3.57	2.69	5.60
<b>Phylum</b>	0.18	0.11	0.11	0.29	0.29	0.20	0.25	0.18	0.14	0.25	0.33	0.10
<b>Clase</b>	0.19	0.13	0.14	0.31	0.25	0.18	0.13	0.12	0.15	0.15	0.20	0.18
<b>Orden</b>	0.19	0.15	0.15	0.33	0.33	0.19	0.14	0.13	0.15	0.17	0.25	0.18
<b>Familia</b>	0.21	0.08	0.07	0.20	0.36	0.11	0.15	0.15	0.09	0.22	0.20	0.07
<b>Género</b>	1.00	0.56	0.83	0.33	0.71	0.67	0.33	0.67	0.57	2.50	0.33	0.50
<b>Relación Arq/Bac de índices de Diversidad<sup>a</sup></b>												
<b>Shannon</b>	0.95	0.87	0.84	0.92	0.83	0.95	0.37	0.93	0.86	0.92	0.51	0.91
<b>Simpson</b>	0.99	0.97	0.97	0.99	0.93	0.99	0.50	0.99	0.94	0.97	0.72	0.98
<b>Simpson E</b>	0.83	0.50	0.59	0.72	0.31	0.67	0.13	0.63	0.43	0.41	0.23	0.50
<b>Simpson Reciproco</b>	0.78	0.41	0.43	0.60	0.27	0.64	0.05	0.56	0.38	0.42	0.09	0.47
<b>Chao1</b>	0.66	0.46	0.55	0.83	0.81	1.32	0.61	0.96	1.36	2.31	0.43	1.29
<b>Diversidad Filogenética</b>	0.45	0.46	0.44	0.54	0.65	0.54	0.29	0.44	0.43	0.73	0.29	0.46
<b>ACE<sup>b</sup></b>	1.08	0.65	0.74	0.77	1.10	1.59	0.68	1.00	1.57	1.79	0.43	1.74

<sup>a</sup>La relación Arquea/Bacteria fue calculada de las tablas de diversidad a partir de datos filtrados y normalizados al mínimo número de secuencias encontradas. <sup>b</sup>ACE:

En algunos estudios en ambientes salinos e hipersalinos la diversidad de Archaea es mayor que la de Bacteria (Canfora et al. 2015) y se acentúa para las Arquea oxidantes de amonio (Wessén et al. 2011; Marusenko et al. 2013; Bollmann, Bullerjahn y McKay 2014). El caso del presente estudio es contrario en los sitios con menores valores de EC, mientras que en los demás la diversidad resultó proporcional. La riqueza de Arquea fue mayor en algunos sitios (Tx-010, Tx-011 y Tx-013) con características fisicoquímicas extremas del suelo (es decir, altos valores de EC y pH) ya que las Archaea generalmente están mejor adaptadas a ambientes halófilos extremos (Oren 2014) sin embargo aún se desconoce el potencial de Taxas desconocidos.

El número de secuencias encontrado en el presente estudio para bacterias es muy bajo en algunos sitios, lo mismo sucede para el caso de Arquea. En algunos estudios como el de Vogt et al. en 2017, se puede observar que a pesar de la obtención de un gran número de secuencias encontradas por los esfuerzos recientes de secuenciación (Illumina), la riqueza se reduce significativamente e incluso es más baja que en otros estudios donde se utilizan herramientas de secuenciación de menor sensibilidad (454). La diversidad encontrada de bacterias y arqueas en estudios de sitios hipersalinos es muy diferente a la de aquellos suelos que no son salinos y aún no hay consenso en cuanto al tamaño o naturaleza de la diversidad generalmente encontrada (Ma y Gong 2013).

Asimismo, las curvas de rarefacción tanto de Bacteria como de Arquea (a excepción de Tx008 y Tx012 para Arquea que parece que lo muestran), no alcanzan el *plateau* y la comparación con Chao1 revela que hay una cantidad de especies propias de cada muestra, resultados similares a lo encontrado por Canfora et al. en 2014.

La alta proporción de OTU/secuencias observada tanto en Arquea como en Bacteria puede deberse a una combinación de factores. Primero, la similitud entre dos secuencias se calcula como el porcentaje que coincide en una alineación de pares. Un umbral de similitud de uso común es del 97%, lo cual se determinó de un estudio

empírico que muestra que un gran número de cepas tienen un 97 % de similitud en la secuencia de 16S rRNA (Konstantinidis y Tiedje 2005). Adicionalmente, se puede encontrar ambigüedad incluso considerando las regiones hipervariables del ARNr 16S, por lo que un organismo puede tener múltiples copias del gen que son diferentes en algunos sitios en un alto porcentaje, o dos organismos pueden tener hasta un 99 % de similitud de secuencias. Los métodos que utilizan la similitud de secuencias no consideran las distancias evolutivas, la cuál es una de las principales causas de ambigüedad. El uso de ciertos métodos, por ejemplo Alineación múltiple, sobre otros, como alineación pareada, puede disminuir la sobreestimación, incluso cuando se considera la similitud y no las distancias evolutivas (Nguyen et al. 2016). El uso de una agrupación de referencia abierta submuestreada mejora la estabilidad de las OTUs (Rideout et al. 2014; He et al. 2015).

En un entorno con condiciones particulares (alta alcalinidad y salinidad), la posibilidad de tener organismos ya informados en la base de datos de referencia (Greengenes v13.5 (<http://greengenes.lbl.gov>)) y una secuenciación adecuada disminuye. El valor correcto de la disimilitud para definir OTU no solo depende del método elegido sino también de la complejidad de la muestra (Chen et al. 2013). Li et al., en 2014, observaron que los microorganismos en ambientes extremos tienen tasas evolutivas más altas que aquellos en ambientes con características más benéficas, es decir, se presenta una tasa mayor de mutación impuesta por condiciones de estrés, proponiéndose este, como el mecanismo principal del proceso evolutivo, que podría influenciar el porcentaje de similitud y la ambigüedad de las secuencias.

Cabe señalar, de forma adicional, que el número total de lecturas obtenidas no es un reflejo del número absoluto de microorganismos presentes, ya que las muestras son solo una fracción de los microorganismos que se encuentran en el medio ambiente (Weiss et al. 2017). Según otros estudios, por ej. Callahan, Mcmurdie y Holmes 2017, será necesario utilizar otras formas distintas de las OTU para la asignación de



taxonomía.

#### 5.3.4. Asignación manual de secuencias

Con la asignación manual de secuencias representativas fue posible detectar un Phylum adicional de Arquea presente en los sitios de muestreo (Woesearchaeota), además de lograr la asignación a nivel de especie, en total 17.4 % (1394 secuencias) del total de secuencias fue asignada a siete especies (>97 % de identidad) dentro de los phyla Euryarchaeota y Thaumarchaeota (Fig.5.6). Adicionalmente, la asignación manual detectó cinco clases, ocho órdenes, 11 familias y 14 géneros (Fig. 5.7). Los órdenes más abundantes fueron Halobacteriales (Halobacteria) seguido de Natrialbales (Halobacteria) y Nitrososphaerales (Nitrososphaeria). De los géneros asignados cuatro fueron identificados igualmente en la asignación por qiime y la asignación manual, entre los adicionales destaca *Halobiforma* (Oren, Arahal y Ventosa 2009), como el segundo más abundante. Detalles de las asignaciones pueden encontrarse en el Apéndice C.

Para el caso de Bacteria, el total de phyla con secuencias asignadas se redujo de 17 a 13, un 47 % pertenecientes al Phylum Actinobacteria, seguido de Proteobacteria (15 %), Chlorobi (12 %) y Gemmatimonadetes (11 %), por otro lado sólo se logró la agrupación del 2 % (44 secuencias) de las secuencias representativas en cinco especies dentro de los phyla Proteobacteria y Firmicutes (Fig. 5.8). En general se asignaron 20 clases, 24 órdenes 32 familias y 21 géneros. Los órdenes más abundantes corresponden a los Nitriliruptorales (Actinobacteria), seguido de Ignavibacteriales (Chlorobi) y Egibacterales (Actinobacteria). Al realizar la asignación manual se descarta la presencia de los phyla: Cyanobacteria, Nitrospirae, OD1, TM7, Verrucomicrobia y WS2, mientras aparecen los phyla Balneolaeota (Hahnke et al. 2016), Rhodothermaeota (Muñoz, Rosselló-Móra y Amann 2016). Deinococcus-Thermus es

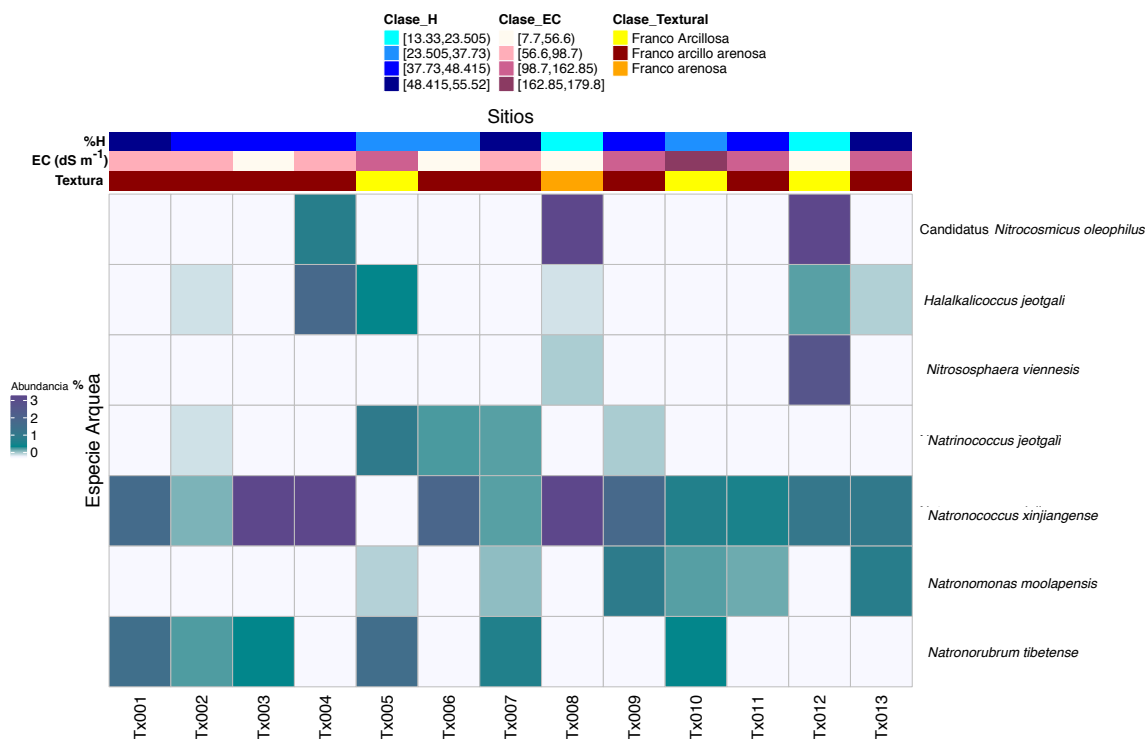


FIGURA 5.6: **Heatmap de abundancia asignada manualmente a Arquea.** Heatmap correspondiente a la abundancia relativa porcentual de especies de Arquea en cada sitio con anotaciones de valores de contenido de agua (%), Conductividad electrolítica (dS m<sup>-1</sup>) y Clasificación textural para cada sitio.

detectada en tres bases de datos mientras que para Greengenes el phylum asignado es [Thermi] (Youssef et al. 2015). De los géneros asignados se comparten siete entre ambas asignaciones, destacando en la asignación manual el género *Egibacter* un género halofílico obligado y alcalifílico facultativo descrito en 2016 (Zhang et al. 2016). Detalles de la asignación se pueden observar en C.

Es importante considerar que si bien la asignación manual implica mucho trabajo, que puede ser absorbido por el uso de herramientas como qiime (Caporaso et al.

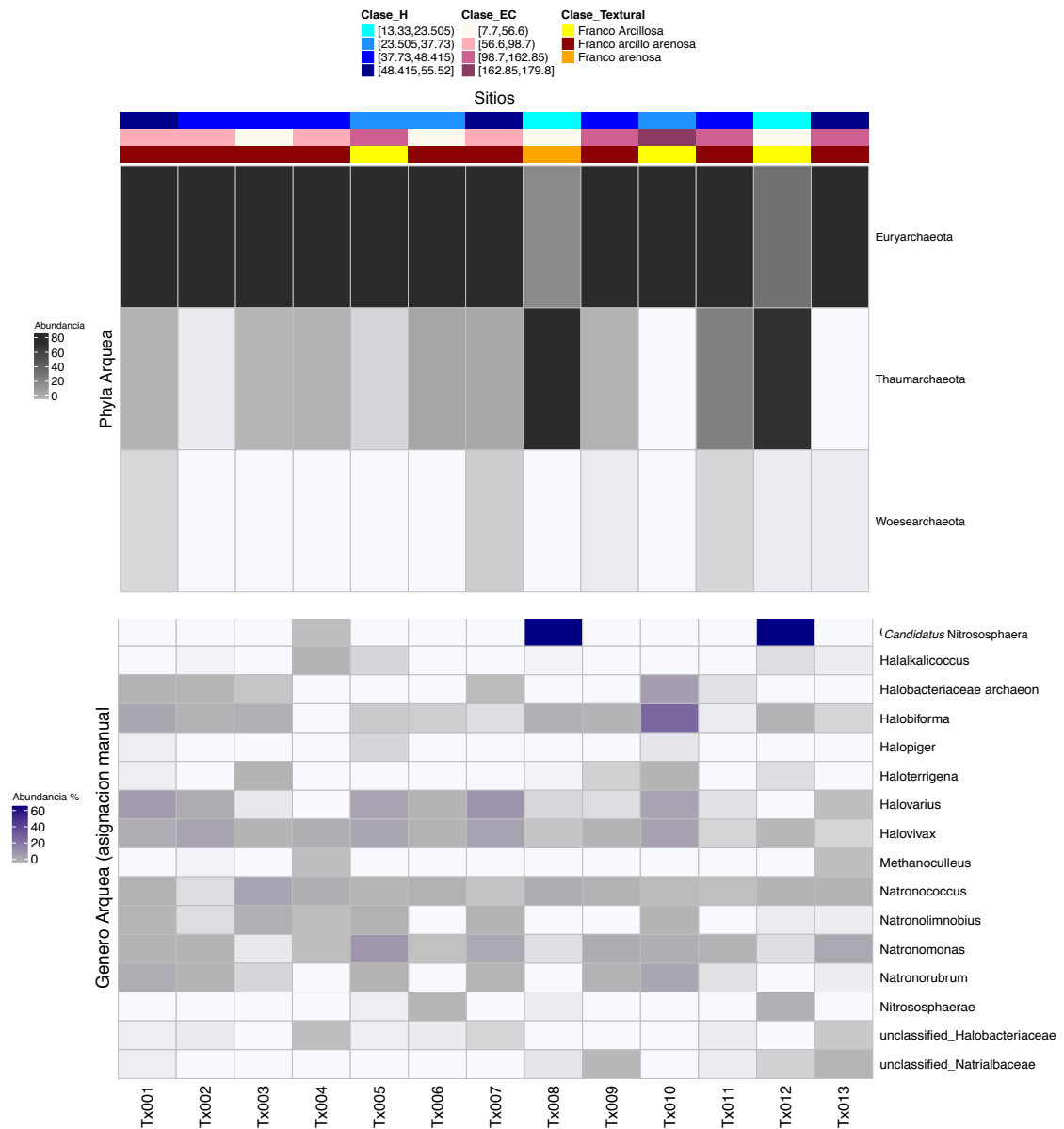


FIGURA 5.7: **Heatmap de abundancia asignada manualmente a Arquea.** Heatmap correspondiente a la abundancia relativa porcentual de Phyla y géneros de Arquea en cada sitio con anotaciones de valores de contenido de agua (%), Conductividad electrolítica ( $\text{dS m}^{-1}$ ) y Clasificación textural para cada sitio.

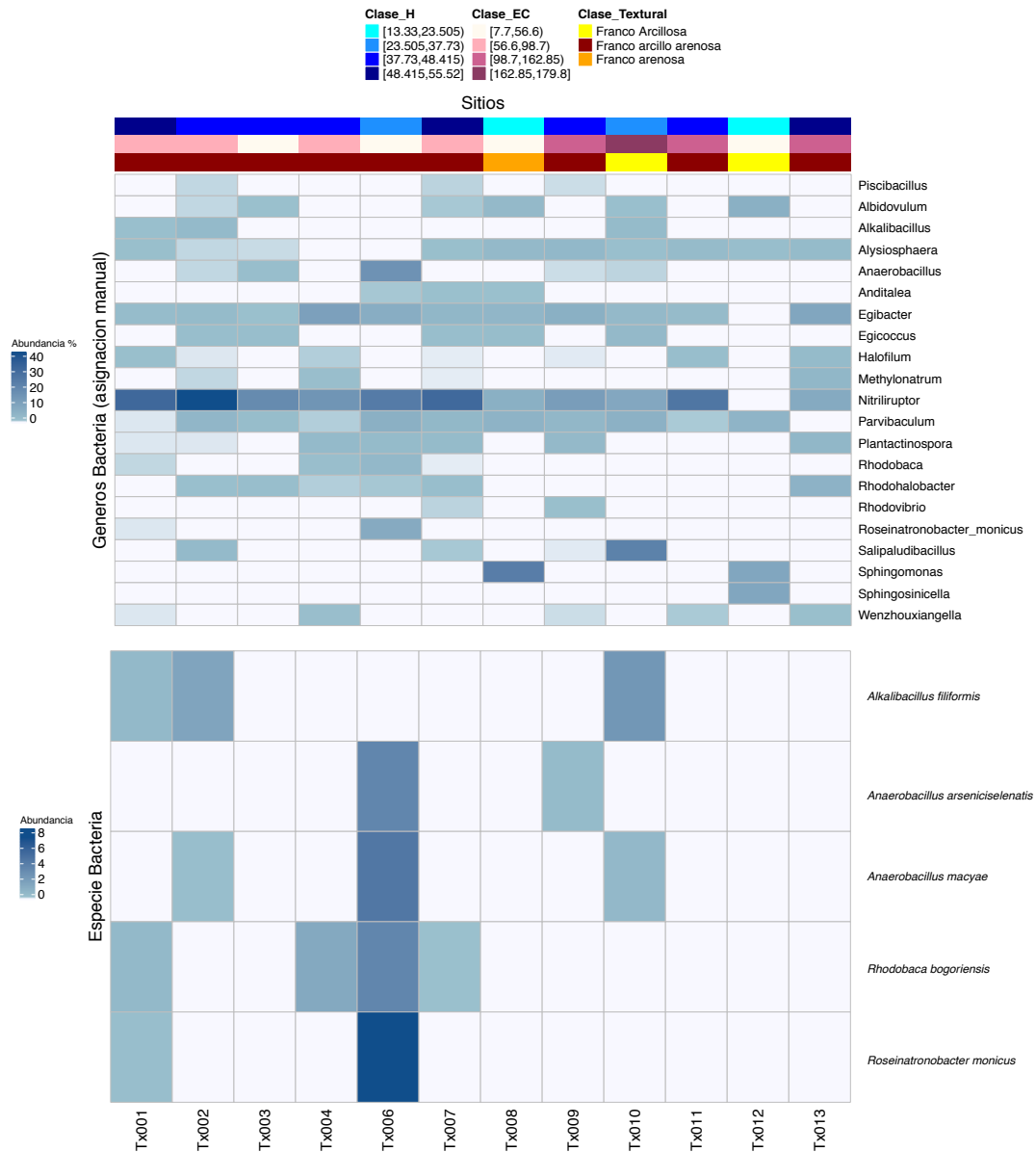


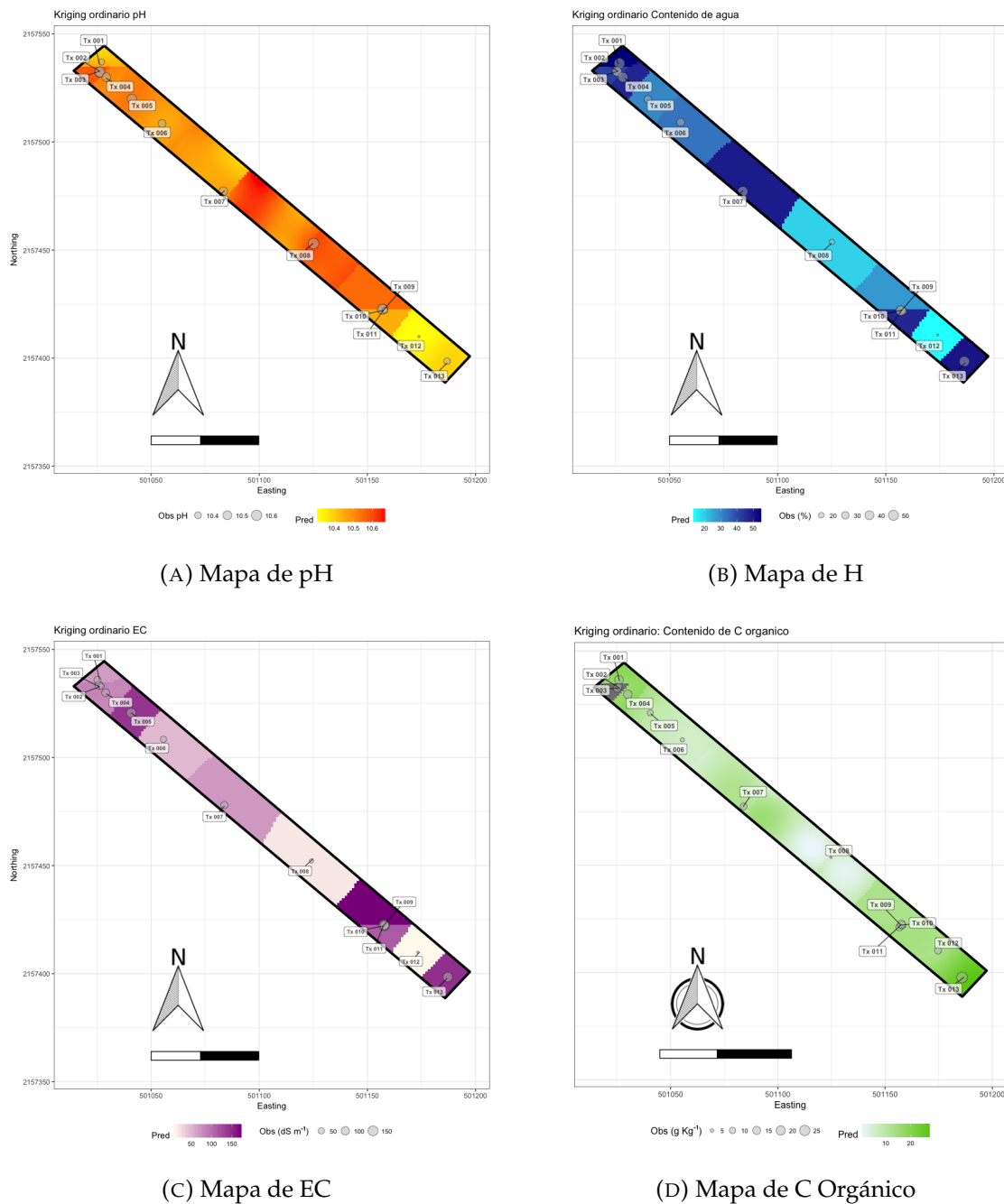
FIGURA 5.8: **Heatmap de abundancia asignada manualmente a Bacteria.** Heatmap correspondiente a la abundancia relativa porcentual de géneros y especies de Bacteria en cada sitio con anotaciones de valores de contenido de agua (%), Conductividad electrolítica ( $\text{dS m}^{-1}$ ) y Clasificación textural para cada sitio.

2010b), se puede usar de forma complementaria a éstas. El creciente uso de herramientas de secuenciación masiva y la mejora de dichas herramientas provocan que las bases de datos no siempre se encuentren al día por lo que una revisión manual de las secuencias encontradas, sobre todo aquellas de poca abundancia o asignadas a un nivel taxonómico bajo puede mejorar considerablemente la asignación taxonómica.

## 5.4. Mapas de Distribución

### 5.4.1. Modelos predictivos de características del suelo

Los suelos estudiados se caracterizaron por valores de pH alcalinos y clasificados texturalmente como Franco arcillo arenosos. Los menores valores de pH ( $<10.2$ ) fueron predichos en un área delimitada de aproximadamente  $180 \text{ m}^2$ , localizado hacia la parte sureste del polígono, sin embargo la variación de valores de pH fue baja como para definir presencia de gradiente. Los menores valores de EC ( $<7 \text{ dS m}^{-1}$ ) fueron predichos de la misma forma más al sureste del polígono y para un área de  $300 \text{ m}^2$  cercana al centro del transecto de muestreo (501110 – 501155 al Este y 2157425 – 2157475 al Norte) y fueron correlacionados visualmente con parches de vegetación. H estuvo espacialmente distribuido de forma más heterogénea con valores que oscilan de 10 a 70 %. Los menores valores predichos para este parámetro coincidieron con los valores más bajos de EC (alrededor del sitio Tx012). Los valores más altos (70 %) fueron predichos en las orillas más al norte y al sur del polígono. El contenido de  $C_i$  se observó distribuido más homogéneamente a través del área puesto que la desviación estándar del parámetro interpolado fue baja (1.95). La predicción kriging del contenido de  $C_o$ , por otro lado, mostró los valores más altos en la orilla sureste del polígono y los más bajos alrededor del sitio Tx008 con un rango de 5 a  $25 \text{ g Kg}^{-1}$ .



**FIGURA 5.9: Mapas de distribución de características del suelo.** Mapas de distribución de (A) valores de pH, (B) EC, (C) Contenido de agua y (D) Contenido de C orgánico. Obs: valores observados en cada sitio del transecto (Tx001-Tx013); Pred: Valores predichos para el parámetro fisicoquímico en un polígono B con 3000 pixeles de 1.29m<sup>2</sup> cada uno.

#### 5.4.1.1. Validación

La validación cruzada (CV-LOO) mostró que algunos sitios (8-16 %) estuvieron sub y sobre-estimados para casi todos los parámetros medidos tanto en el polígono de mayor área como el de menor área. El Error estándar promedio (ASE) y la raíz cuadrada del error estándar (RMSE) fueron similares entre si para la predicción del pH y H pero no para la EC y el Ci, arcilla y limo en el caso del polígono mayor, mientras que para el polígono menor los valores de los errores obtenidos fueron similares a los obtenidos para el polígono mayor a excepción de H, que presentó valores de ASE y RMSE no cercanos entre si y el contenido de Ci cuyo caso fue el contrario (Tabla 5.10)

Esto confirma una variabilidad adecuada en los valores predichos para la mayoría las características del suelo medidas. Algunas de las características fueron descritas más adecuadamente en el polígono mayor (como es el caso del contenido de agua) y otras fueron mejor representadas en el polígono menor (como el contenido de C inorgánico). El RMSEr para el contenido de agua fue de 1.108 y de 1.378 para el contenido de arcilla en el primer polígono considerado (mayor), indicando que el modelo reflejó menos del 50 % de la variabilidad en los puntos de validación pero fue adecuado en el polígono menor para el caso del contenido de agua (Tabla 5.10). Por lo tanto las predicciones para esas características en el polígono mayor no fueron completamente satisfactorias.

#### 5.4.2. Mapas de Distribución de comunidades Arquea

Para el polígono A, de mayor área de pixel, la diversidad Shannon de Archaea a través del área de predicción fue de 2 a 6, el mismo caso en el polígono B. Alrededor del sitio Tx007, cerca del centro del transecto se predice una diversidad mayor de Arquea (Fig.5.10a). Por otro lado, la riqueza osciló de 50 a 250 para ambos polígonos,

TABLA 5.10: **Validación de la interpolación.** Índices calculados para los mapas generados por kriging ordinario en dos polígonos de distintas áreas (3000 píxeles de A:2.82 m<sup>2</sup> y B:1.29 m<sup>2</sup> cada uno) para las características del suelo y ciertos phyla de Arquea y Bacteria.

Parametro	Polígono A			Polígono B		
	ASE <sup>a</sup>	RMSE <sup>b</sup>	RMSEr <sup>c</sup>	ASE	RMSE	RMSEr
pH	0.07	0.09	0.89	0.06	0.09	0.89
EC <sup>d</sup>	42.05	32.65	0.62	41.96	32.70	0.62
H <sup>e</sup>	15.67	14.27	1.11	7.46	4.41	0.34
Ci <sup>f</sup>	2.30	1.20	0.58	1.55	1.20	0.58
Silt	6.89	4.06	0.53	3.69	6.22	0.81
Clay	3.20	6.14	1.38	3.17	6.14	1.38
Arquea Shannon	0.71	0.60	0.53	0.38	0.60	0.53
Arquea Chao1	39.06	46.43	0.63	23.50	46.43	0.63
Bacteria Shannon	0.37	0.20	0.44	0.25	0.26	0.57
Bacteria Chao1	129.36	34.56	0.28	100.88	36.91	0.30
Arq/Bac <sup>g</sup> Shannon	0.11	0.10	0.53	0.11	0.10	0.53
Arq/Bac Chao1	0.35	0.48	0.89	0.56	0.50	0.94
Euryarchaeota	0.16	0.30	1.18	0.05	0.31	1.16
Thaumarchaeota	0.17	0.29	1.11	0.05	0.29	1.11
Proteobacteria	0.10	0.04	0.31	0.10	0.04	0.31
Actinobacteria	0.13	0.09	0.77	0.13	0.09	0.77
Bacteroidetes	0.03	0.01	0.45	0.03	0.01	0.45
Gemmatimonadetes	0.10	0.02	0.17	0.10	0.03	0.31
Chlorobi	0.17	0.03	0.35	0.05	0.03	0.35
[Thermi]	0.01	0.01	0.92	0.01	0.01	0.92
Firmicutes	0.06	0.10	1.62	0.04	0.01	0.14
Woesearchaeota <sup>1</sup>				0.0013	0.0025	1.08

<sup>a</sup>Error estándar promedio; <sup>b</sup>Raíz cuadrada del error; <sup>c</sup>Raíz cuadrada del error normalizado (relativo); <sup>d</sup>Conductividad electrolítica; <sup>e</sup>Contenido de agua (H); <sup>f</sup>Contenido de C inorgánico; <sup>g</sup>Relación Arquea/Bacteria; <sup>1</sup>Phyla encontrado en asignación manual de secuencias



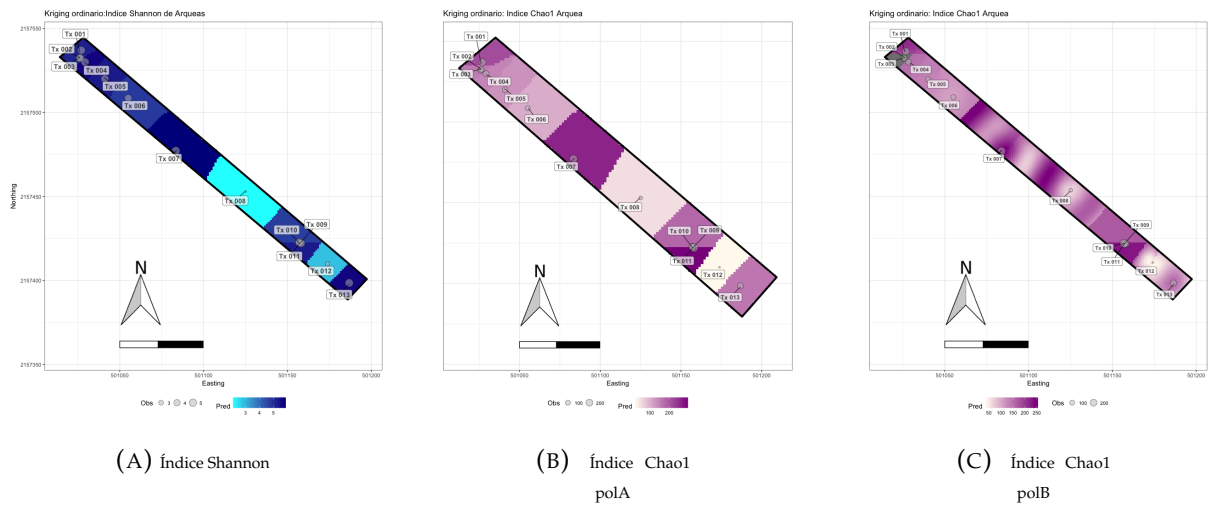
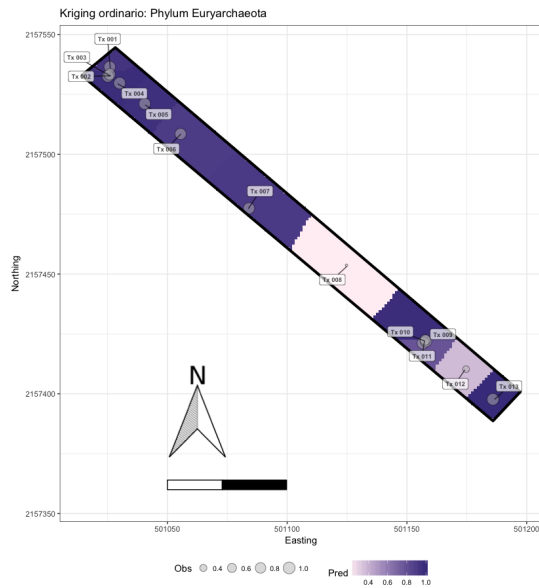


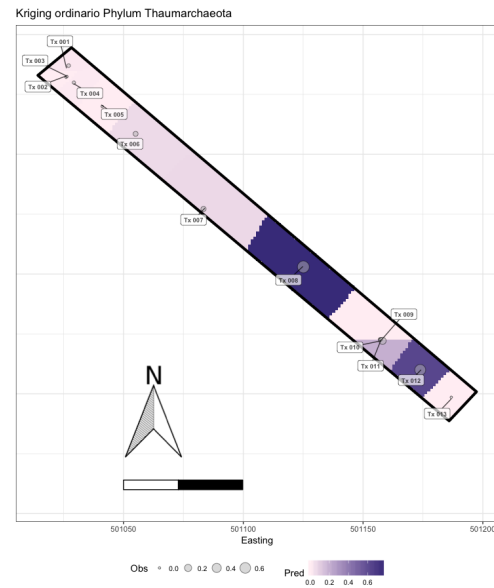
FIGURA 5.10: Mapas de distribución de Índices de  $\alpha$  Diversidad Arquea. Mapas de distribución de Arquea para (A) El índice Shannon, (B) el índice Chao1 en el polígono A y (C) el índice Chao1 en el polígono B. Obs: valores observados en cada sitio del transecto (Tx001-Tx013); Pred: Valores predichos para el parámetro fisicoquímico en polígonos con 3000 píxeles polA de 2.82 m<sup>2</sup> y polB de 1.29 m<sup>2</sup> cada uno.

sin embargo la distribución en el polígono de menor área mostró más claramente las zonas de posible transición y reveló la falta de intensidad de muestreo en puntos intermedios (Fig. 5.10b y 5.10c). La distribución de Thaumarchaeota mostró dos zonas (parches) de alta abundancia dentro del área de predicción que muestran correlación negativa con la predicción de la abundancia de Euryarchaeota y a su vez, se encuentra en zonas de bajos valores de EC (Figs. 5.11a, 5.11b y 5.9).

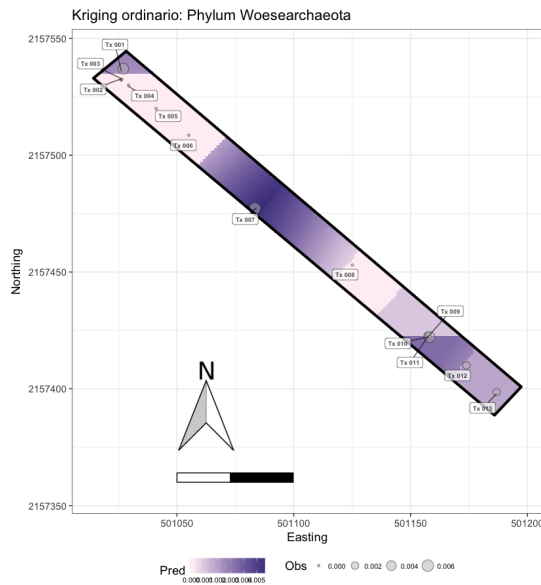
Euryarchaeota se presentó regularmente a lo largo del área de predicción con las más bajas abundancias, incluso nulas, en las zonas donde se presentó alta abundancia de Thaumarchaeota. Patrones similares se observan en el polígono de menor tamaño de píxel (1.29 m<sup>2</sup>). Se pueden observar cambios súbitos en la abundancia de phylas de Arquea, especialmente donde la conductividad electrolítica y el contenido de agua son menores. En el caso del Phylum Woesearchaeota, identificado gracias a la asignación manual, se presentó a lo largo del transecto con pequeñas abundancias,



(A) Euryarchaeota



(B) Thaumarchaeota



(C) Woesearchaeota

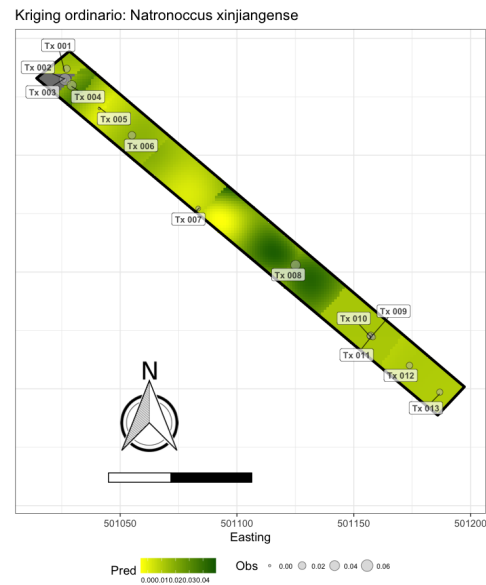
(D) *Natronococcus xinjiangense*

FIGURA 5.11: Mapas de distribución de Arquea. Mapas de distribución de las abundancias relativas de los phyla (A) Euryarchaeota, (B) Thaumarchaeota, (C) Woesearchaeota y la especie (D) *Natronococcus xinjiangense*. Obs: valores observados en cada sitio del transecto (Tx001-Tx013); Pred: Valores predichos para el parámetro fisicoquímico en un polígono B con 3000 pixeles de 1.29m<sup>2</sup> cada uno.

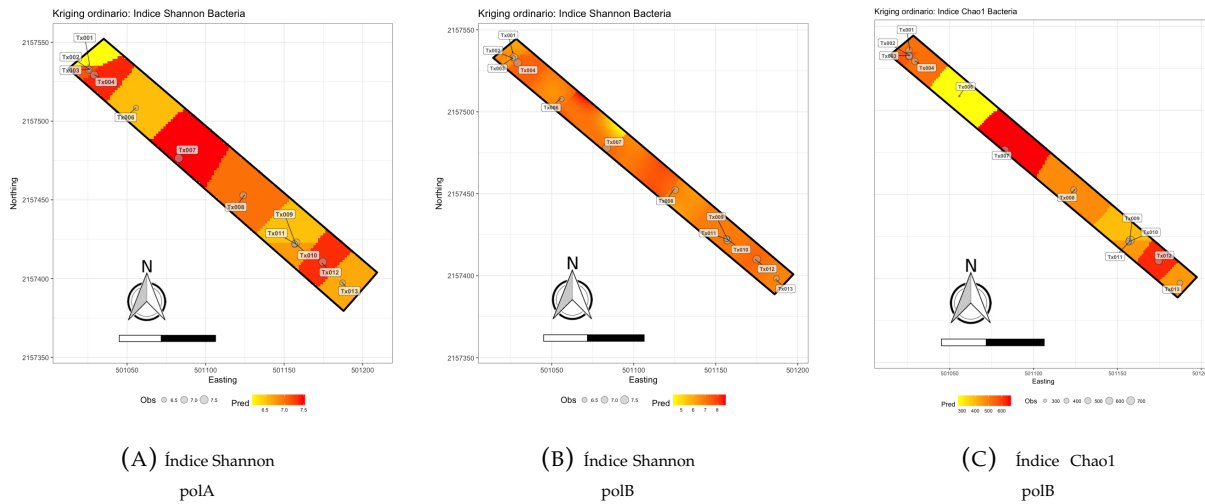
sin embargo se puede observar la mayor abundancia alrededor del sitio Tx007 disminuyendo su abundancia en aproximadamente 27 m hasta desaparecer en el sitio Tx008 y aumentando de nuevo gradualmente su abundancia hacia el sur-este donde muestra en los sitios Tx010 a Tx012 una abundancia un poco menor que en Tx008 (Fig. 5.11c). Dentro de las especies destacan *Natronomonas xinjiangense* que presentó la mayor variabilidad a lo largo del transecto con las mayores abundancias alrededor de los sitios Tx008 y Tx004 (Fig. 5.11d) y *Nitrososphaera viennensis* que se encontró alrededor del sitio Tx012 con la más alta abundancia, mientras que fue un poco menos alrededor de Tx008, siguiendo el mapa de distribución de Thaumarchaeota y sin presencia en los demás sitios del transecto.

Los índices ASE y RMSE (Tabla 5.10) fueron similares entre si para las comunidades de Arquea indicando buena variabilidad de la predicción, sin embargo el RMSE fue mayor a 70 % por lo que las predicciones no fueron completamente satisfactorias, sugiriendo un mayor esfuerzo de muestreo.

Es posible que zonas de transición de la distribución microbiana sean más fácilmente observadas a distancias menores de las consideradas en el presente estudio.

### 5.4.3. Mapas de Distribución de comunidades bacterianas

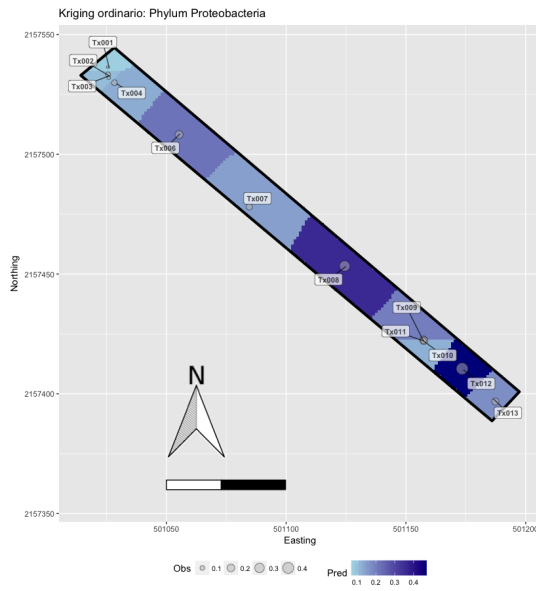
La predicción de la riqueza y diversidad para bacteria (observadas mediante el estimador Chao1 y el índice Shannon) mostró los valores mayores en parches alrededor de los sitios Tx012 y Tx007 (Fig. 5.12). La abundancia relativa de los grupos bacterianos de mayor abundancia, por ejemplo Actinobacteria y Proteobacteria, fue menos variable que la de la mayoría de los grupos de arqueas, pero la abundancia relativa de otros phyla, por ej. Chlorobi y Firmicutes, mostraron ser más heterogéneos. Por ejemplo, la abundancia relativa de Firmicutes varió entre <0.1 % en Tx008 a 25.0 % en Tx010 a pesar de la poca distancia entre ambos sitios.



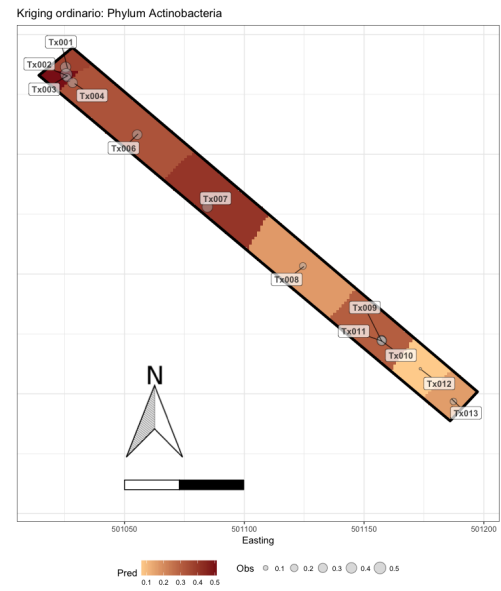
**FIGURA 5.12: Mapas de distribución de índices de  $\alpha$  diversidad Bacteria.** Mapas de distribución de Bacteria para (A) El índice Shannon en el polígono A, (B) el índice Shannon en el polígono B y (C) el índice Chao1. Obs: valores observados en cada sitio del transecto (Tx001-Tx013); Pred: Valores predichos para la distribución en polígonos con 3000 pixeles polA de 2.82 m<sup>2</sup> y polB de 1.29 m<sup>2</sup> cada uno.

Las predicciones para la abundancia de Proteobacteria muestran los mayores valores alrededor de Tx008 y Tx012 aunque la variabilidad a lo largo del transecto no fue grande, es decir Proteobacteria se presentó en todos los sitios con una abundancia mayor al 10 %, al igual que Actinobacteria que presentó los menores valores de abundancia en los sitios donde Proteobacteria fue mayor (Fig.5.13a 5.13b). En el caso de Bacteroidetes las mayores abundancias fueron observadas en el sur-este del transecto y alrededor de Tx004, por otro lado Gemmatimonadetes fue más homogéneo a través de todo el transecto (Fig. 5.13c 5.13d). Gemmatimonadetes fue observado anteriormente en suelos hipersalinos (Canfora et al. 2014) y parece no verse afectado por la EC. Esto podría deberse a que posee géneros y especies que pueden adaptarse a una amplia gama de condiciones diferentes e incluye especies que no se ven afectadas por la salinidad (Ma y Gong 2013; Canfora et al. 2014).

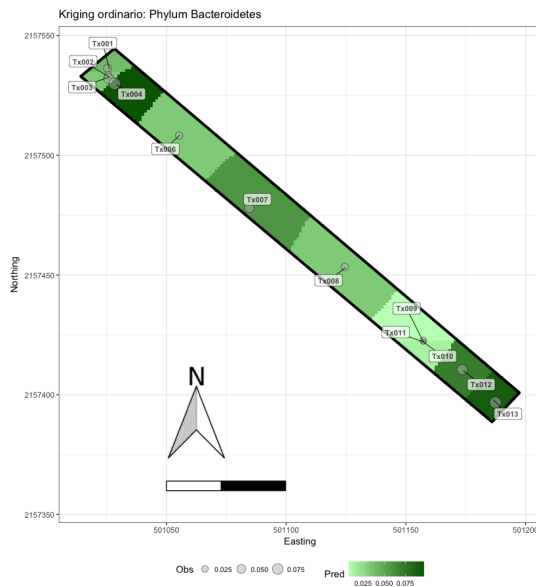
Firmicutes y Chlorobi, como ya se mencionó, mostraron una distribución más heterogénea a pesar de ser phylas menores (Fig. 5.14a y 5.14c). La transición de la



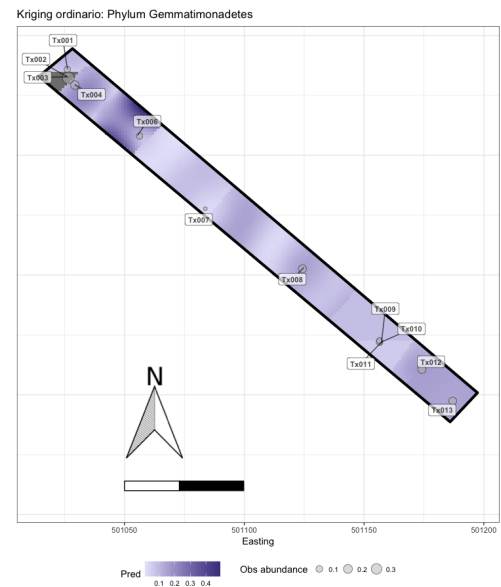
(A) Proteobacteria



(B) Actinobacteria



(C) Bacteroidetes



(D) Gemmatimonadetes

FIGURA 5.13: **Mapas de distribución de Bacteria.** Distribución de las abundancias relativas de los phyla (A) Proteobacteria, (B) Actinobacteria, (C) Bacteroidetes y (D) Gemmatimonadetes. Obs: valores observados en cada sitio del transecto (Tx001-Tx013); Pred: Valores predichos para la distribución de abundancia relativa en un polígono B con 3000 píxeles de 1.29m<sup>2</sup> cada uno.

distribución de Thermi puede ser observada más claramente de los sitios Tx006 a Tx008 y de Tx009 a Tx013 en dirección sur-este o viceversa (5.14b). De las pocas especies detectadas algunas como *Rhodobaca bogoriensis* (Fig. 5.14d) y *Reseinatronobacter monicus* muestran presencia en pocos sitios (Tx004-Tx006).

El ASE y RMSE fue similar entre si para la mayoría de las comunidades excepto para la riqueza bacteriana en ambos polígonos, lo que descarta que la predicción haya dado una buena variabilidad para éste parámetro. Por otro lado el RMSEr mostró que las predicciones de Firmicutes en el polígono mayor, Thermi en ambos polígonos y Actinobacteria (77 %) no fueron completamente satisfactorias (Tabla 5.10), posiblemente debido a la falta de muestras en sitios a menores o más regulares distancias. Las predicciones para Proteobacteria, Bacteroidetes, Gemmatimonadetes, Chlorobi (en ambos polígonos de predicción) y Firmicutes en el polígono menor reflejaron la distribución de dichos phyla de forma adecuada. Al igual que para Arqueas sería posible observar patrones de transición de algunas comunidades a distancias menores a las consideradas.

## 5.5. Factores que conducen la distribución de comunidades

A partir de los diferentes análisis multivariados se determinó qué características son las que conducen la variabilidad de las comunidades. Un conjunto de análisis tanto de correlaciones como de ordenación nos pueden dar una pista de qué características propician la presencia o ausencia de comunidades, y su análisis en conjunto con los mapas de distribución espacial nos llevan a definir patrones presentes en dichas distribuciones.

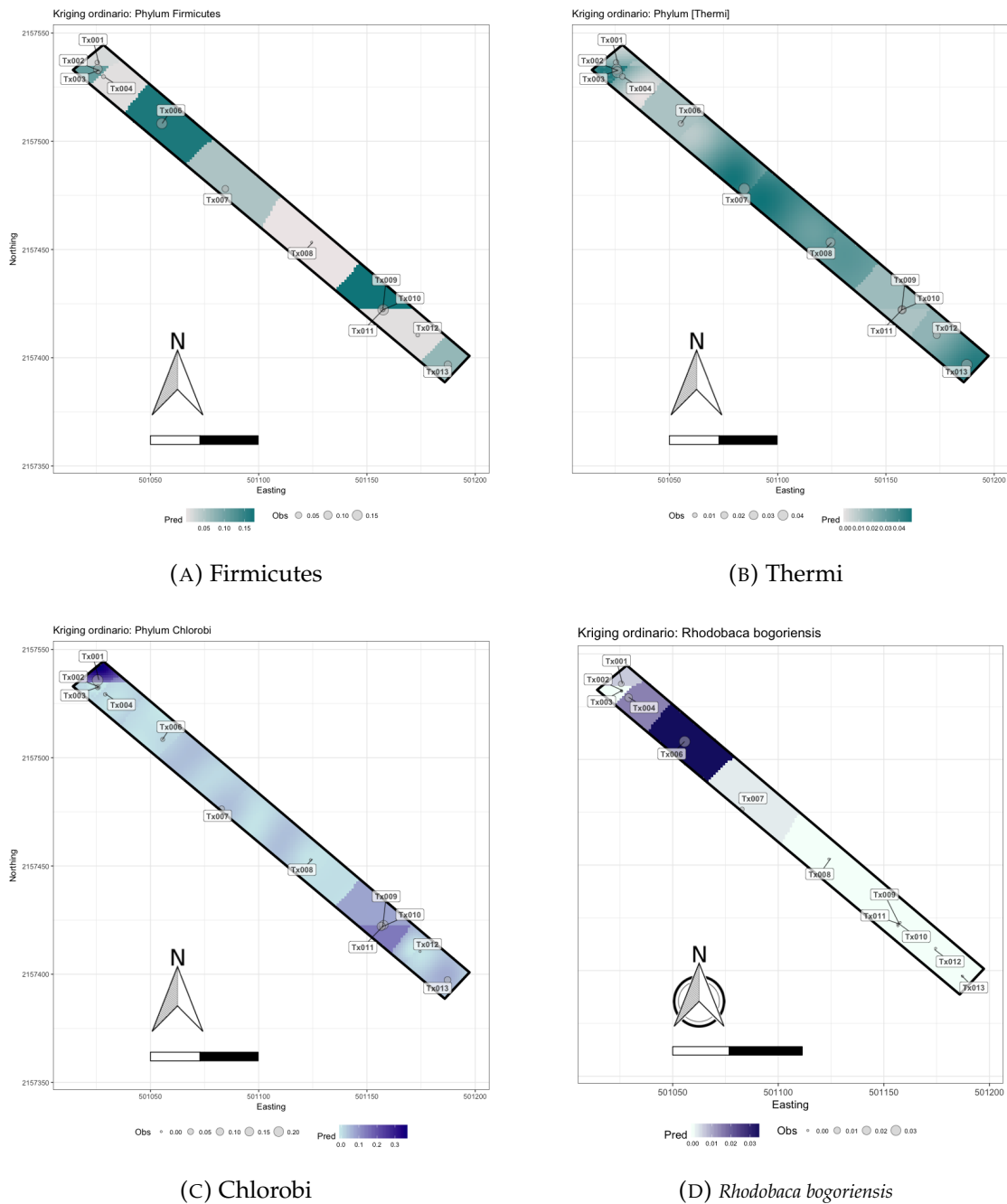


FIGURA 5.14: **Mapas de distribución de Bacteria.** Distribución de las abundancias relativas de los phyla (A) Firmicutes, (B) Thermi, (C) Chlorobi y (D) *Rhodobaca bogoriensis*. Obs: valores observados en cada sitio del transecto (Tx001-Tx013); Pred: Valores predichos para la distribución de abundancia relativa en un polígono B con 3000 píxeles de 1.29m<sup>2</sup> cada uno.

### 5.5.1. Correlaciones: comunidades y características

#### 5.5.1.1. Arqueas y sus conductores

Los índices de diversidad estuvieron correlacionados positivamente con el contenido de agua (Shannon  $\rho=0.74$ ) y el contenido de C inorgánico (Shannon  $\rho = 0.71$ ), y negativamente con el contenido de limo (Shannon  $\rho=-0.57$ ) ( $p<0.05$ ).

En el caso de los phyla de Arquea, Euryarchaeota presentó una correlación positiva con la conductividad electrolítica ( $\rho=0.7$ ) al contrario de Thaumarchaeota ( $\rho=-0.71$ ), mientras que el phyla Woesearchaeota, encontrado por asignación manual, se correlacionó negativamente con el pH del suelo (5.15). Mientras tanto, en el caso de los géneros, la abundancia relativa de Cenarcheum estuvo correlacionada positiva y significativamente con el contenido de arena ( $\rho=0.76$ ) y negativamente con el contenido de limo ( $\rho=-0.67$ ), C. Nitrososphaera se correlacionó negativamente tanto con la EC como con el contenido de C inorgánico ( $\rho<-0.6$ ) por otro lado la abundancia de Natronomonas mostró una correlación positiva con EC ( $\rho=0.71$ ,  $p<0.05$ ) y Halostagnicola presentó una correlación negativa con el contenido de C inorgánico (Figura 5.15). El análisis de correlaciones también reveló que el género Halobacteriales cambia según las coordenadas, al Este negativamente y al Norte positivamente.

A nivel de especies detectadas, la mayoría estuvo correlacionada con la conductividad electrolítica, tanto positiva como negativamente, a excepción de *Natronococcus jeotagli* y *Natronorubrum tibetense* que no mostraron estar correlacionadas con las características medidas y *Halalkalicoccus jeotagli* que presentó correlación negativa con el contenido de arena (Fig. 5.16). *Nitrososphaera viennensis* mostró, además, correlación negativa con el contenido de agua y *Natronomonas moolapensis* correlación positiva con el contenido de C inorgánico, además de cambiar con las coordenadas (Fig. 5.16). Las especies que no lograron ser asignadas mostraron relación tanto con la conductividad como con el contenido de C inorgánico, lo que sin duda se relaciona



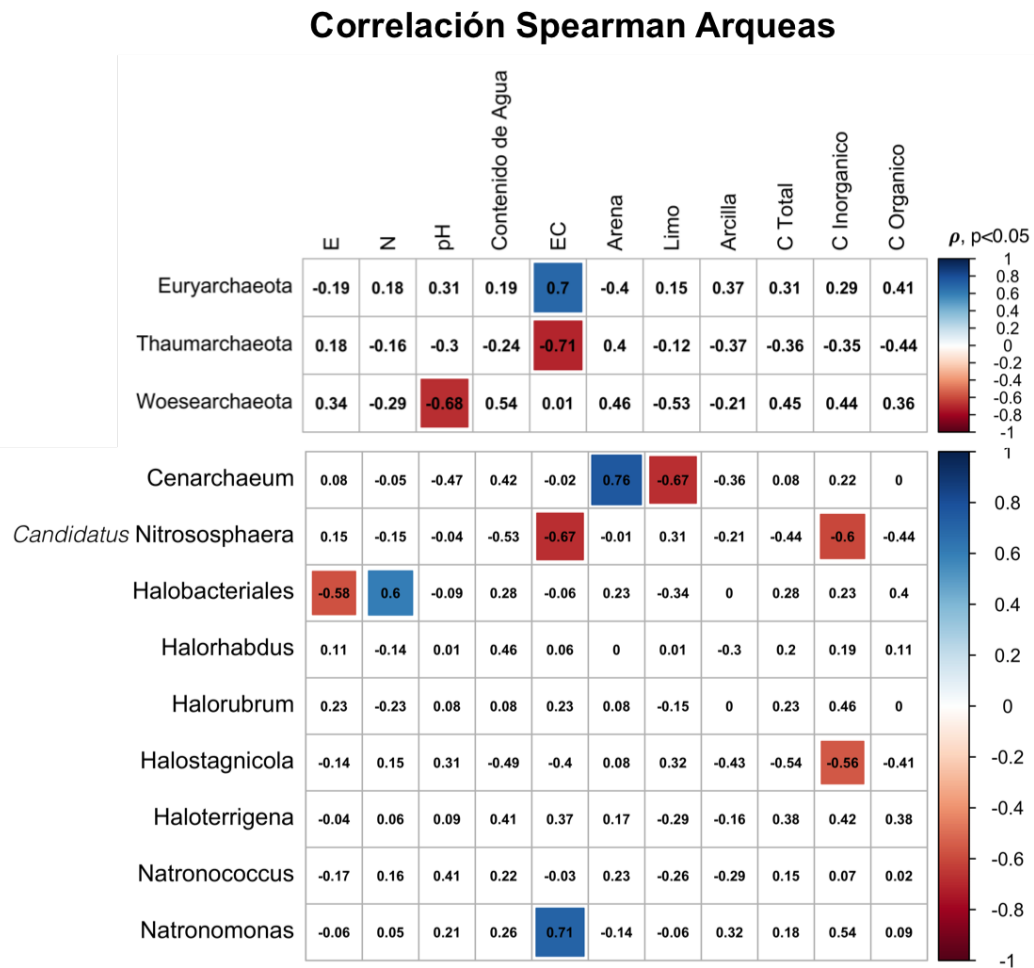


FIGURA 5.15: **Correlaciones Spearman para Phyla y Géneros Arquea.** Heatmap con coeficientes de correlación  $\rho$  entre arqueas y características del suelo, incluido Este y Norte. Según la escala de color se muestran únicamente aquellos con una correlación significativa ( $p < 0.05$ ). E: Este; N: Norte; EC: Conductividad electrolítica

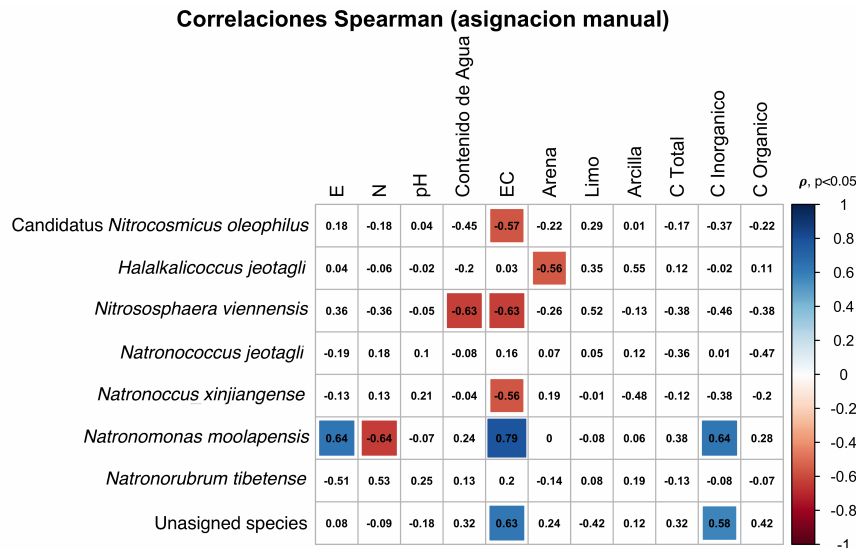


FIGURA 5.16: **Correlaciones Spearman para especies Arquea.** Heatmap con coeficientes de correlación ( $\rho$ ) entre especies de arqueas y características del suelo, incluido Este (E) y Norte (N). Según la escala de color se muestran únicamente aquellos con una correlación significativa ( $p < 0.05$ ). EC: Conductividad electrolítica

con especies aún desconocidas presentes en los ambientes salinos y alcalinos extremos, hecho que también se ve reflejado con la cantidad de secuencias obtenidas, los índices de riqueza y las curvas de rarefacción.

El CAP agrupó la mayoría de los sitios en el cuadrante superior e inferior izquierdo, excepto los sitios Tx008 y Tx012 (Fig. 5.17). Los sitios Tx008 y Tx012 estuvieron caracterizados por un amplio valor para la dimensión 1 (CAP1), por ejemplo un valor alto en la abundancia de *C. Nitrososphaera*, un alto contenido de limo y bajos valores de pH, EC y H. Los demás sitios fueron separados por las diferencias en la abundancia relativa de la mayoría de los géneros Arquea (*Haloterrigena* y *Halostagnicola*, y diferencias en las características del suelo tal como el contenido de C inorgánico y orgánico, el contenido de arcilla y arena). El análisis basado en divergencia que compara las comunidades de Arquea, es decir, UniFrac PCoA, mostró una clara separación entre las muestras Tx-008 y Tx-012 y las demás que se observan

agrupadas juntas (Figs. D.4a y D.4b).

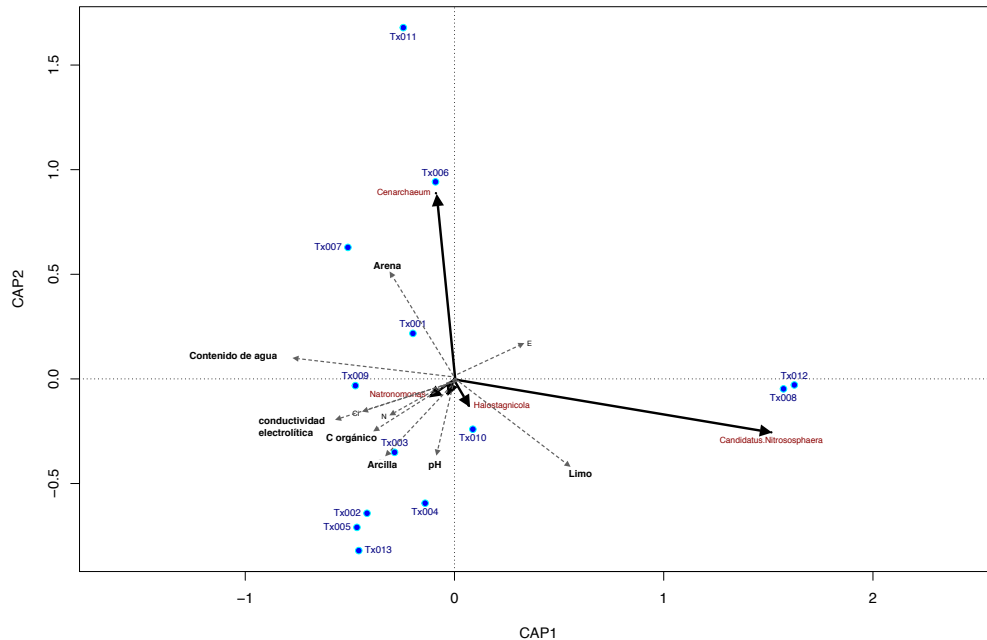


FIGURA 5.17: **CAP de Géneros Arquea.** Análisis canónico de coordenadas principales (CAP) con score para sitios (puntos azules), sumas ponderadas de scores de especies (Vectores continuos) y scores biplot (líneas discontinuas), aplicando distancias Bray para los géneros de Arquea asignados (Dimensión 1 explica el 44.5 % de la variación y Dimensión 2 explicando el 16.89 % de la variación).

De acuerdo a los análisis BIOENV, una combinación de contenido de agua, pH y contenido de arcilla fue el mejor conjunto de parámetros que explicaría la composición de las comunidades de Arqueas (Tabla 5.11).

### 5.5.1.2. Bacterias y sus conductores

No se presentó correlación alguna entre los parámetros de  $\alpha$  diversidad y las características del suelo medidas.

La abundancia relativa de Chlorobi estuvo correlacionada negativa y significativamente ( $p < 0.01$ ) con el contenido de limo ( $\rho = -0.74$ ) y positivamente con el contenido de agua ( $\rho = 0.73$ ), mientras que Spirochaetes mostró una correlación positiva con C inorgánico ( $\rho = 0.62$ ,  $p < 0.05$ ) Figura 5.18.

El contenido de Agua estuvo correlacionado con la abundancia relativa de los phyla Chlorobi ( $\rho = 0.73$ ) y Proteobacteria ( $\rho = -0.68$ ), además Chlorobi presentó correlaciones con la mayoría de los parámetros fisicoquímicos, positivamente con arena y C inorgánico y negativamente con limo ( $p < 0,05$ ). (Figura 5.18). Por otro lado al realizar el análisis de correlación con los phyla bacterianos asignados manualmente se encontró que algunos phyla se encuentran correlacionados con las mismas características que la asignación por Qiime, mientras que otros presentaron otras correlaciones. La abundancia relativa del género B-42 estuvo correlacionada negativamente de forma significativa al contenido de C orgánico ( $\rho = -0.68$ ) y KSA1 positivamente con el contenido de C inorgánico ( $\rho = 0.89$ ,  $p < 0.01$ ) (Fig. 5.18). El análisis basado en la divergencia comparando comunidades bacterianas, es decir, PCoA UniFrac no ponderado, separa las muestras Tx-008 y Tx012 de las otras muestras de suelo, mientras que PCoA UniFrac ponderado separa las muestras de suelo en tres grupos, uno con Tx-008 y Tx012, uno con Tx-002, Tx003, Tx-004, Tx-006 y Tx-010 y el resto (Figs. D.4c y D.4d).

El CAP de los phyla bacterianos y las características del suelo separó los sitios Tx008 y Tx012 de los otros sitios (figura 5.19a) y dichas muestras estuvieron caracterizadas por una gran abundancia relativa de Proteobacteria (5.19a). Por otro lado el análisis CAP con los géneros bacterianos mostró que las muestras Tx008 y Tx012 tuvieron altas abundancias relativas de Amaricoccus, Kaistobacter, Paracoccus y Pontibacter a diferencia de los sitios restantes (figura 5.19b). De forma general, la combinación de contenido de agua, EC y contenido de limo en conjunto fueron los que mejor correlacionaron con la abundancia relativa de los Phyla bacterianos (Tabla

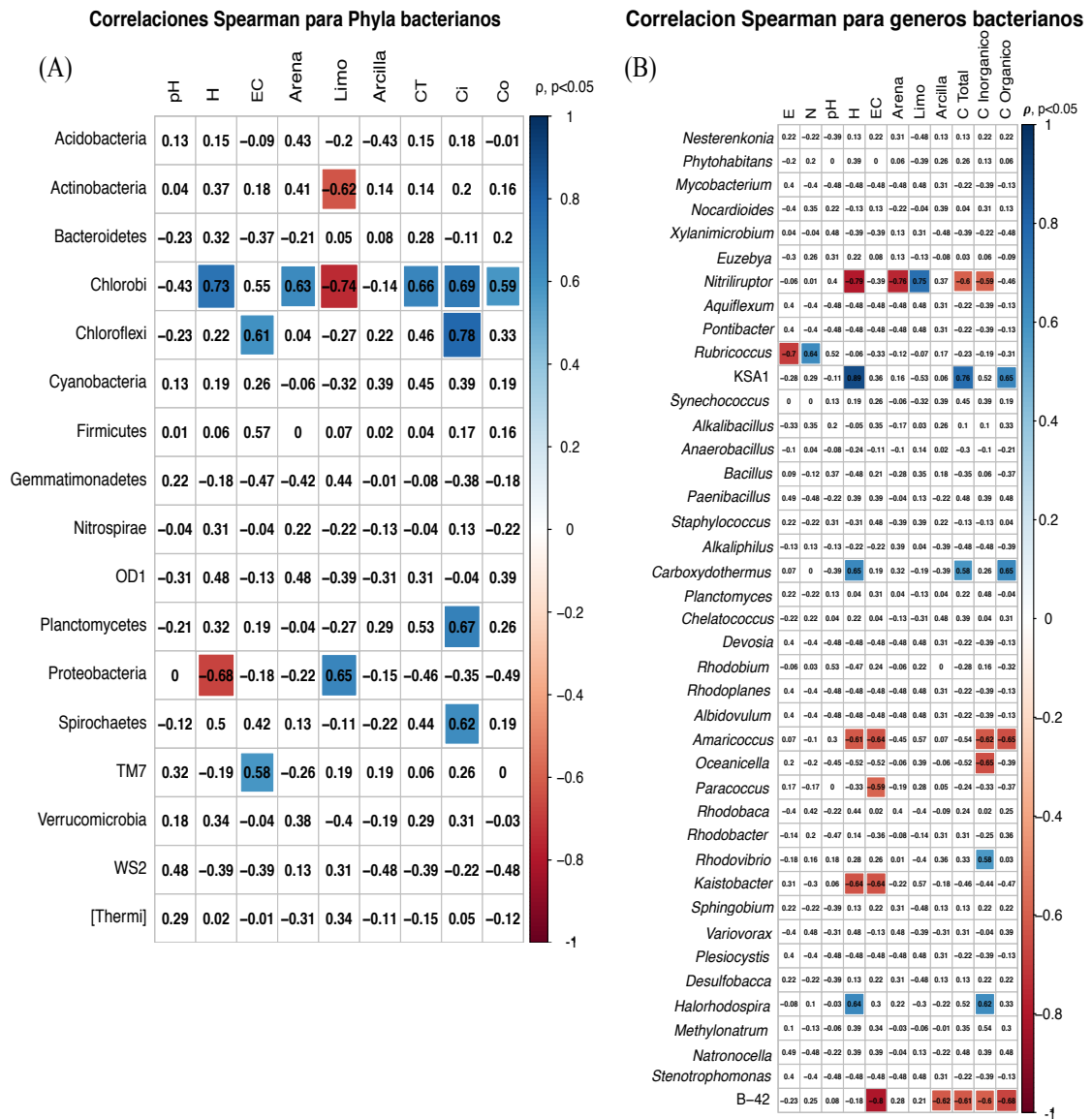
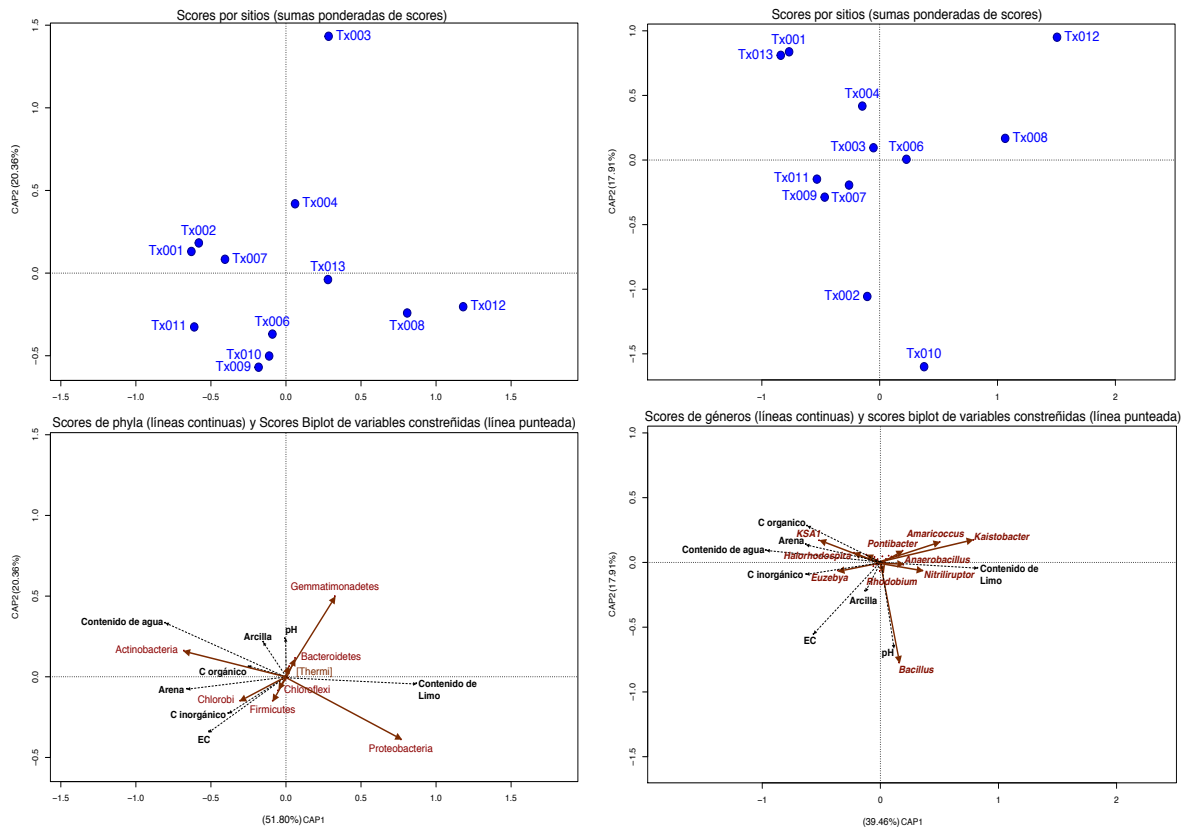


FIGURA 5.18: Correlaciones Spearman para géneros y phyla de Bacteria. Heatmap con coeficientes de correlación ( $\rho$ ) entre (A) phylas de bacterias, (B) géneros de Bacteria y características del suelo, incluido Este (E) y Norte (N) para phyla. Según la escala de color se muestran únicamente aquellos con una correlación significativa ( $p < 0.05$ ). EC: Conductividad electrolítica; CT: contenido Total de C; Ci: contenido de C inorgánico; Co: contenido de C orgánico.



(A) CAP Phyla Bacteria

(B) CAP géneros Bacteria

FIGURA 5.19: **Análisis CAP de Bacteria.** Análisis canónico de coordenadas principales (CAP) con score para sitios (arriba, puntos en azul), sumas ponderadas de scores de especies y scores biplot (abajo) aplicando distancias Bray para (A) los phyla bacterianos asignados (Dimensión 1 explica el 51.8% de la variación y Dimensión 2 explicando el 20.36% de la variación) y (B) los géneros bacterianos (Dimensión 1 explicando el 39.46% de la variación y Dimensión 2 explicando el 17.91% de la variación).

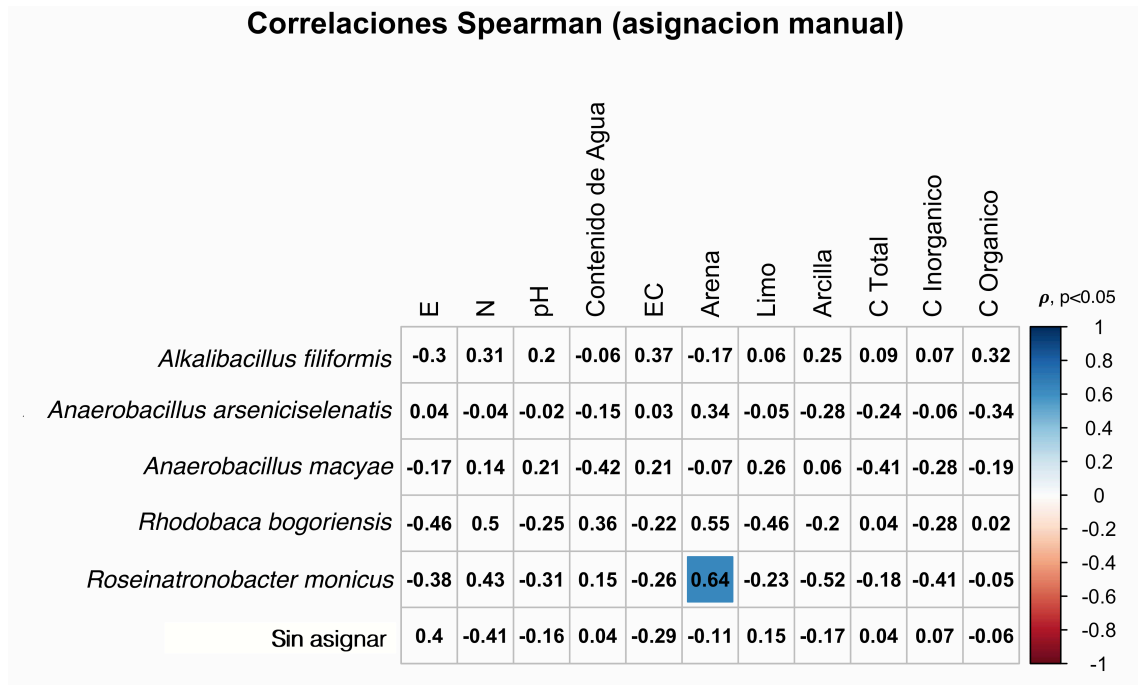


FIGURA 5.20: **Correlaciones Spearman para especies de Bacteria.** Heatmap con coeficientes de correlación ( $\rho$ ) entre especies de Bacteria y características del suelo, incluido Este (E) y Norte (N) para phyla. Según la escala de color se muestran únicamente aquellos con una correlación significativa ( $p < 0.05$ ). EC: Conductividad electrolítica; C: Carbono.

5.11).

En el presente estudio, la abundancia relativa de Chlorobi mostró la mayor variabilidad a lo largo del transecto y se correlacionó negativamente con el contenido de limo y positivamente con el contenido de agua. La abundancia relativa de Firmicutes también fue muy variable, pero no controlada por las características del suelo en este ecosistema. Los firmicutes son ubicuos en el suelo y se pueden encontrar en la mayoría de los ecosistemas (por ejemplo Liu et al. 2014 y Santini, Warren y Kendr 2015). Methylostrum fue descrito por primera vez en lagos hipersalinos en Rusia por Sorokin et al. en 2007. Lo describieron como perteneciente a la Gammaproteobacteria que fue capaz de crecer en sustratos C1. Clasificado como un género metilótrofo,

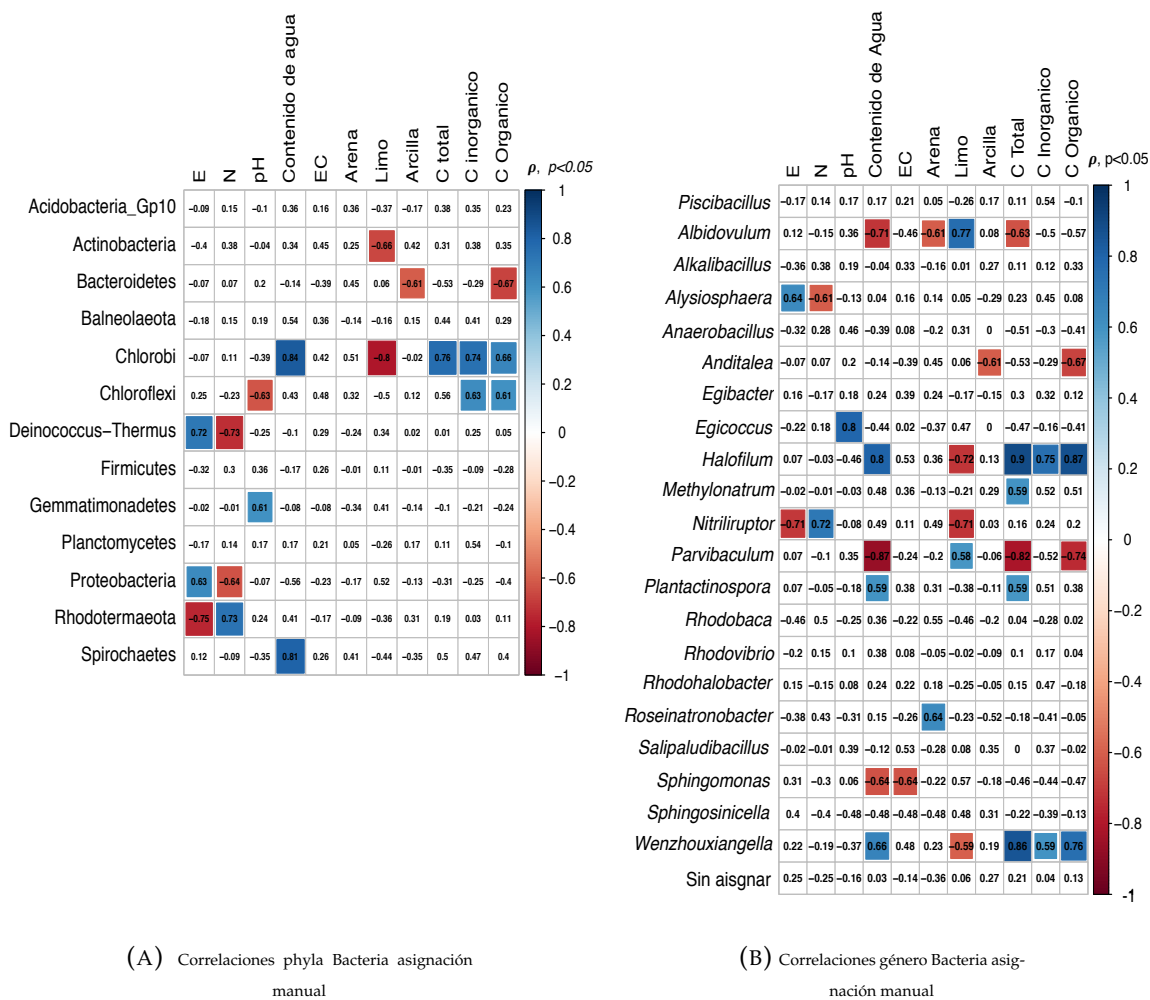


FIGURA 5.21: Correlaciones Spearman para asignación manual de Bacteria. Correlaciones de (A) phyla y (B) géneros bacterianos con propiedades fisicoquímicas del suelo.



**TABLA 5.11: Conjunto de variables con la mejor correlación para las comunidades microbianas.** Distancias Euclidianas de las variables escaladas que tienen la máxima correlación  $r$  con las disimilitudes en las comunidades (Función BIOENV del paquete Vegan en R). La tabla muestra los coeficientes Spearman ( $r$ ) para la correlación entre las comunidades bacterianas y arquea del suelo y las propiedades fisicoquímicas del mismo.

	Subconjuntos de Parámetros	Coefficiente $r$
Comunidades Arquea Phylum	pH H <sup>a</sup> EC <sup>b</sup> Arcilla	0.6386146
Comunidades Arquea géneros	pH H Arcilla	0.6133534
Comunidades Bacteria Phylum	H EC Limo	0.5359622
Comunidades Bacteria géneros	pH H Arena Arcilla	0.6764083

<sup>a</sup>H: % agua ; <sup>b</sup>EC: Conductividad electrolítica.

*Methylostratum* puede crecer bajo condiciones de saturación de sal y alcalófilas, cuyo pH óptimo alcanza 10.5. En el presente estudio, este género mostró una alta correlación con el contenido inorgánico de C, y se detectó a pH 10.6, lo cual es consistente con los hallazgos de Simachew et al. 2015. Se identificó *Methylostratum* solo en un sitio salino intermedio (25 %) dominado por Archaea. KSA1 y B-42, el segundo y tercer género más abundante en este estudio, son grupos bacterianos poco conocidos. Fueron encontrados por primera vez en sedimentos lacustres comparables al suelo del antiguo lago Texcoco. La división candidata KSA1, el segundo género más abundante encontrado en este estudio, se observó por primera vez en un ambiente de lodo negro rico en sulfuros hipersalinos en un pantano y se clasificó como perteneciente a Bacteroidetes (Tanner et al. 2000; McDonald et al. 2012). Se encontraron miembros de B-42 en sedimentos de agua dulce y lodo activado de una planta de tratamiento de aguas residuales (Sánchez et al. 2011, [número de acceso Genbank: FN598000]; Hamamura et al. 2014, [número de acceso: KC852965.1]; Chen et al. 2015). También se encontraron en suelos contaminados con Hidrocarburos policíclicos aromáticos

(PAHs) (Tejeda-Agredano et al. 2013, [número de acceso: JQ771979.1]).

## 5.6. Patrones de distribución espacial

Se intentaron describir los cambios en las comunidades y su dependencia a las características del suelo, detallando patrones de distribución espacial a través de un transecto de muestreo. Estudios previos del área heterogénea extrema en la zona federal del exlago de Texcoco proporcionaron pistas para elegir un área, establecer una escala y una dirección de muestreo para evaluar la influencia relativa de las características del suelo medidas. De dichos estudios, se sospechaba la presencia de gradientes en los valores de pH y EC del suelo a lo largo del transecto de muestreo, sin embargo, a la escala seleccionada, sólo se observó una alta variabilidad en los valores de EC con un rango reducido en los valores de pH. En muchos estudios se ha encontrado que el pH es un importante impulsor de la composición bacteriana incluso cuando no está presente una gran variación (Constancias et al. 2015), sin embargo, en un suelo extremo tal como el del exlago de Texcoco y a una escala local, la influencia directa del pH sobre los microorganismos está sobrepasada por otras características. Los parámetros tales como EC y el contenido de agua exhibieron una alta variabilidad a lo largo del transecto y de igual forma se encuentran altamente correlacionados y el patrón de distribución irregular observado para ambos mostró una estrecha relación con los cambios abruptos en las comunidades Archaea. Por lo tanto, podemos relacionar los parches de vegetación en el área extrema con los bajos valores de EC y la presencia del phylum Thaumarchaeota. Se han detectado miembros de Thaumarchaeota en sitios oligotróficos y se han correlacionado con un menor pH y contenido de sal, especialmente en bajas concentraciones de amonio (Wessén et al. 2011; Marusenko et al. 2013; Bollmann, Bullerjahn y McKay 2014), parámetro que no fue medido en el presente estudio.

Los mapas de distribución bacteriana revelaron un patrón con poca variabilidad a través del transecto para los phyla más abundantes (Actinobacteria, Proteobacteria y Bacteroidetes) aunque para Proteobacteria se observan las mayores abundancias en los sitios con valores menores de EC. La poca variabilidad de dichas Taxas puede deberse a que poseen géneros más adaptados a condiciones de salinidad y alcalinidad (Ma y Gong 2013). Por otro lado para phyla menos abundantes tal como Chlorobi y Firmicutes se puede observar un comportamiento más heterogéneo, y se observó que la distribución de Chlorobi depende de la mayoría de las características medidas del suelo en el presente estudio, a diferencia de Firmicutes que no estuvo correlacionada con característica alguna.

La elección de la escala es importante para definir patrones espaciales. Se ha observado que a escala local algunos factores muestran homogeneidad espacial, como se encontró con el pH en este estudio. Difiere de los estudios a gran escala, donde el pH puede ser más variable y, en consecuencia, ser el principal impulsor de la distribución de las comunidades, como lo muestra Navarro-Noya et al. (2015). La variabilidad espacial de phyla a escala local puede ser baja ya que los miembros del mismo phylum pueden compartir características y actividades metabólicas similares, mientras que los cambios en la estructura y composición se destacan por la abundancia de orden/familias (Crits-Christoph et al. 2013). Aunque las variaciones temporales no se investigaron en este estudio, se ha informado una fuerte correlación en escalas temporales y espaciales (Lauber et al. 2013; Regan et al. 2014; Wasserstrom et al. 2017).

## Capítulo 6

### Conclusión

*«When radium was discovered, no one knew that it would prove useful in hospitals. The work was one of pure science. And this is a proof that scientific work must not be considered from the point of view of the direct usefulness of it.»*

---

*Marie Curie*

- Las características del suelo medidas mostraron dependencia espacial de fuerte a media, observada a través de la presencia de autocorrelación espacial a diferentes distancias.
- Las características del suelo y las comunidades Arquea y Bacteria se presentaron en forma periódica a través del transecto de muestreo.
- Se observaron distribuciones de comunidades del suelo en un ambiente extremadamente heterogéneo utilizando un enfoque geoestadístico.
- La distribución del phylum Thaumarchaeota se observó más abundante en las localidades con los valores más bajos de EC y H, en presencia de parches de

*Distichlis spicata* L. y en consecuencia, el phylum Euryarchaeota estuvo ausente o con muy baja abundancia en estos puntos. La distribución del phylum Woeisearchaeota, que se encontró en baja abundancia, no mostró seguir un patrón dependiente de EC o H si no con la distribución del pH, con el cuál presentó correlación negativa. En cuanto a la distribución de especies, la especie encontrada en mayor abundancia, *Natronococcus xinjiangense*, siguió una distribución un tanto opuesta a EC pero también mostrando parches de alta abundancia.

- Las bacterias fueron menos afectadas por factores fisicoquímicos, pero la presencia o ausencia de grupos bacterianos se relacionó con H, pH y EC.
- La distribución de los phyla bacterianos fue igualmente periódica, se observaron parches de mayor abundancia para Proteobacteria en áreas de bajos valores de EC y H. La variabilidad de las abundancias para los phyla dominantes fue menor que las de sus contraparte Arquea, sin embargo en phyla menores como Firmicutes y Chlorobi la variabilidad de abundancias relativas en el polígono fue mayor.
- El kriging ordinario para algunas características del suelo y comunidades explicó adecuadamente la variabilidad en los sitios de muestreo. Para otros, sin embargo, se requiere un muestreo más intensivo.
- El principal factor de conducción para explicar la estructura de las comunidades microbianas fue H, seguido de EC. A diferencia de lo observado a otras escalas en este mismo suelo o en suelos con características no extremas la variabilidad del pH fue menor y por lo tanto su efecto en la distribución fue sobrepasado por las ya mencionadas características.

- El uso de nuevas herramientas ómicas y análisis geoestadísticos puede determinar la funcionalidad geoquímica en un área y ayudar a diseñar nuevas prácticas de gestión para ecosistemas extremos o aquellos con actividad antropogénica.

# Apéndice A

## Fisicoquímicos y exploratorios

### A.1. Estadística descriptiva

#### A.1.1. Box-plot fisicoquímicos

Gráficas descriptivas de la distribución en Box-plot para las características del suelo medidas a lo largo del transecto (Fig. A.1).

#### A.1.2. Pruebas de normalidad y tendencia

Gráficas de evaluación de la distribución y tendencia de las características del suelo (Figs. A.2 y A.3).

### A.2. Geoexploratorios

Con el fin de conocer los datos espacialmente se realizan análisis geoexploratorios a cada variable medida en el suelo (Figs. A.4 y A.5), para lo cual se utiliza el paquete GeoR en el entorno R (Ribeiro y Diggle 2016) y el paquete ggmap (Kahle y Wickham 2013).

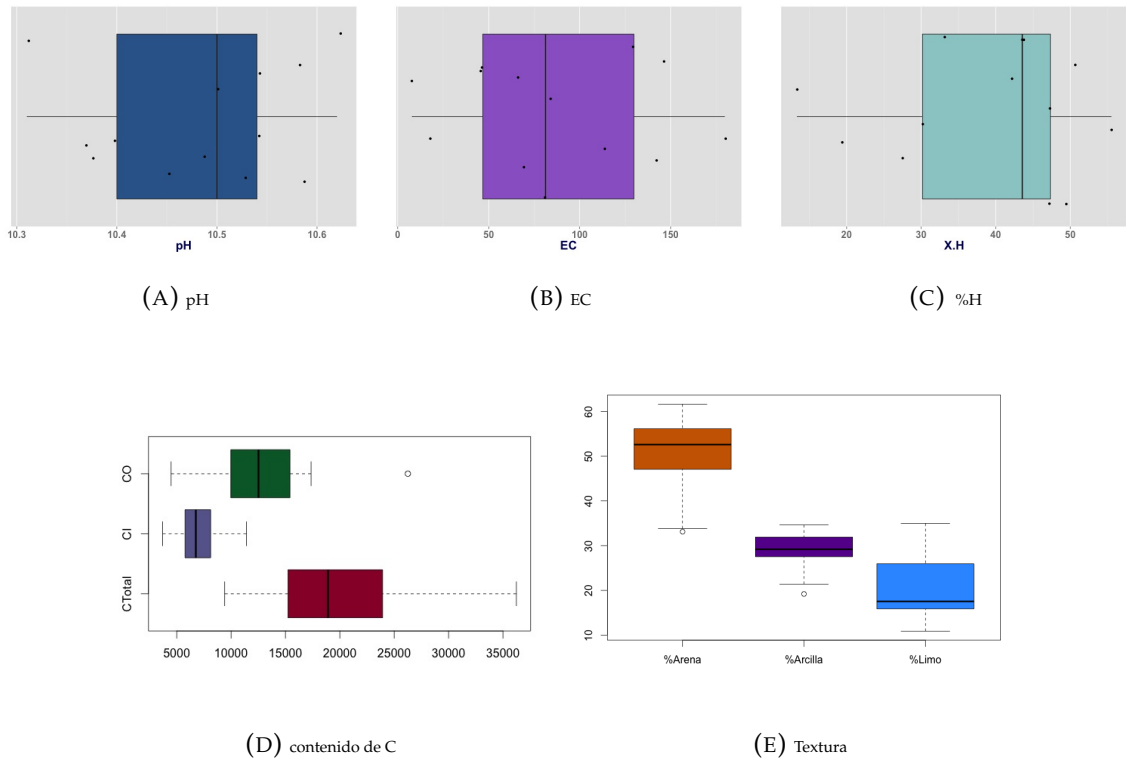
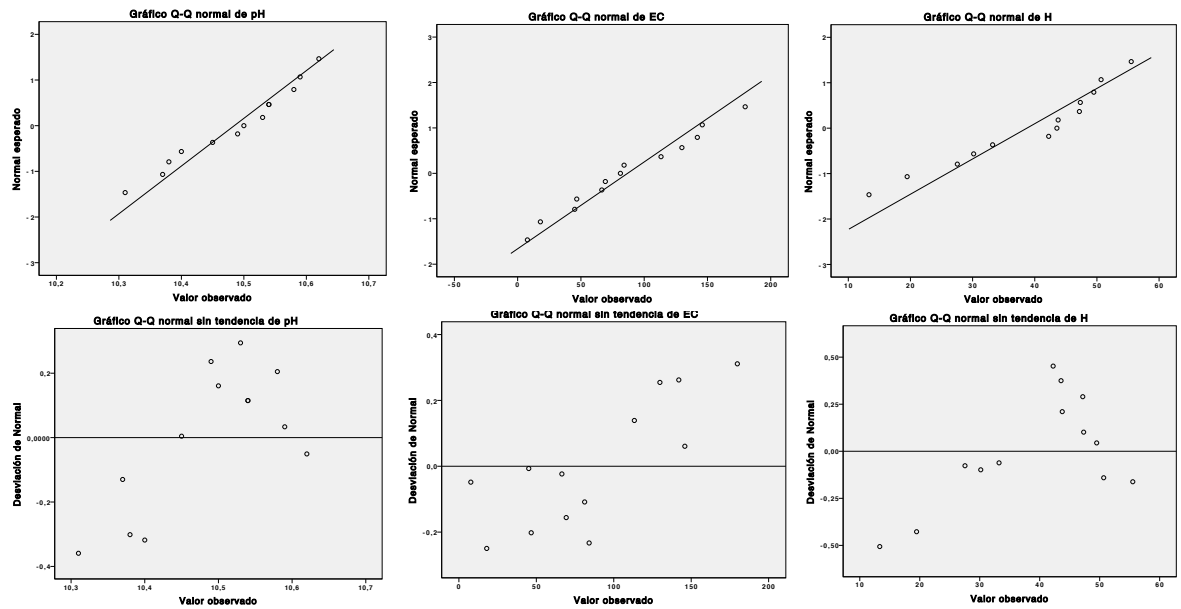


FIGURA A.1: **Box-plot Físicoquímicos** Estadística descriptiva de (A) pH, (B) Conductividad electrolítica, (C) contenido de Agua, (D) contenido de C y (E) Tamaño de partícula

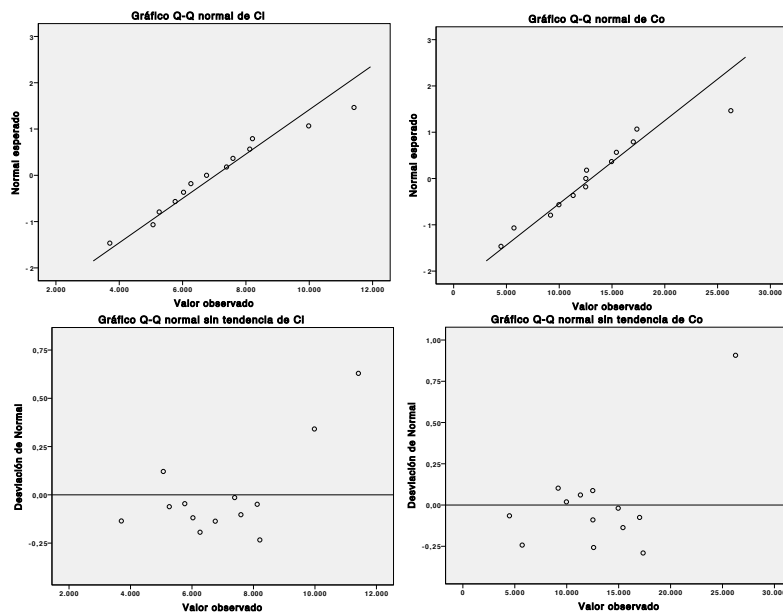




(A) pH

(B) EC

(C) %H



(D) Ci

(E) Co

FIGURA A.2: Q-Q plot Físicoquímicos Pruebas de normalidad de (A) pH, (B) conductividad electrolítica, (C) contenido de Agua, (D) contenido de C inorgánico y (E) Contenido de C orgánico

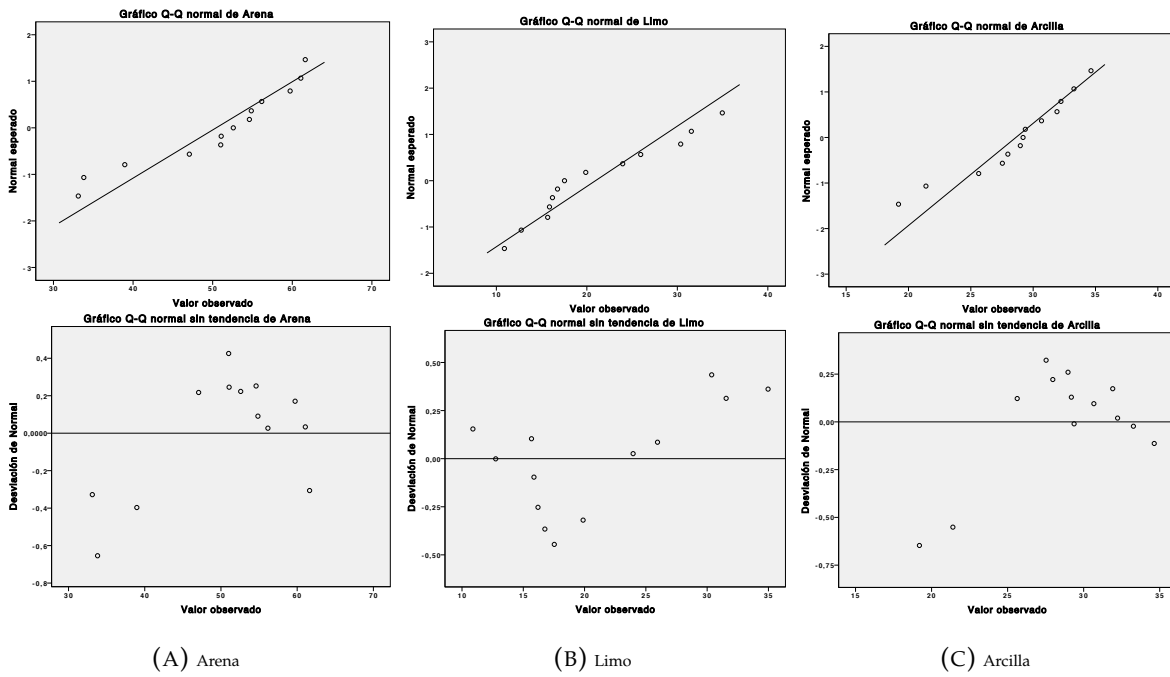
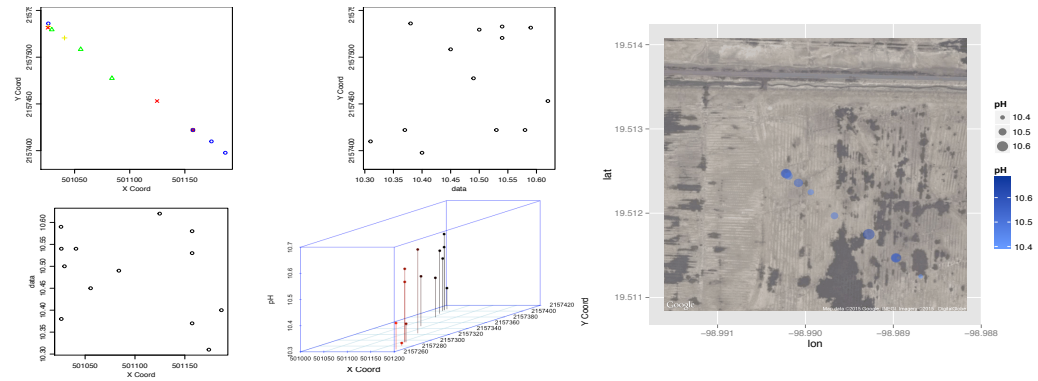
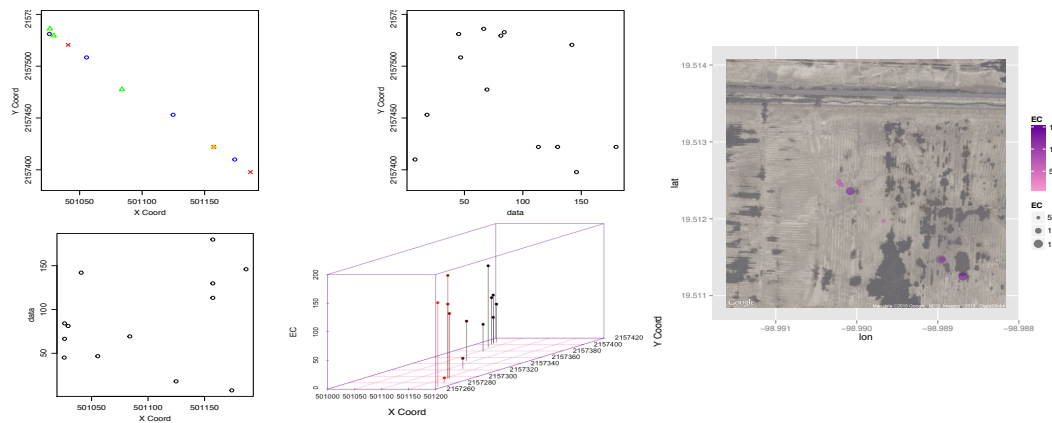


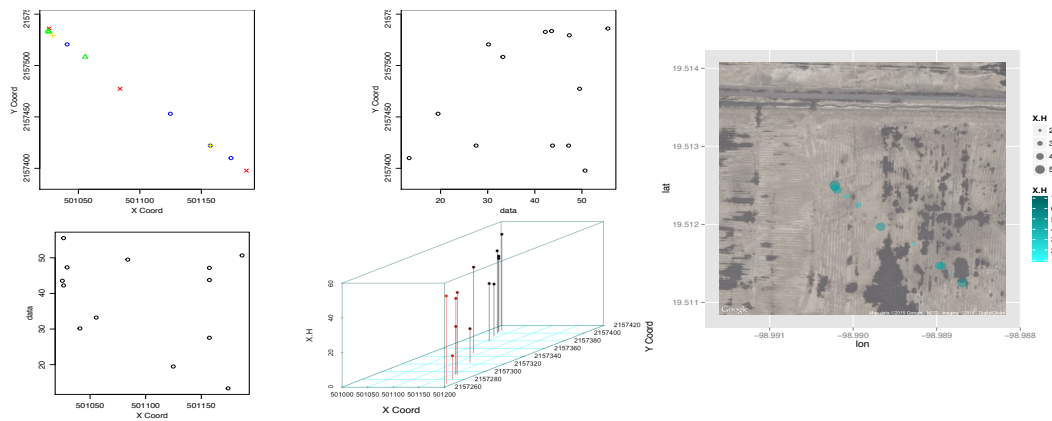
FIGURA A.3: Q-Q plot Físicoquímicos Pruebas de normalidad del contenido de (A) arena, (B) arcilla y (C) limo.



(A) Situación geoespacial de pH

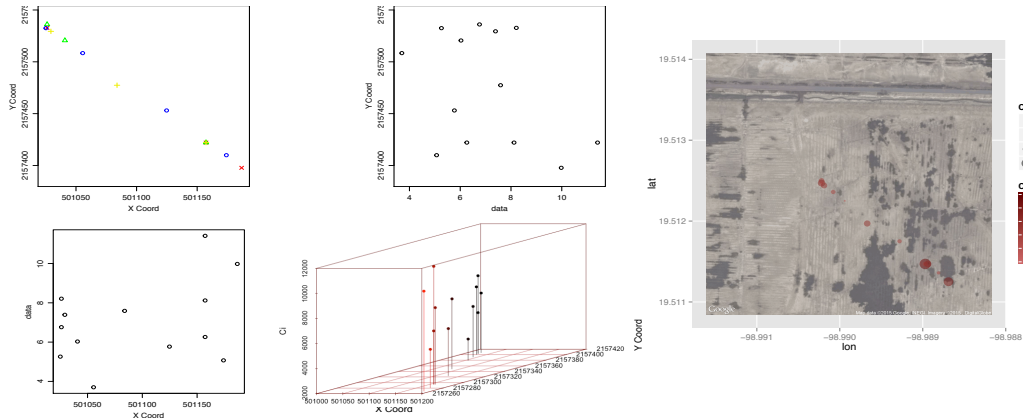


(B) Situación geoespacial de EC

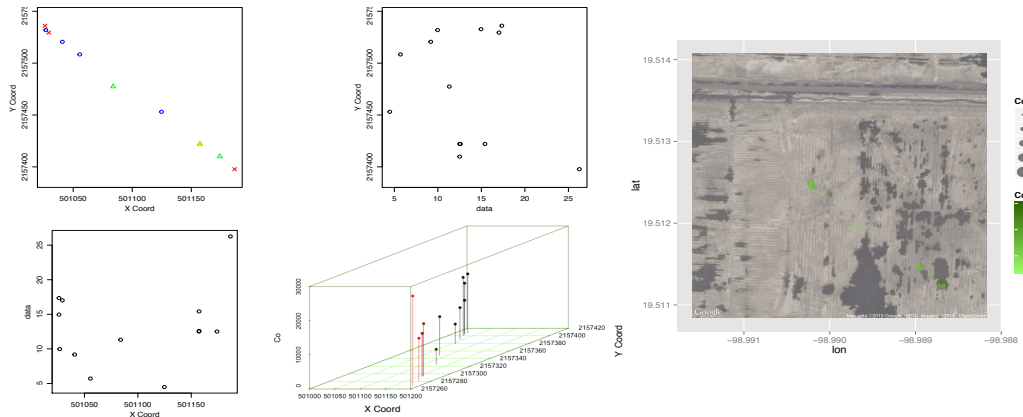


(C) Situación geoespacial de %H

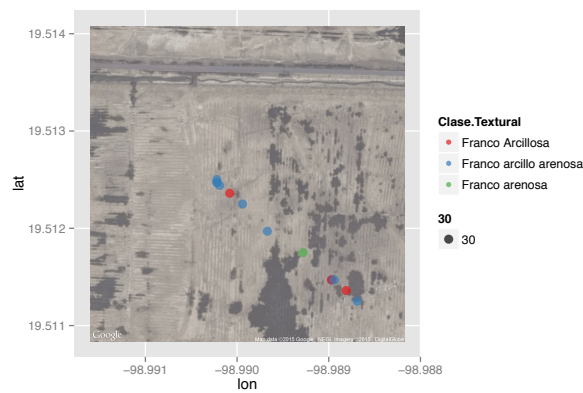
FIGURA A.4: **Geoexploratorios** Situación geoespacial de (A) pH, (B) conductividad electrolítica, (C) contenido de Agua



(A) Situación geoespacial de Ci



(B) Situación geoespacial de Co



(C) Situación geoespacial de la clase textural

FIGURA A.5: **Geoexploratorios** Situación geoespacial de (A) contenido de C inorgánico, (B) contenido de C orgánico y (C) clasificación textural

# Apéndice B

## Estructura Espacial

### B.1. Autocorrelogramas

Determinación de la dependencia espacial y estructura espacial de las características del suelo medidas a lo largo del transecto (Fig. B.1)

### B.2. Semivariogramas

Para detectar tendencias y ver el comportamiento de la semivarianza a cada par de distancias se observan los semivariogramas de nube (Fig. B.2).

Para determinar la dependencia espacial y la estructura presentada por las características del suelo también se realizan los semivariogramas por clase de distancia y el modelado usando un ajuste periódico (Fig. B.3)

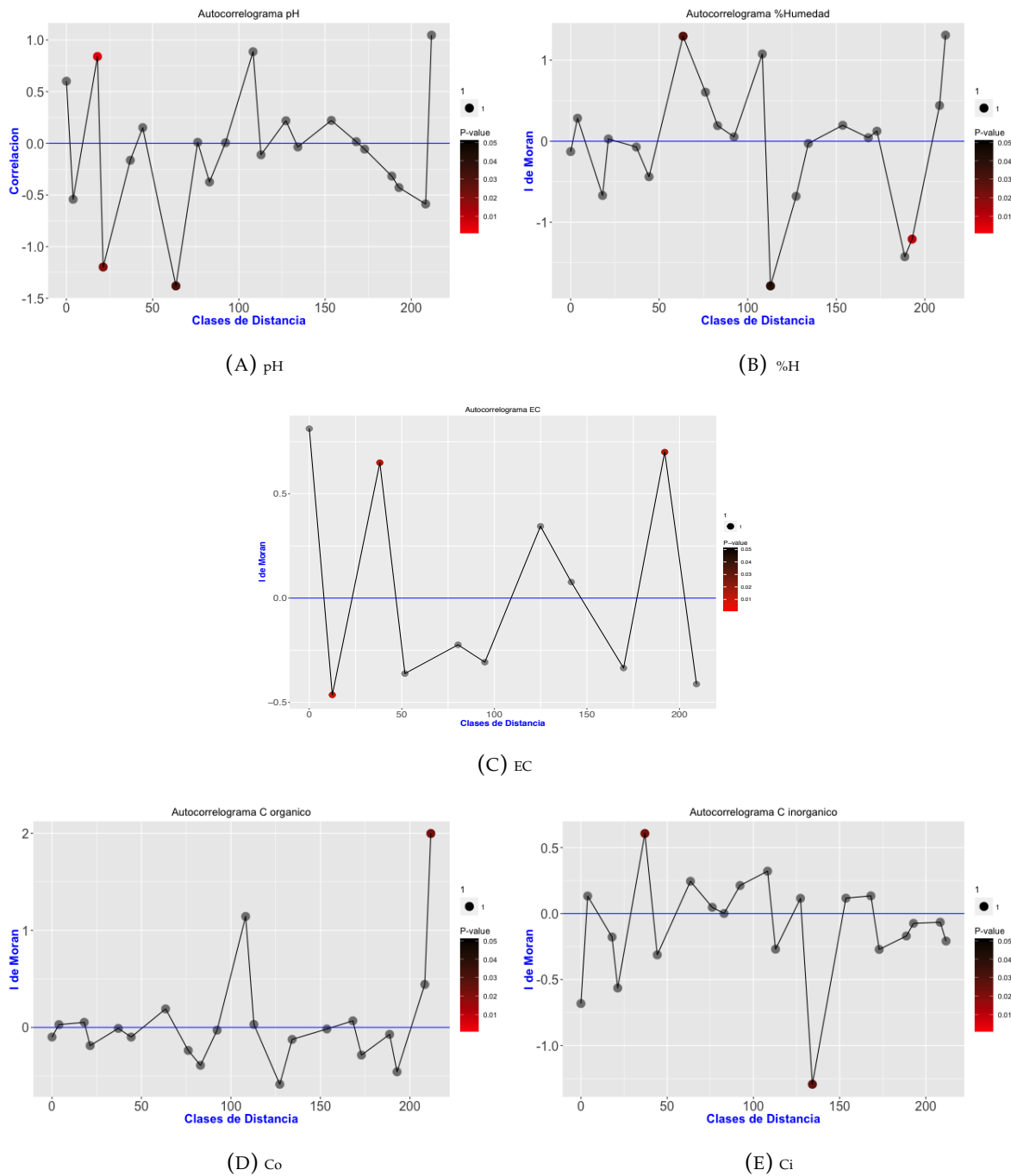


FIGURA B.1: **Autocorrelogramas.** Autocorrelación a diferentes clases de distancia de (A) pH, (B) EC, (C) contenido de agua, (D) contenido de C orgánico y (E) contenido de C inorgánico.

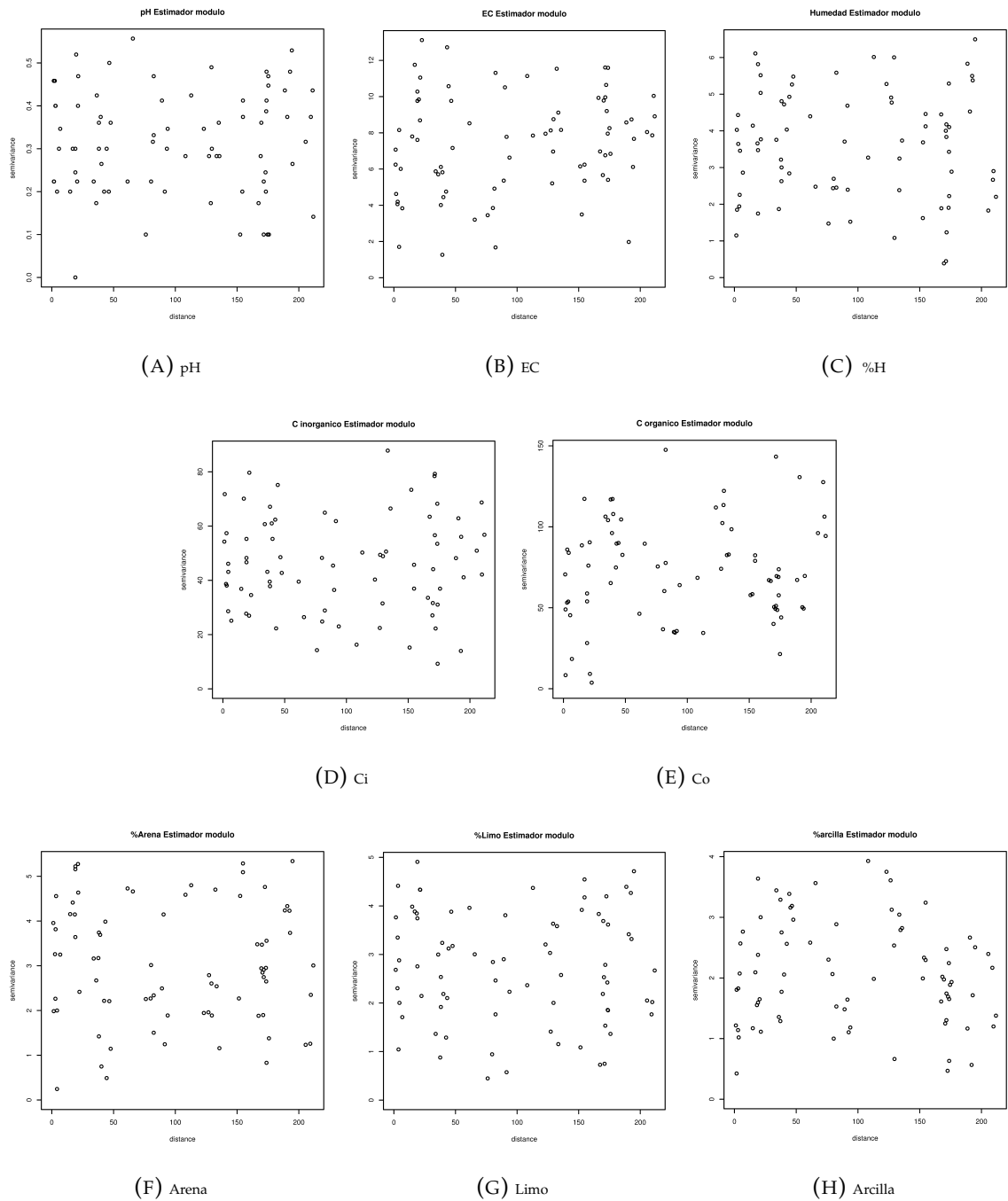


FIGURA B.2: **Semivariogramas** Estimador modulo en semivariogramas tipo nube (A) pH, (B) EC, (C) contenido de agua, (D) contenido de C inorgánico, (E) contenido de C orgánico, (F) contenido de arena, (G) contenido de limo y (H) contenido de arcilla

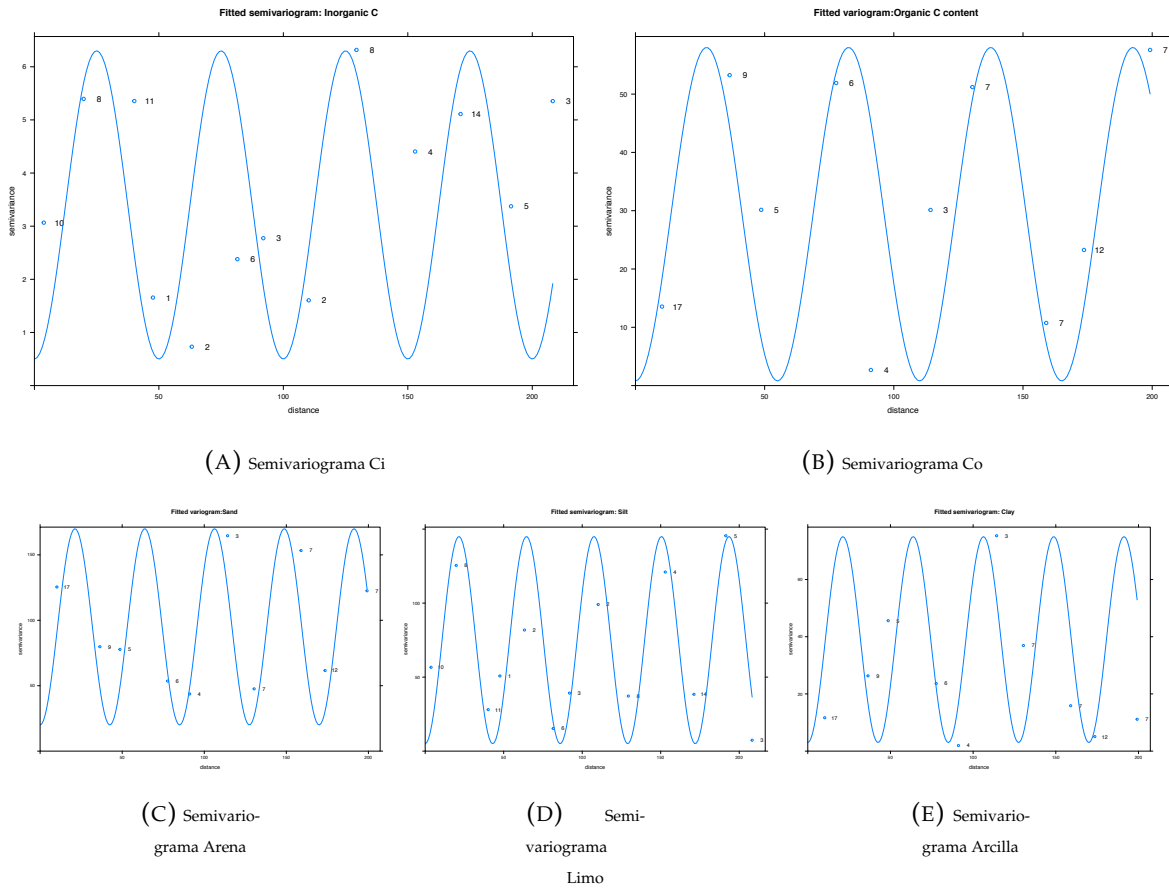


FIGURA B.3: **Semivariogramas** Ajuste de semivariogramas para (A) contenido de C inorgánico, (B) contenido de C orgánico, (C) arena, (D) limo y (E) arcilla



# Apéndice C

## Estructura de comunidades

### C.1. Curvas de rarefacción

Curvas de rarefacción de número de secuencias contra OTUs asignadas y contra Chao1 para Bacterias y Arqueas sin filtrar (Fig. C.1).

### C.2. Gráficas de niveles Taxonómicos

#### C.2.1. Asignación qiime

Resumen de Taxas asignados a clase, orden y familia para Arquea y Bacteria (Figs. C.2, C.3, C.4 y C.5).

#### C.2.2. Asignación manual

Resumen de Taxas asignados manualmente a clase, orden y familia para Arquea y Bacteria (Figs. C.6, C.7 y C.8).

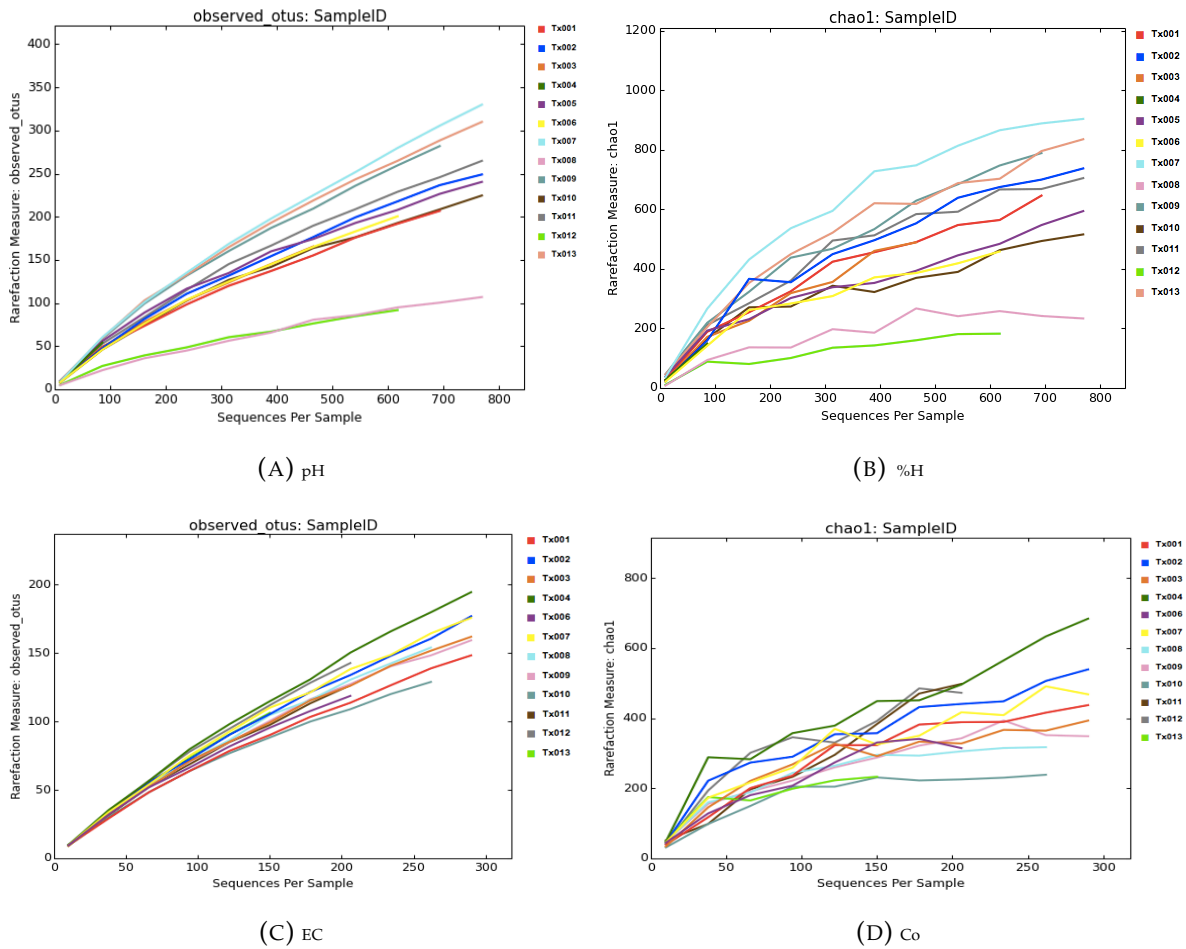
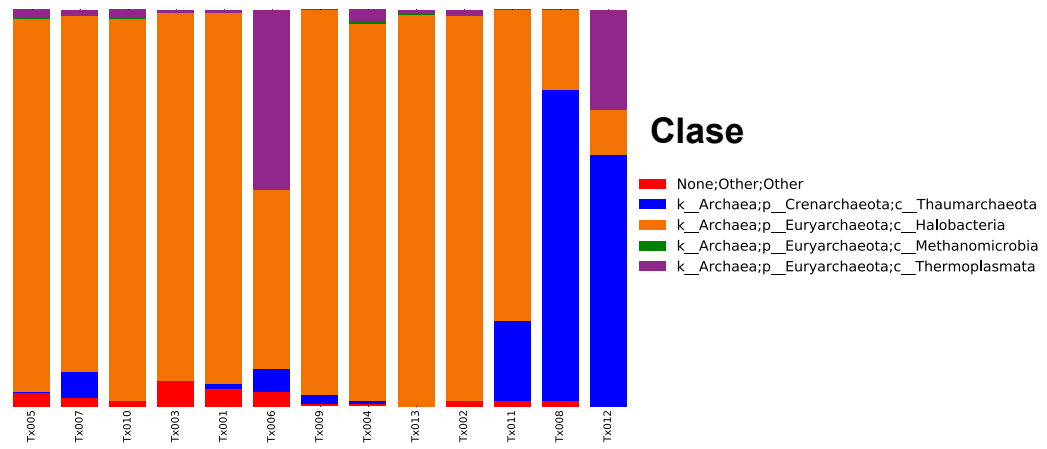
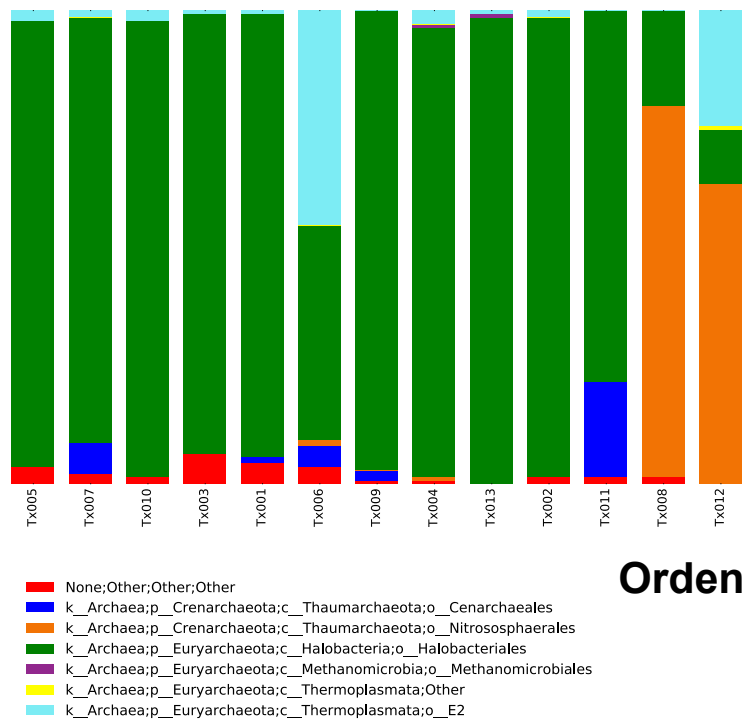


FIGURA C.1: **Curvas de rarefacción.** Número de secuencias por muestra contra (A) OTUs observados de Arquea, (B) índice Chao1 de Arquea, (C) OTUs observados de Bacteria y (D) índice Chao1 de Bacteria.



(A) Clase Arquea



(B) Orden Arquea

FIGURA C.2: **Resumen de Taxas.** Asignaciones de Arquea para (A) clase y (B) orden.

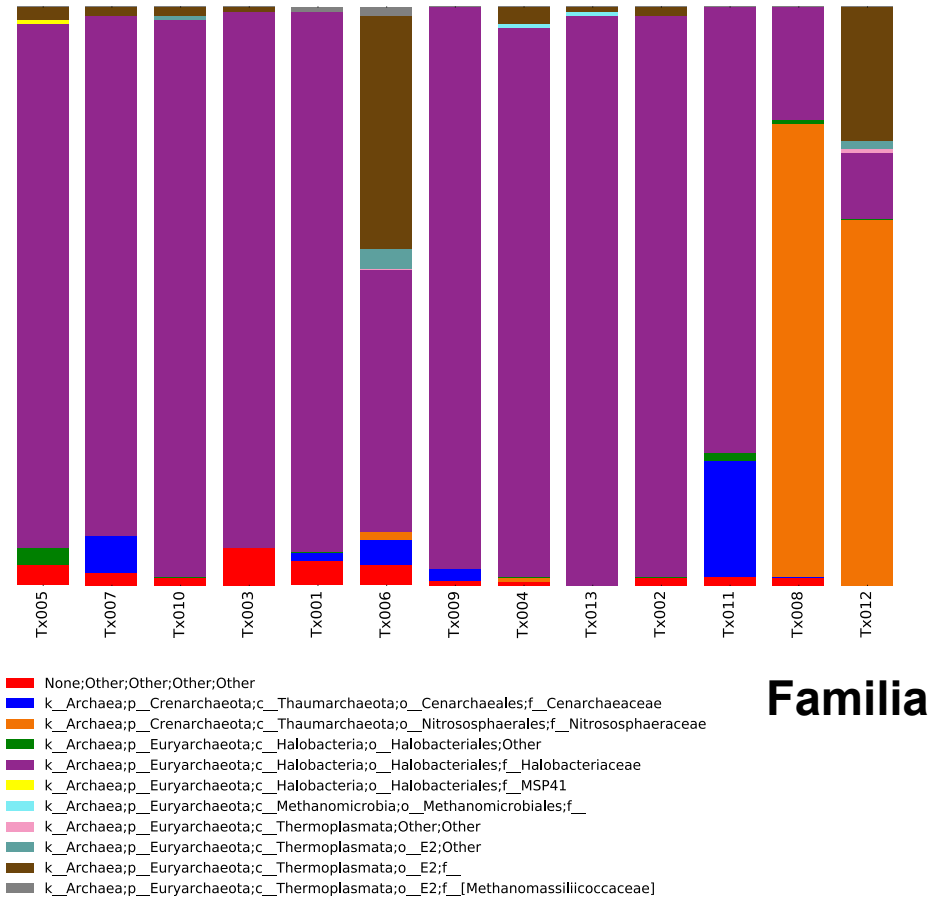
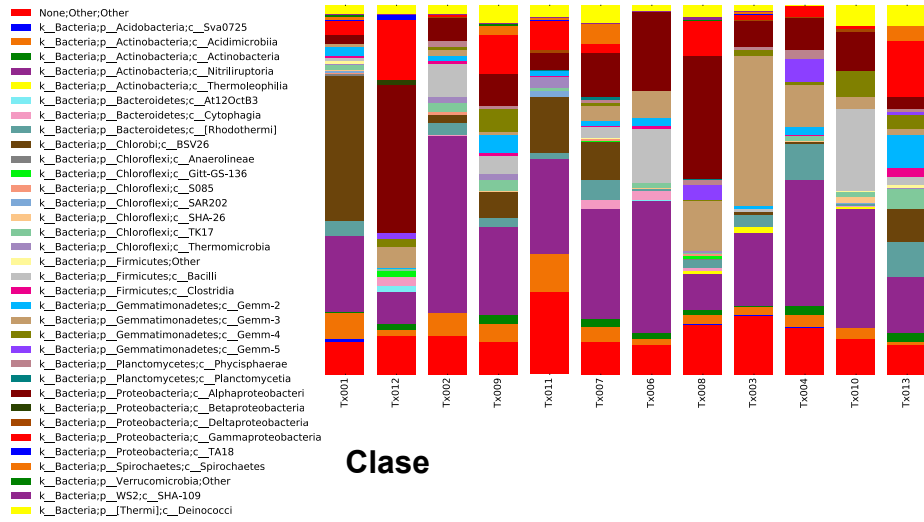
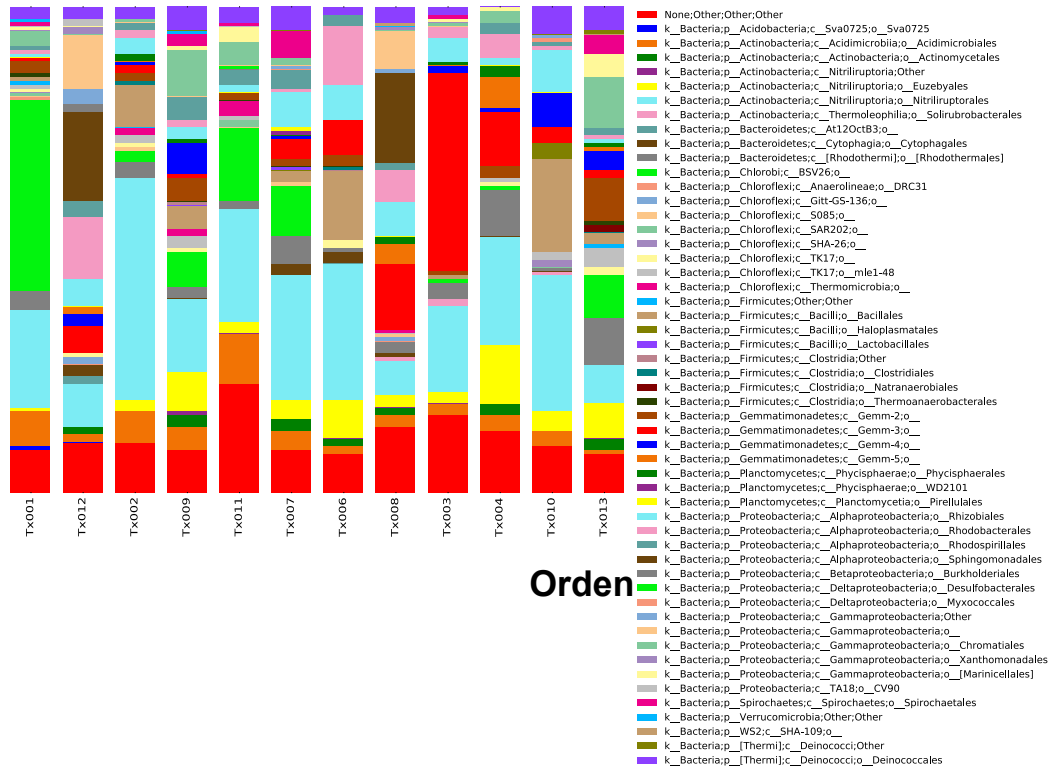


FIGURA C.3: **Resumen de Taxas.** Asignaciones de Arquea para familia.



(A) Clase Bacteria



(B) Orden Bacteria

FIGURA C.4: Resumen de Taxas. Asignaciones de Bacteria para (A) clase y (B) orden.

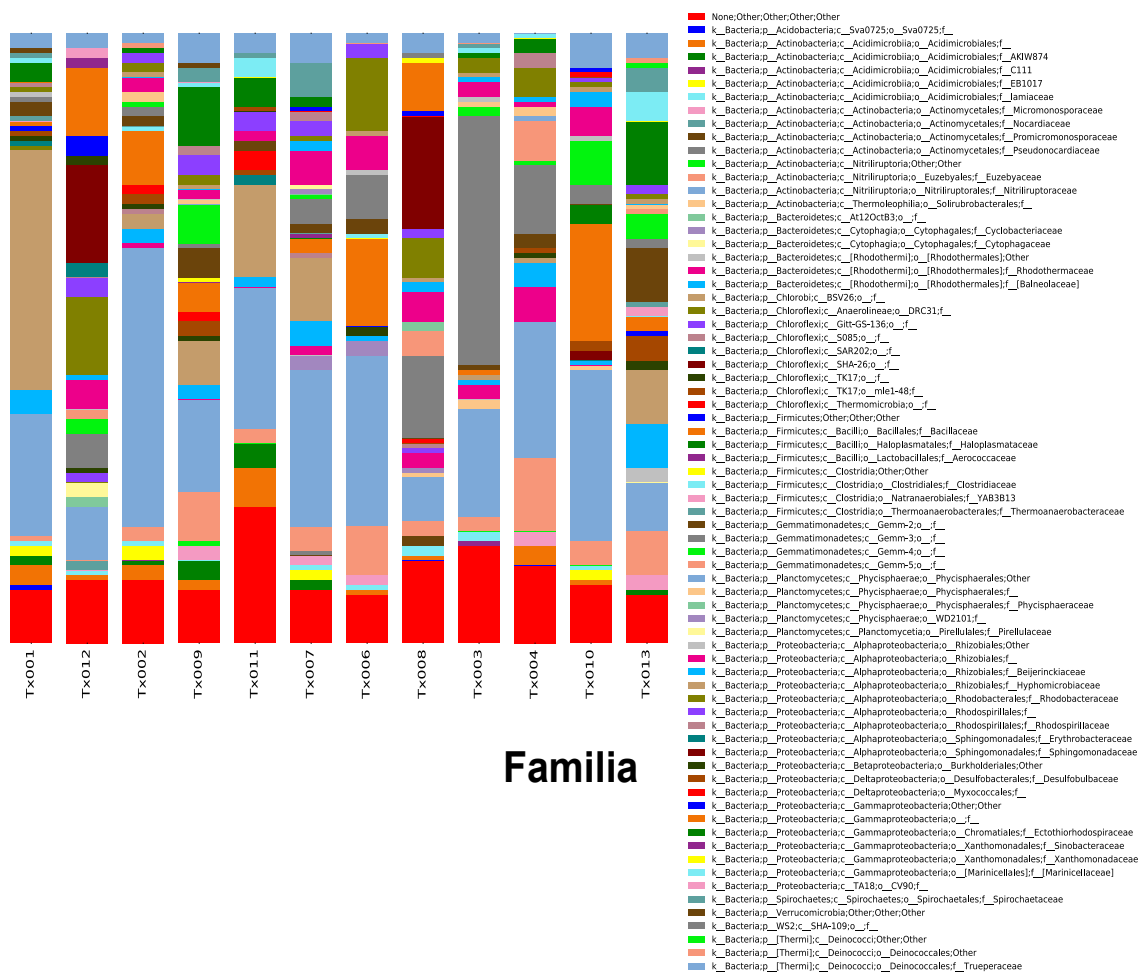
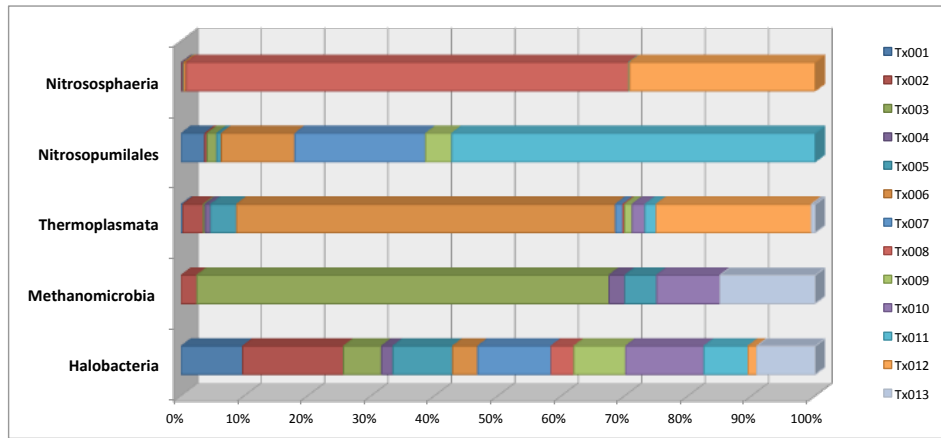
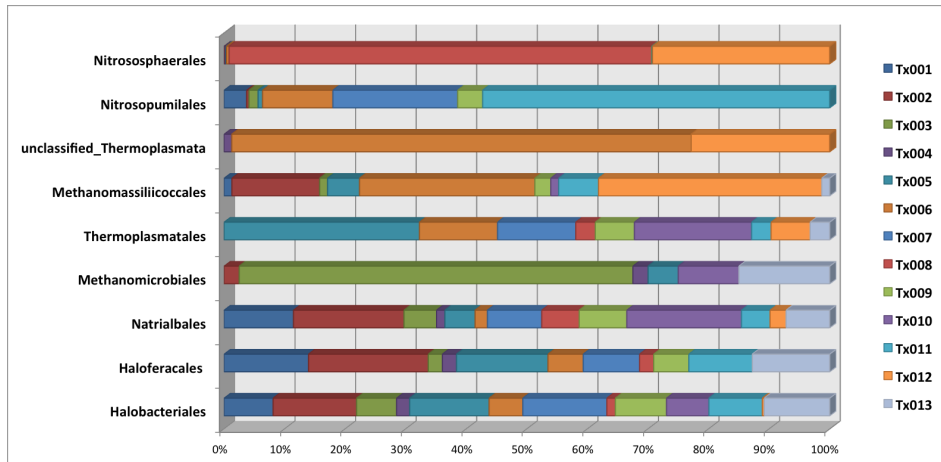


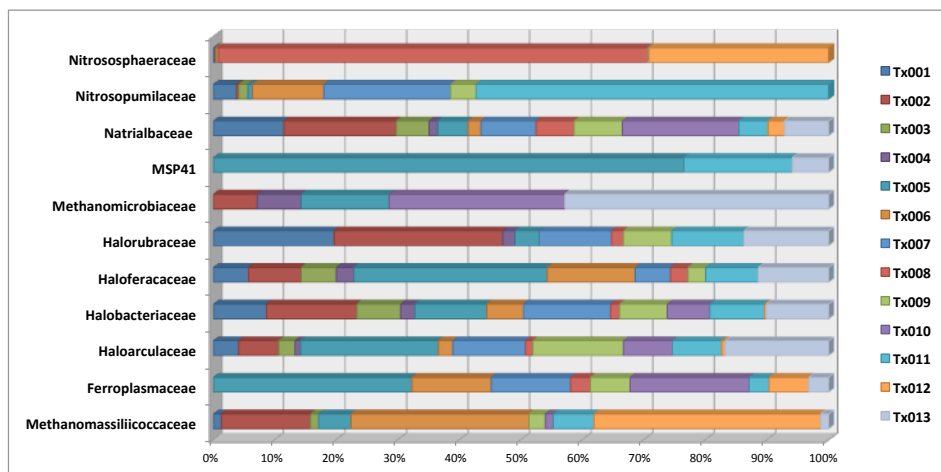
FIGURA C.5: Resumen de Taxas. Asignaciones de Bacteria para familia.



(A) Clase Arquea

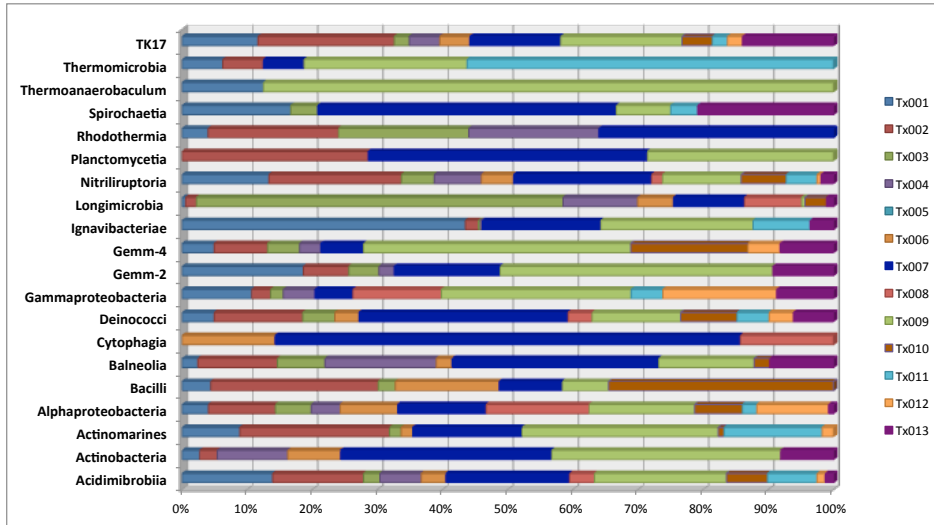


(B) Orden Arquea

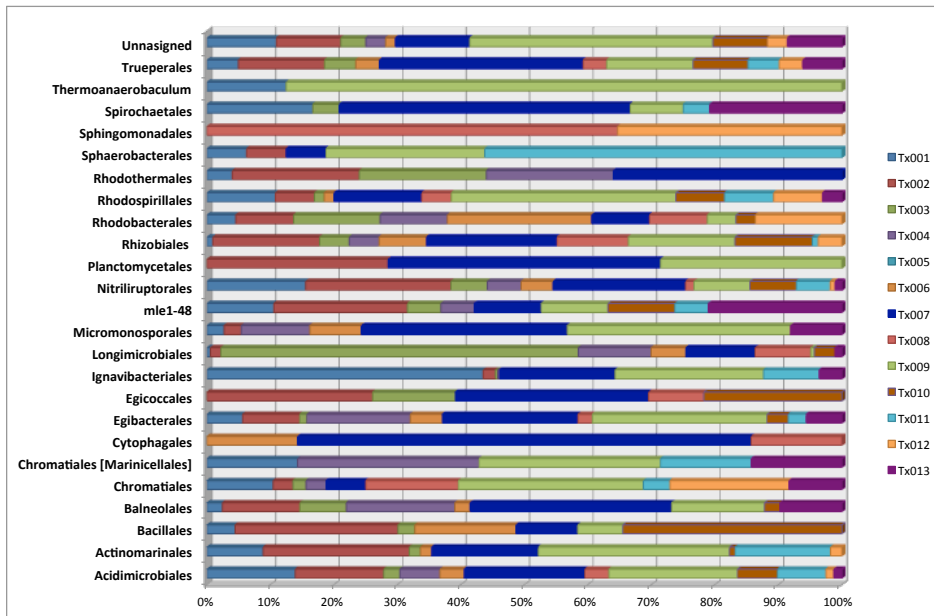


(C) Familia Arquea

FIGURA C.6: **Resumen de Taxas.** Asignaciones manuales de Arquea para (A) clase, (B) orden y (C) familia.



(A) Clase Bacteria



(B) Orden Bacteria

FIGURA C.7: Resumen de Taxas. Asignaciones manuales de Bacteria para (A) clase y (B) orden.



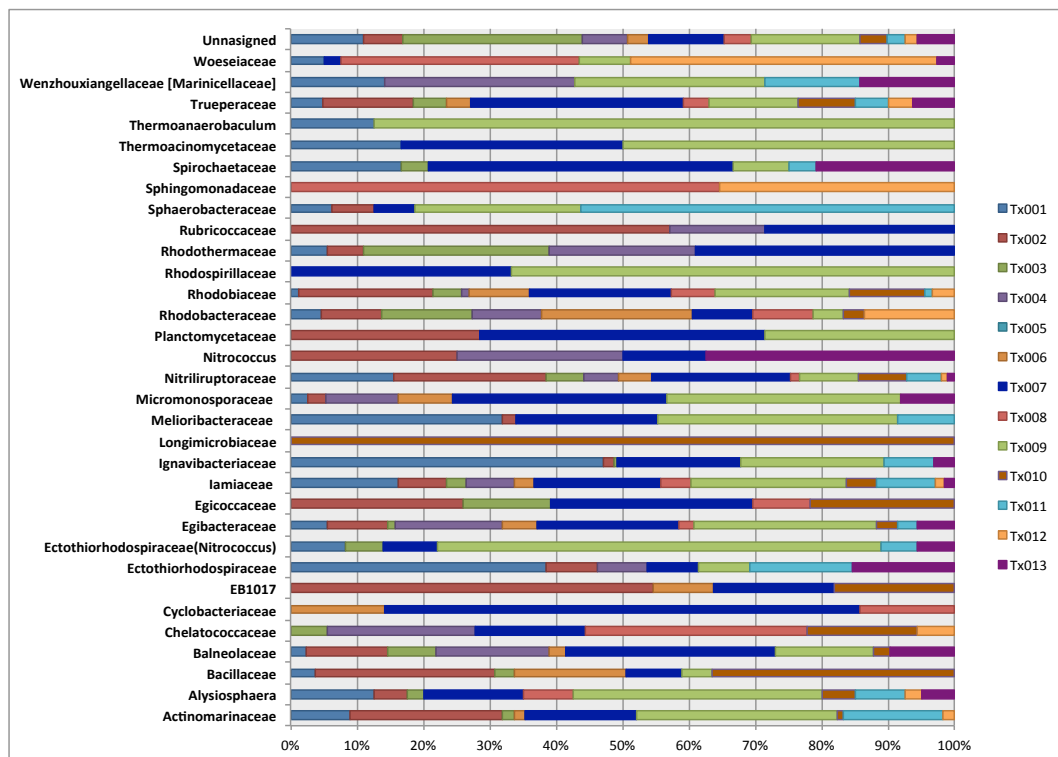


FIGURA C.8: Resumen de Taxas. Asignaciones de Bacteria para familia.

### **C.3. Asignación manual**

La asignación manual de secuencias, tanto para Arquea como para Bacteria, se realizó tomando en cuenta cada OTU y comparando contra diferentes bases de datos (Tablas C.1 y C.2).

TABLA C.1: Asignación manual de secuencias representativas para Arquea

OTU	No acceso	Phylum	Clase	Orden	Familia	Genero	Especie	% <sup>a</sup>	seq <sup>b</sup>	Similares
100780	NR_134097.1; FR773157	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	Nitrososphaeraceae	Nitrososphaera	<i>Nitrososphaera viennensis</i>	98	15	NCUR.OTU219; OTU324; OTU611; OTU1555; OTU1796; OTU1919; OTU1931; OTU473
235648	NR_113533.1	Euryarchaeota	Halobacteria	Natrialbales	Natrialbaeae	Natronococcus	<i>Natronococcus jeolgali</i>	98	13	242718; 252039; 249568; 239680; 242268; 237630; 257313; 238871; 251323; NCUR.OTU27; OTU47; OTU74; OTU87; OTU98; OTU105; OTU106; OTU138; OTU146; OTU245; OTU385; OTU386; OTU450; OTU473; OTU520; OTU526; OTU527; OTU552; OTU576; OTU625; OTU641; OTU644; OTU645; OTU655
236030	KR012975.1; JF449404	Euryarchaeota	Halobacteria	Halobacteriales	Halobacteriaceae	Halobacteriaceae archaeon		99	104	
236234	NR_118098.1; AF255608	Euryarchaeota	Thermoplasmata	Methanomassiliococcales	Methanomassiliococaceae			82	76	NCUR.OTU442; OTU668; OTU1643; OTU1703
240803	NR_113536.1; AB125108	Euryarchaeota	Halobacteria	Natrialbales	Natrialbaeae	Natronolimnobius		96	59	NCUR.OTU633; OTU1498; OTU1811
241952	KY486188.1	Euryarchaeota	Halobacteria	Halobacteriales	Halobacteriaceae	Halovarius		95	309	
242030	KF739010.1	Euryarchaeota	Halobacteria	Natrialbales	Natrialbaeae	Halobiforma		96	330	NCUR.OTU1920; OTU1992
242718	LT634701.1; AJ270233	Euryarchaeota	Halobacteria	Halobacteriales	Halobacteriaceae			91	2473	NCUR.OTU732





Tabla C.1 Continuación de página anterior.

OTU	No acceso	Phylum	Clase	Orden	Familia	Genero	Especie	% <sup>a</sup>	seq <sup>b</sup>	Similares
New.CleanUp. ReferenceOTU1473	KT165221	Euryarchaeota	Halobacteria	Halobacteriales	Halobacteriaceae	unclassified Halobacteriaceae		92	15	NCUR.OTU269; OTU373; OTU387; OTU653; OTU680; OTU735; OTU897; OTU1162; OTU1241; OTU1244; OTU1305; OTU1420; OTU1487
New.CleanUp. ReferenceOTU159	AB685273.1	Euryarchaeota	Halobacteria	Natrialbales	Natrialbaeae	Natronococcus	<i>Natronococcus xinjiangense</i>	98	125	
New.CleanUp. ReferenceOTU1757	FR832414	Euryarchaeota	Methanomicrobia	Methanomicrobiales				<50	26	NCUR.OTU903
New.CleanUp. ReferenceOTU1787	EU731296	Woesearchaeota						65	14	NCUR.OTU1767
New.CleanUp. ReferenceOTU18	LT634707.1	Euryarchaeota	Halobacteria	Halobacteriales	Haloarculaeeae			94	89	NCUR.OTU716; OTU972; OTU1161
New.CleanUp. ReferenceOTU1805	KJ644206.1; AB576125	Euryarchaeota	Halobacteria	Haloferacales	Haloferacaceae (Haloferacaceae en RDP)			90	51	
New.CleanUp. ReferenceOTU1834	LT634707.1; AV498645	Euryarchaeota	Halobacteria	Halobacteriales	Haloarculaeeae	Natronomonas	<i>Natronomonas moellapensis</i>	97	18	
New.CleanUp. ReferenceOTU202	KU663028.1	Euryarchaeota	Halobacteria	Natrialbales	Natrialbaeae	Halopiger		95	6	
New.CleanUp. ReferenceOTU23	NR_133034.1; KC430940; KT247978.1	Euryarchaeota	Halobacteria	Natrialbales	Natrialbaeae	unclassified Natrialbaeae		90	19	NCUR.OTU1521
New.CleanUp. ReferenceOTU258	NR_041513.1	Euryarchaeota	Thermoplasmata	unclassified Thermoplasmata				83	236	NCUR.OTU921
New.CleanUp. ReferenceOTU609	JF449405.1	Euryarchaeota	Halobacteria	Natrialbales	Natrialbaeae	Haloterrigena		96	22	
New.CleanUp. ReferenceOTU700	CF011070	Thaumarchaeota	Nitrososumiliales	Nitrososumiliales	Nitrososumilaceae			92	267	NCUR.OTU408; OTU1257
New.CleanUp. ReferenceOTU730	AJ862839.1	Euryarchaeota	Methanomicrobia	Methanomicrobiales	Methanomicrobiaceae			82	7	
New.CleanUp. ReferenceOTU848		Euryarchaeota	Methanomicrobia	Methanomicrobiales	Methanomicrobiaceae	Methanoculleus		96	7	

Tabla C.1 Continuación de página anterior.

OTU	No acceso	Phylum	Clase	Orden	Familia	Genero	Especie	% <sup>a</sup>	seq <sup>b</sup>	Similares
-----	-----------	--------	-------	-------	---------	--------	---------	----------------	------------------	-----------

<sup>a</sup> %T porcentaje de identidad, <sup>b</sup>Seq: total de secuencias, <sup>c</sup>NCUR: NewCleanUp.Reference

TABLA C.2: Asignación manual de secuencias representativas para Bacteria

OTU	No acceso	Phylum	Clase	Orden	Familia	Genero	Especie	%1 <sup>a</sup>	seq <sup>b</sup>	Similares
1142512	JN178434	Actinobacteria	Actinomarinnes	Actinomarinales	Actinomarinaceae			98	113	432982
4312325	JX240583	Actinobacteria	Nitriiruptoria	Nitriiruptorales	Nitriiruptoraceae	Nitriiruptor		98	572	113150
251360	KJ817496.1; KP307797; EU925618	Firmicutes	Bacilli	Bacillales	Bacillaceae	Salipaludibacillus		98	49	NewCleanUp. Reference OTU451
1132615	FJ152660; JN177984	Actinobacteria	Nitriiruptoria	Egibacterales	Egibacteraceae	Egibacter		99	105	NewCleanUp. Reference OTU1029
113787	KP174551.1; AY913401	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhodobiaceae	Parvibaculum		98	78	NewCleanUp. Reference OTU468
880753	KJ817461.1; HQ697793	Proteobacteria	Alphaproteobacteria	Rhizobiales	Chelatococcaceae			98	18	
745987	NR_043914.1	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Roseinatronobacter	<i>Roseinatronobacter monticus</i>	99	10	
NewCleanUp. Reference OTU7	NR_125596.1	Bacteroidetes	Cytophagia	Cytophagales	Cyclobacteriaceae	Anditalea		98	7	
NewCleanUp. Reference OTU9	EF150738.1; JF935173	Deinococcus- Thermus	Deinococci	Trueperales	Trueperaceae			90	81	NewCleanUp. Reference OTU196
NewCleanUp. Reference OTU14	FJ152623	Gemmatimonadetes	Longimicrobia	Longimicrobiales				97	177	NewCleanUp. Reference OTU229
NewCleanUp. Reference OTU37	HM438076; JQ427519	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Alysiosphaera	Alysiosphaera		99	40	NewCleanUp. Reference OTU168
NewCleanUp. Reference OTU64	JQ428680	Chlorobi	Ignavibacteriae	Ignavibacteriales	Ignavibacteriaceae			90	199	NewCleanUp. Reference OTU93
NewCleanUp. Reference OTU126	JN178582.1; NR_147752.1	Actinobacteria	Nitriiruptoria	Egibacterales	Egibacteraceae			94	73	NewCleanUp. Reference OTU324



Tabla C.2 continuación de página anterior

OTU	No acceso	Phylum	Clase	Orden	Familia	Genero	Especie	%I <sup>a</sup>	seq <sup>b</sup>	Similares
New.CleanUp. Reference OTU146	JQ428680; NR_074796.1; HM187401	Chlorobi	Ignavibacteriae	Ignavibacteriales	Meliobacteraceae			86	47	
New.CleanUp. Reference OTU885	NR_074796.1; HM187401	Chlorobi	Ignavibacteriae	Ignavibacteriales				84	55	New.CleanUp. Reference OTU1269
New.CleanUp. Reference OTU160	FJ152946.1; FJ152573	Actinobacteria	Nitrospiratia	Nitrospiratales	Nitrospiraceae			94	180	New.CleanUp. Reference OTU568
New.CleanUp. Reference OTU163	EU138973; AY673309.1; AY221045	Actinobacteria	Acidimicrobia	Acidimicrobiales	EB1017			95	11	
New.CleanUp. Reference OTU175	AY396017; JX240523.1; JN178059; FM864227	Firmicutes	Bacilli	Bacillales	Bacillaceae	Piscibacillus		95	7	
New.CleanUp. Reference OTU186	HM487963; HQ675422.1	Gemmatimonadetes	Gemm-2					96	43	New.CleanUp. Reference OTU499
New.CleanUp. Reference OTU231	FJ153017; JQ798462	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae			90	6	
New.CleanUp. Reference OTU235	JQ427679; FM209138	Actinobacteria	Acidimicrobia	Acidimicrobiales	lamiaceae			94	68	New.CleanUp. Reference OTU489
New.CleanUp. Reference OTU241	HQ397387	Proteobacteria	Gammaaproteobacteria	Chromatiales	Ectothiorhodospiraceae	Halofilum		97	13	
New.CleanUp. Reference OTU255	JQ426305	Gemmatimonadetes	Gemm-4					98	61	New.CleanUp. Reference OTU1043
New.CleanUp. Reference OTU313	FJ517010; KP174541.1	Proteobacteria	Gammaaproteobacteria	Chromatiales [Marinicellales]	Wenzhouxiangellaceae [Marinicellaceae]	Wenzhouxiangella		96	7	

Tabla C.2 continuación de página anterior

OTU	No acceso	Phylum	Clase	Orden	Familia	Genero	Especie	%1ª	seq <sup>b</sup>	Similares
New.CleanUp. Reference OTU340	FJ152969.1; FJ152901	Baincologiaota	Baincologia	Baincologiales	Baincologiales	Rhodohalobacter	Rhodohalobacter	98	21	New.CleanUp. Reference OTU922
New.CleanUp. Reference OTU351	JX240499; JQ428063	Baincologiaota	Baincologia	Baincologiales	Baincologiales			93	20	New.CleanUp. Reference OTU375
New.CleanUp. Reference OTU380	DQ463733.2	Chlorobi	Ignavibacteriae	Ignavibacteriales				84	7	
New.CleanUp. Reference OTU405	KY176966.1; JF236117	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingosinicella		96	6	
New.CleanUp. Reference OTU415	JN177967; NR_074379	Chloroflexi	Thermomicrobia	Sphaerobacterales	Sphaerobacteraceae			92	16	New.CleanUp. Reference OTU1388
New.CleanUp. Reference OTU502	NR_117452.1; KX989462.1; GU187912	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhodobiaceae			94	11	
New.CleanUp. Reference OTU506	EU423303.1; HM447693	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas		98	28	New.CleanUp. Reference OTU943
New.CleanUp. Reference OTU526	FJ748779	Proteobacteria	Gammaproteobacteria	Chromatiales	Woeseiaceae			92	39	New.CleanUp. Reference OTU1093
New.CleanUp. Reference OTU536	AM935790; HQ857646	Chloroflexi	TK17	mle1-48				95	19	New.CleanUp. Reference OTU831
New.CleanUp. Reference OTU547	KX785123.1; FJ876430; Z48306	Firmicutes	Bacilli	Bacillales	Bacillaceae			94	7	
New.CleanUp. Reference OTU564	EU006088.1; GQ464373; DQ789390	Proteobacteria	Gammaproteobacteria	Chromatiales	Nitroccoccus	Methylonatronum		94	8	

Tabla C.2 continuación de página anterior

OTU	No acceso	Phylum	Clase	Orden	Familia	Genero	Especie	%I <sup>a</sup>	seq <sup>b</sup>	Similares
New.CleanUp. Reference OTU597	CP020382.1; KF219109; FJ152912	Rhodotermaea	Rhodothermia	Rhodothermales	Rhodothermaeae			90	18	New.CleanUp. Reference OTU804
New.CleanUp. Reference OTU599	KJ535426.1; GQ264238; FJ153023	Chloroflexi	TK17					95	24	New.CleanUp. Reference OTU1450
New.CleanUp. Reference OTU616	KT1943989; EF208688	Planctomycetes	Planctomycetia	Planctomycetales	Planctomycetaceae			92	7	
New.CleanUp. Reference OTU701	KR108283.1; DQ480589; JQ428009	Rhodotermaea	Rhodothermia	Rhodothermales	Rubricoccaceae			90	7	
New.CleanUp. Reference OTU711	JX241025	Acidobacteria Gp10	Thermoanaerobactulum	Thermoanaerobactulum	Thermoanaerobactulum			91	8	
New.CleanUp. Reference OTU719	KY775646.1; JQ862432; ADGO01000419; FJ152655	Actinobacteria	Actinobacteria	Micromonosporales	Micromonosporaceae	Plantactinospora		95	37	New.CleanUp. Reference OTU726
New.CleanUp. Reference OTU768	KR605110.2	Actinobacteria	Nitriiruptoria	Egiccocales	Egiccocaceae	Egiccoccus		96	23	New.CleanUp. Reference OTU1250
New.CleanUp. Reference OTU860	KJ852652.1; LN610499; JN178396	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anaerobacillus		97	13	
New.CleanUp. Reference OTU895	KY775668.1; LN610499; AY032601	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anaerobacillus	<i>Anaerobacillus mucyae</i>	97	7	
New.CleanUp. Reference OTU933	NR_114942.1; AJ865469	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anaerobacillus	<i>Anaerobacillus arseniciselenatis</i>	99	6	
New.CleanUp. Reference OTU897	FM956479.1	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae			96	30	New.CleanUp. Reference OTU1233

Tabla C.2 continuación de página anterior

OTU	No acceso	Phylum	Clase	Orden	Familia	Genero	Especie	% <sup>a</sup>	seq <sup>b</sup>	Similares
New.CleanUp. Reference OTU926	FM956479; KF088606	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Albidovulum		99	17	
New.CleanUp. Reference OTU1069	KU720568.1; FJ152570; AF114783	Proteobacteria	Gammaproteobacteria	Chromatiales	Ectothiorhodospiraceae (Nitrococcus)			94	36	New.CleanUp. Reference OTU1451
New.CleanUp. Reference OTU1119	DQ157468.1	Firmicutes	Bacilli	Bacillales	Bacillaceae	Alkalibacillus	<i>Alkalibacillus filiformis</i>	99	12	
New.CleanUp. Reference OTU1497	HM132878; AJ493661	Firmicutes	Bacilli	Bacillales	Bacillaceae	Alkalibacillus		99	6	
New.CleanUp. Reference OTU1155	NR_028820.1; HG965770; JQ739013	Spirochaetes	Spirochaeta	Spirochaetales	Spirochaetaceae			92	24	New.CleanUp. Reference OTU1267
New.CleanUp. Reference OTU1336	AB042060	Firmicutes	Bacilli	Bacillales	Thermo- actinomycetaceae			86	6	
New.CleanUp. Reference OTU1353	KY489984.1; AJ987846; FJ152947; DQ432379	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Rhodovibrio		97	9	
New.CleanUp. Reference OTU1367	KY489984.1; JQ427519; HM438076.1	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillales			90	16	
New.CleanUp. Reference OTU1417	AF248638	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Rhodobaca	<i>Rhodobaca logovinis</i>	99	9	

<sup>a</sup> %I. Porcentaje de identidad, <sup>b</sup> Seq: total de secuencias

# Apéndice D

## Correlaciones y gráficos de ordinación

Gráficas realizadas como complemento multivariado a los análisis de correlación.

### D.1. PCA

El PCA se realizó para los análisis fisicoquímicos y los diferentes niveles taxonómicos de Arquea y Bacteria. Según el PCA de las características del suelo (Fig. D.1), Los sitios Tx009, Tx011 y Tx013 son agrupados en el cuadrante superior derecho y se caracterizan por gran cantidad de C y una alta EC y opuestos a los sitios Tx003, Tx006 y Tx008. Los sitios Tx008 y Tx012 se muestran distanciados de los demás sitios.

Las figuras D.2 y D.3 muestran los PCA para tres niveles taxonómicos de Arquea y Bacteria.

### D.2. Beta diversidad PCoA

Los gráficos basados en divergencia fueron obtenidos para Arquea y Bacteria (Fig. D.4)

Otros índices de Beta diversidad calculados fueron Bray-Curtis, Binary-Jaccard y Euclideana (Figs. D.5a y D.5b).



FIGURA D.1: **PCA de características fisicoquímicas.** Las Elipses agrupan la clase textural de cada sitio, PC1 explica el 37.8% de la variabilidad mientras PC2 explica el 26.5%.



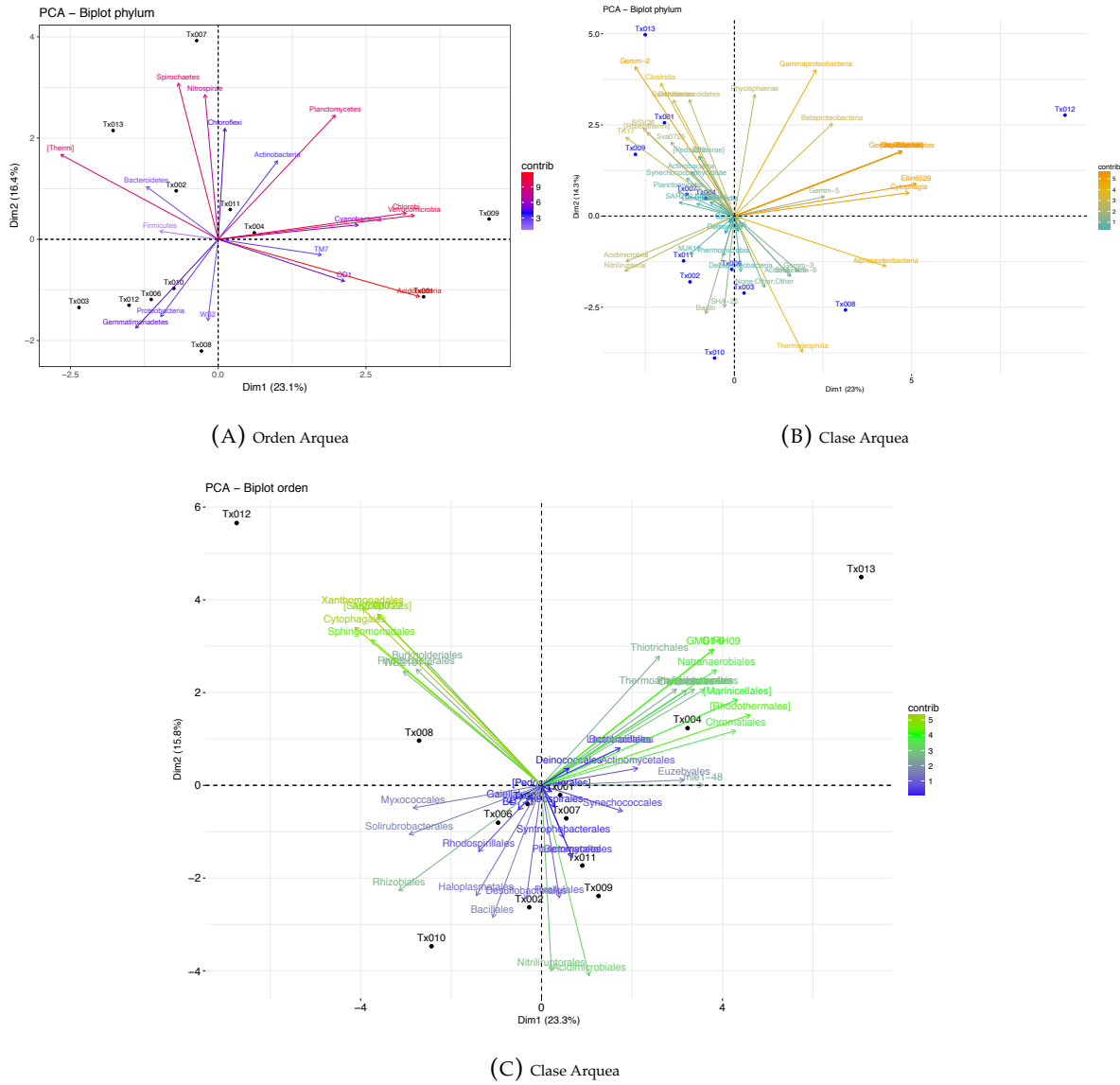


FIGURA D.3: Análisis de componentes principales (PCA) Bacteria. Análisis PCA para (A) phylum, (B) clase y (C) orden de Bacteria.



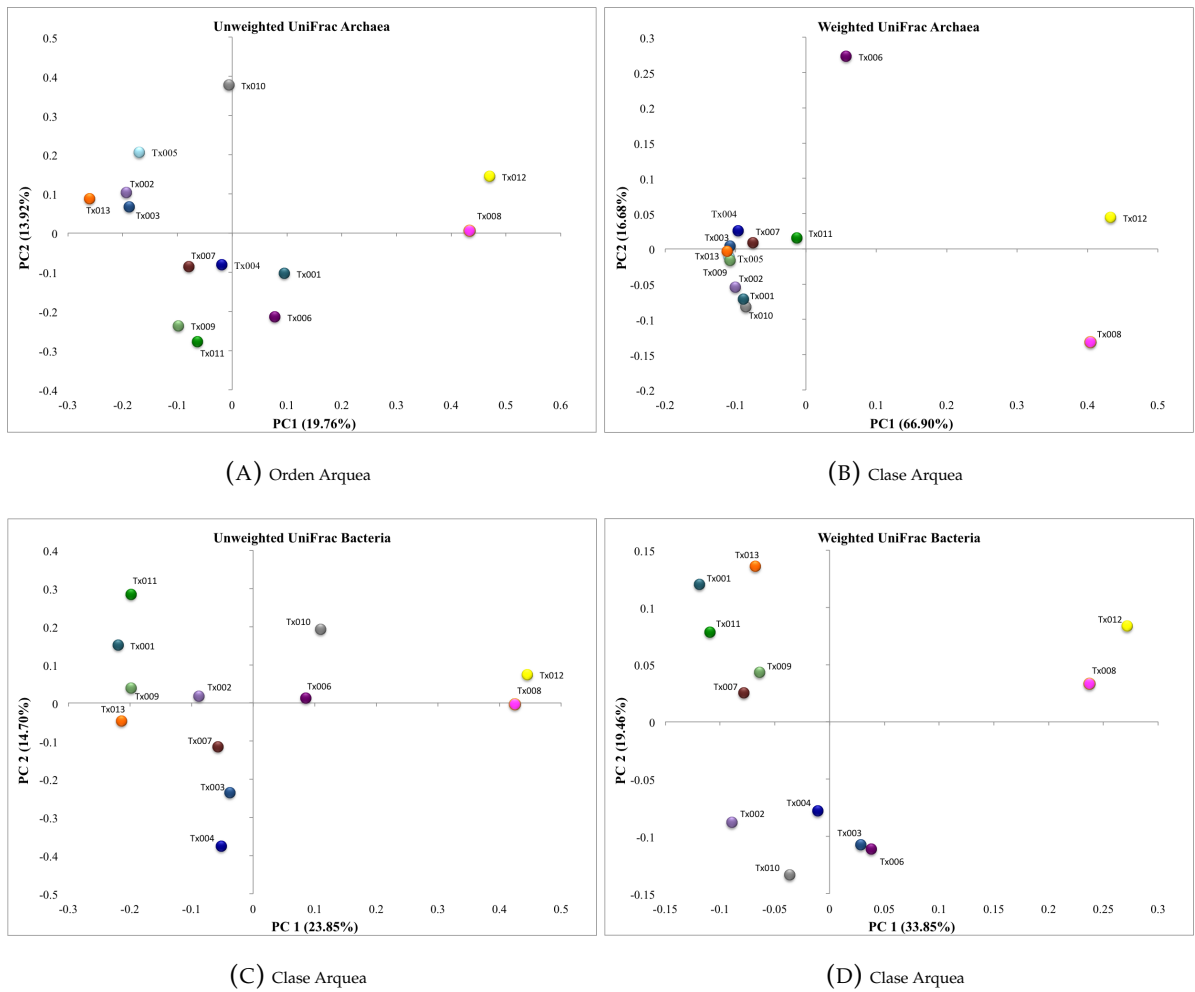


FIGURA D.4: Análisis de coordenadas principales (PCoA). (A) unweighted UniFrac Arquea, (B) weighted UniFrac Arquea, (C) unweighted UniFrac Bacteria y (D) weighted UniFrac Bacteria.

Bray-Curtis													
Sitio	Tx001	Tx002	Tx003	Tx004	Tx005	Tx006	Tx007	Tx008	Tx009	Tx010	Tx011	Tx012	Tx013
Tx001	0.00	0.53	0.88	0.86	0.85	0.95	0.69	0.89	0.69	0.65	0.88	0.89	0.88
Tx002	0.53	0.00	0.81	0.78	0.81	0.89	0.71	0.94	0.68	0.74	0.83	0.91	0.83
Tx003	0.88	0.81	0.00	0.61	0.84	0.78	0.71	0.94	0.76	0.91	0.76	0.96	0.80
Tx004	0.86	0.78	0.61	0.00	0.79	0.73	0.68	0.93	0.67	0.86	0.71	0.95	0.63
Tx005	0.89	0.81	0.84	0.79	0.00	0.89	0.78	0.98	0.80	0.87	0.83	0.97	0.70
Tx006	0.95	0.89	0.78	0.73	0.89	0.00	0.81	0.99	0.86	0.97	0.81	1.00	0.88
Tx007	0.69	0.71	0.71	0.68	0.78	0.81	0.00	0.94	0.71	0.75	0.71	0.92	0.85
Tx008	0.89	0.94	0.94	0.93	0.98	0.99	0.94	0.00	0.95	0.93	0.99	0.38	0.96
Tx009	0.69	0.68	0.76	0.67	0.80	0.86	0.71	0.95	0.00	0.78	0.70	0.94	0.72
Tx010	0.65	0.74	0.91	0.86	0.87	0.97	0.75	0.93	0.78	0.00	0.93	0.90	0.93
Tx011	0.88	0.83	0.76	0.71	0.83	0.81	0.71	0.99	0.70	0.93	0.00	0.96	0.78
Tx012	0.89	0.91	0.96	0.95	0.97	1.00	0.92	0.38	0.94	0.90	0.96	0.00	0.96
Tx013	0.88	0.83	0.80	0.63	0.70	0.88	0.85	0.96	0.72	0.93	0.78	0.96	0.00

Bin-Jaccard													
Sitio	Tx001	Tx002	Tx003	Tx004	Tx005	Tx006	Tx007	Tx008	Tx009	Tx010	Tx011	Tx012	Tx013
Tx001	0.00	0.78	0.89	0.86	0.91	0.95	0.83	0.89	0.84	0.84	0.93	0.94	0.92
Tx002	0.78	0.00	0.83	0.84	0.83	0.92	0.84	0.95	0.86	0.88	0.90	0.97	0.90
Tx003	0.89	0.83	0.00	0.81	0.88	0.87	0.85	0.94	0.85	0.92	0.87	0.96	0.86
Tx004	0.86	0.84	0.81	0.00	0.85	0.85	0.82	0.92	0.79	0.88	0.85	0.96	0.78
Tx005	0.91	0.83	0.88	0.85	0.00	0.90	0.83	0.98	0.88	0.88	0.89	0.98	0.81
Tx006	0.95	0.92	0.87	0.85	0.90	0.00	0.90	0.98	0.91	0.96	0.89	1.00	0.91
Tx007	0.83	0.84	0.85	0.82	0.83	0.90	0.00	0.94	0.88	0.83	0.87	0.97	0.90
Tx008	0.89	0.95	0.94	0.92	0.98	0.98	0.94	0.00	0.93	0.94	0.99	0.86	0.96
Tx009	0.84	0.86	0.85	0.79	0.88	0.91	0.88	0.93	0.00	0.87	0.82	0.97	0.81
Tx010	0.84	0.88	0.92	0.88	0.88	0.96	0.83	0.94	0.87	0.00	0.96	0.95	0.93
Tx011	0.93	0.90	0.87	0.85	0.89	0.89	0.87	0.99	0.82	0.96	0.00	0.99	0.87
Tx012	0.94	0.97	0.96	0.96	0.98	1.00	0.97	0.86	0.97	0.95	0.99	0.00	0.97
Tx013	0.92	0.90	0.86	0.78	0.81	0.91	0.90	0.96	0.81	0.93	0.87	0.97	0.00

Euclidean													
Sitio	Tx001	Tx002	Tx003	Tx004	Tx005	Tx006	Tx007	Tx008	Tx009	Tx010	Tx011	Tx012	Tx013
Tx001	0.00	19.80	43.17	34.93	42.94	50.64	24.98	103.04	25.46	36.33	40.07	79.99	36.85
Tx002	19.80	0.00	42.99	35.27	43.86	51.63	27.64	104.10	23.66	41.71	42.54	81.33	38.05
Tx003	43.17	42.99	0.00	24.54	44.25	46.26	28.32	105.03	34.26	51.19	40.62	83.53	36.85
Tx004	34.93	35.27	24.54	0.00	37.95	40.17	23.02	102.19	24.54	45.72	34.03	79.92	24.25
Tx005	42.94	43.86	44.25	37.95	0.00	51.85	37.23	105.01	36.72	53.01	44.61	83.34	27.78
Tx006	50.64	51.63	46.26	40.17	51.85	0.00	42.73	107.69	45.89	57.98	48.15	86.89	46.54
Tx007	24.98	27.64	28.32	23.02	37.23	42.73	0.00	101.65	21.35	40.30	26.80	78.80	29.80
Tx008	103.04	104.10	105.03	102.19	105.01	107.69	101.65	0.00	102.14	106.40	104.85	31.81	102.67
Tx009	25.46	23.66	34.26	24.54	36.72	45.89	21.35	102.14	0.00	41.57	33.70	79.46	28.74
Tx010	36.33	41.71	51.19	45.72	51.01	57.98	40.30	106.40	41.57	0.00	51.09	84.49	47.52
Tx011	40.07	42.54	40.62	34.03	44.61	48.15	26.80	104.85	33.70	51.09	0.00	82.96	36.52
Tx012	79.99	81.33	83.53	79.92	83.34	86.89	78.80	31.81	79.46	84.49	82.96	0.00	80.52
Tx013	36.85	38.05	36.85	24.25	27.78	46.54	29.80	102.67	28.74	47.52	36.52	80.52	0.00

(A) Beta Diversidad Arquea

Bray-Curtis												
Sitio	Tx001	Tx002	Tx003	Tx004	Tx006	Tx007	Tx008	Tx009	Tx010	Tx011	Tx012	Tx013
Tx001	0.00	0.74	0.86	0.90	0.86	0.80	0.98	0.76	0.93	0.70	0.97	0.82
Tx002	0.74	0.00	0.85	0.83	0.76	0.77	0.95	0.78	0.77	0.82	0.91	0.78
Tx003	0.86	0.85	0.00	0.82	0.82	0.79	0.86	0.88	0.90	0.91	0.93	0.90
Tx004	0.90	0.83	0.82	0.00	0.86	0.82	0.95	0.81	0.86	0.95	0.94	0.86
Tx006	0.86	0.76	0.82	0.86	0.00	0.86	0.90	0.88	0.89	0.92	0.96	0.92
Tx007	0.80	0.77	0.79	0.82	0.86	0.00	0.94	0.73	0.86	0.74	0.95	0.74
Tx008	0.98	0.95	0.86	0.95	0.90	0.94	0.00	0.95	0.92	0.98	0.70	0.98
Tx009	0.76	0.78	0.88	0.81	0.80	0.73	0.95	0.00	0.82	0.77	0.92	0.74
Tx010	0.93	0.77	0.90	0.86	0.89	0.86	0.92	0.82	0.00	0.95	0.93	0.89
Tx011	0.70	0.82	0.91	0.95	0.92	0.74	0.98	0.77	0.95	0.00	0.97	0.80
Tx012	0.97	0.91	0.93	0.94	0.96	0.95	0.70	0.92	0.93	0.97	0.00	0.95
Tx013	0.82	0.78	0.90	0.86	0.92	0.74	0.98	0.74	0.89	0.80	0.95	0.00

Binary-Jaccard												
Sitio	Tx001	Tx002	Tx003	Tx004	Tx006	Tx007	Tx008	Tx009	Tx010	Tx011	Tx012	Tx013
Tx001	0.00	0.85	0.87	0.91	0.91	0.89	0.98	0.81	0.93	0.76	0.97	0.87
Tx002	0.85	0.00	0.87	0.90	0.84	0.81	0.96	0.84	0.84	0.87	0.92	0.80
Tx003	0.87	0.87	0.00	0.87	0.86	0.86	0.92	0.91	0.92	0.94	0.93	0.93
Tx004	0.91	0.90	0.87	0.00	0.91	0.87	0.96	0.88	0.91	0.96	0.95	0.89
Tx006	0.91	0.84	0.86	0.91	0.00	0.92	0.93	0.90	0.92	0.94	0.96	0.93
Tx007	0.89	0.81	0.86	0.87	0.92	0.00	0.96	0.82	0.90	0.84	0.96	0.83
Tx008	0.98	0.96	0.92	0.96	0.93	0.96	0.00	0.96	0.94	0.99	0.75	0.98
Tx009	0.81	0.84	0.91	0.88	0.90	0.82	0.96	0.00	0.88	0.81	0.95	0.82
Tx010	0.93	0.84	0.92	0.91	0.92	0.90	0.94	0.88	0.00	0.95	0.94	0.92
Tx011	0.76	0.87	0.94	0.96	0.94	0.84	0.99	0.81	0.95	0.00	0.98	0.86
Tx012	0.97	0.92	0.93	0.95	0.96	0.96	0.75	0.95	0.94	0.98	0.00	0.96
Tx013	0.87	0.80	0.93	0.89	0.93	0.83	0.98	0.82	0.92	0.86	0.96	0.00

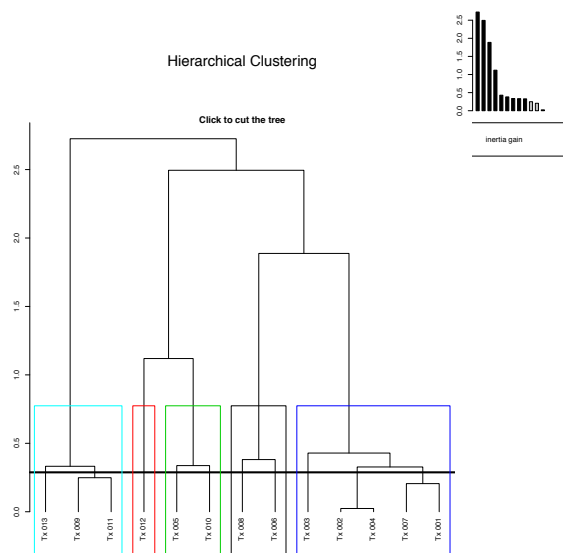
Euclidean												
Sitio	Tx001	Tx002	Tx003	Tx004	Tx006	Tx007	Tx008	Tx009	Tx010	Tx011	Tx012	Tx013
Tx001	0.00	25.06	30.10	28.18	28.46	23.87	31.14	26.00	32.47	25.42	31.84	24.37
Tx002	25.06	0.00	25.42	22.54	21.68	21.54	27.39	21.49	25.34	24.08	27.64	22.49
Tx003	30.10	25.42	0.00	23.71	24.17	21.12	25.73	24.82	29.39	26.46	28.64	23.54
Tx004	28.18	22.54	23.71	0.00	23.02	20.54	25.61	19.85	26.04	25.10	26.42	21.12
Tx006	28.46	21.68	24.17	23.02	0.00	22.63	26.23	23.62	28.00	25.65	28.04	23.75
Tx007	23.87	21.54	21.12	20.54	22.63	0.00	24.66	18.87	26.19	19.85	26.00	18.44
Tx008	31.14	27.39	25.73	25.61	26.23	24.66	0.00	25.50	29.50	27.31	22.05	25.42
Tx009	26.00	21.49	24.82	19.85	23.62	18.87	25.50	0.00	25.50	21.77	25.69	18.87
Tx010	32.47	25.34	29.39	26.04	28.00	26.19	29.50	25.50	0.00	29.56	30.40	26.31
Tx011	25.42	24.08	26.46	25.10	25.65	19.85	27.31	21.77	29.56	0.00	28.21	22.41
Tx012	31.84	27.64	28.64	26.42	28.04	26.00	22.05	25.69	30.40	28.21	0.00	26.04
Tx013	24.37	22.49	23.54	21.12	23.75	18.44	25.42	18.87	26.31	22.41	26.04	0.00

(B) Beta Diversidad Bacteria

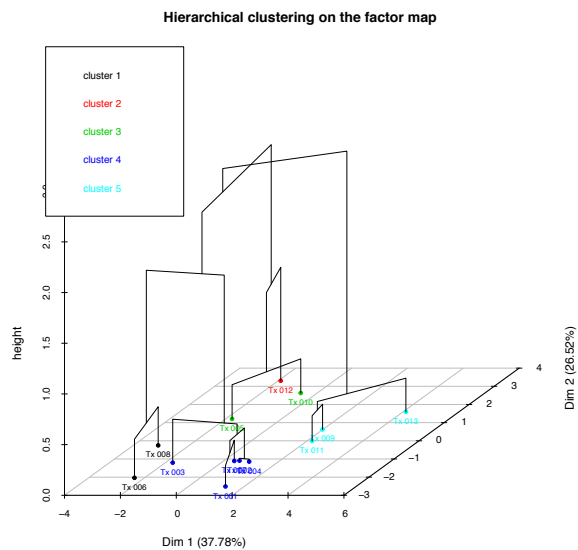
FIGURA D.5: Análisis de Beta Diversidad (A) Arquea y (B) Bacteria.

## **D.3. Clustering**

Se realizaron análisis de cluster con el fin de agrupar los sitios según distancias euclidianas de acuerdo a sus características fisicoquímicas.



(A) Hclust de Factores



(B) Clase Arquea

FIGURA D.6: **Análisis de cluster usando un árbol jerárquico por distancias euclidianas.** Para los sitios de acuerdo a sus características físico-químicas (A) árbol que muestra el cluster jerárquico y el conglomerado junto con la ganancia de inercia y (B) mapa de individuos (sitios) incluyendo el árbol. Los clúster se muestran de diferente color.

# Bibliografía

- Amato, M. (1983). «Determination of Carbon 12C and 14C in Plant and Soil». En: *Soil Biology & Biochemistry* 15.5, págs. 611-612.
- Andrew, D.R., Fitak, R.R., Munguia-Vega, A., Racolta, A., Martinson, V.G. y Dontsova, K. (2012). «Abiotic factors shape microbial diversity in Sonoran Desert soils». En: *Applied and Environmental Microbiology* 78.21, págs. 7527-37. ISSN: 1098-5336. DOI: [10.1128/AEM.01459-12](https://doi.org/10.1128/AEM.01459-12). URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3485727&tool=pmcentrez&rendertype=abstract>.
- Auguet, J.C. y Casamayor, E.O. (2013). «Partitioning of Thaumarchaeota populations along environmental gradients in high mountain lakes». En: *FEMS microbiology ecology* 84.October, págs. 154-164. DOI: [10.1111/1574-6941.12047](https://doi.org/10.1111/1574-6941.12047).
- Barbosa, A.M. (2016). *fuzzySim: Fuzzy similarity in species distributions*. R package version 1.7.8/r95. URL: <https://R-Forge.R-project.org/projects/fuzzysim/>.
- Bissett, A., Richardson, A. E., Baker, G., Wakelin, S. y Thrall, P. H. (2010). «Life history determines biogeographical patterns of soil bacterial communities over multiple spatial scales». En: *Molecular Ecology* 19, págs. 4315-4327. DOI: [10.1111/j.1365-294X.2010.04804.x](https://doi.org/10.1111/j.1365-294X.2010.04804.x).

- Bivand, R.S., Pebesma, E. y Gómez-Rubio, V. (2013a). *Applied Spatial Data Analysis with R*. Ed. por Robert Gentleman, Kurt Hornik y Giovanni G Parmigiani. Segunda Ed. New York, USA: Springer Science+Business Media, pág. 413. ISBN: 9781461476177. DOI: [10.1007/978-1-4614-7618-4](https://doi.org/10.1007/978-1-4614-7618-4).
- (2013b). *Applied spatial data analysis with R, Second edition*. Springer, NY. URL: <http://www.asdar-book.org/>.
- Bjørnstad, O.N. (2016). *ncf: Spatial Nonparametric Covariance Functions*. R package version 1.1-7. URL: <https://CRAN.R-project.org/package=ncf>.
- Bollmann, A., Bullerjahn, G.S. y McKay, R.M. (2014). «Abundance and Diversity of Ammonia-Oxidizing Archaea and Bacteria in Sediments of Trophic End Members of the Laurentian Great Lakes, Erie and Superior». En: *PLOS one* 9.5, págs. 1-11. DOI: [10.1371/journal.pone.0097068](https://doi.org/10.1371/journal.pone.0097068).
- Bru, D., Ramette, A., Saby, N.P.A., Dequiedt, S., Ranjard, L., Jolivet, C., Arrouays, D. y Philippot, L. (2011). «Determinants of the distribution of nitrogen-cycling microbial communities at the landscape scale». En: *The ISME journal* 5.3, págs. 532-542. ISSN: 1751-7370. DOI: [10.1038/ismej.2010.130](https://doi.org/10.1038/ismej.2010.130). URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3105713&tool=pmcentrez&rendertype=abstract>.
- Budakoglu, M., Kurt, H., Karaman, M., Kumru, M., Kumral, M. y Akarsubaşı, A.T. (2014). «Archaeal microbial diversity of hypersaline Lake Acıgöl, Denizli, Turkey». En: *Geomicrobiology Journal* 31.6, págs. 454-460. ISSN: 0149-0451. DOI: [10.1080/01490451.2013.866994](https://doi.org/10.1080/01490451.2013.866994). URL: <http://www.tandfonline.com/doi/abs/10.1080/01490451.2013.866994>.
- Bundy, L.G. y Bremner, J.M. (1972). «A Simple Titrimetric Method for Determination of Inorganic Carbon in Soils<sup>1</sup>». En: *Soil Science Society of America Journal* 36, págs. 273-275. DOI: [10.2136/sssaj1972.03615995003600020021x](https://doi.org/10.2136/sssaj1972.03615995003600020021x).

- Cadotte, M.W., Jonathan Davies, T., Regetz, J., Kembel, S.W., Cleland, E. y Oakley, T.H. (2010). «Phylogenetic diversity metrics for ecological communities: Integrating species richness, abundance and evolutionary history». En: *Ecology Letters* 13.1, págs. 96-105. ISSN: 1461023X. DOI: [10.1111/j.1461-0248.2009.01405.x](https://doi.org/10.1111/j.1461-0248.2009.01405.x).
- Callahan, B. J., Mccurdie, P. J. y Holmes, S. P. (2017). «Exact sequence variants should replace operational taxonomic units in marker-gene data analysis». En: *ISME* 11.12, págs. 2639-2643. ISSN: 1751-7362. DOI: [10.1038/ismej.2017.119](https://doi.org/10.1038/ismej.2017.119). URL: <http://dx.doi.org/10.1038/ismej.2017.119>.
- Cambardella, C. A., Moorman, T. B., Novak, J., Parkin, T. B., Karlen, R. F., Turco, R. F. y Konopka, A. E. (1994). «Field-Scale variability of soil properties in central Iowa soils». En: *Soil Science Society of America Journal* 58, págs. 1501-1511. DOI: [10.2136/sssaj1994.03615995005800050033x](https://doi.org/10.2136/sssaj1994.03615995005800050033x).
- Canfora, L., Bacci, G., Pinzari, F., Lo Papa, G., Dazzi, C. y Benedetti, A. (2014). «Salinity and bacterial diversity: To what extent does the concentration of salt affect the bacterial community in a saline soil?» En: *PLOS one* 9.9, págs. 1-15. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0106662](https://doi.org/10.1371/journal.pone.0106662). URL: <http://dx.plos.org/10.1371/journal.pone.0106662>.
- Canfora, L., Lo Papa, G., Vittori Antisari, L., Bazan, G., Dazzi, C. y Benedetti, A. (2015). «Spatial microbial community structure and biodiversity analysis in “extreme” hypersaline soils of a semiarid Mediterranean area». En: *Applied Soil Ecology* 93, págs. 120-129. ISSN: 09291393. DOI: [10.1016/j.apsoil.2015.04.014](https://doi.org/10.1016/j.apsoil.2015.04.014). URL: <http://linkinghub.elsevier.com/retrieve/pii/S092913931500116>.
- Caporaso, J. G., Bittinger, K., Bushman, F. D., Desantis, T. Z., Andersen, G. L. y Knight, R. (2010a). «PyNAST: A flexible tool for aligning sequences to a template alignment». En: *Bioinformatics* 26.2, págs. 266-267. ISSN: 14602059. DOI: [10.1093/bioinformatics/btp636](https://doi.org/10.1093/bioinformatics/btp636).

- Caporaso, J. G. et al. (2010b). «QIIME allows analysis of high-throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing». En: *Nature Methods* 7.5, págs. 335-336. ISSN: 1548-7091. DOI: [10.1038/nmeth0510-335](https://doi.org/10.1038/nmeth0510-335). URL: <http://dx.doi.org/10.1038/nmeth0510-335>.
- Ceja-Navarro, J.A., Rivera, F. N., Patiño-Zúñiga, L., Govaerts, B., Marsch, R., Vila-Sanjurjo, A. y Dendooven, L. (2010). «Molecular characterization of soil bacterial communities in contrasting zero tillage systems». En: *Plant and Soil* 329.1-2, págs. 127-137. ISSN: 0032-079X. DOI: [10.1007/s11104-009-0140-9](https://doi.org/10.1007/s11104-009-0140-9). URL: <http://link.springer.com/10.1007/s11104-009-0140-9>.
- Chabala, L.M., Mulolwa, A. y Lungu, O. (2017). «Application of ordinary kriging in mapping soil organic carbon in Zambia». En: *Pedosphere* 27.2, págs. 338-343. ISSN: 1002-0160. DOI: [10.1016/S1002-0160\(17\)60321-7](https://doi.org/10.1016/S1002-0160(17)60321-7). URL: [http://dx.doi.org/10.1016/S1002-0160\(17\)60321-7](http://dx.doi.org/10.1016/S1002-0160(17)60321-7).
- Chen, W., Zhang, C. K., Cheng, Y., Zhang, S. y Zhao, H. (2013). «A Comparison of Methods for Clustering 16S rRNA Sequences into OTUs». En: *PloS one* 8.8, págs. 1-8. DOI: [10.1371/journal.pone.0070837](https://doi.org/10.1371/journal.pone.0070837).
- Chen, Y., Dai, Y., Wang, Y., Wu, Z., Xie, S. y Liu, Y. (2015). «Distribution of bacterial communities across plateau freshwater lake and upslope soils». En: *Journal of Environmental Sciences* 43, págs. 61-69. ISSN: 1001-0742. DOI: [10.1016/j.jes.2015.08.012](https://doi.org/10.1016/j.jes.2015.08.012). URL: <http://dx.doi.org/10.1016/j.jes.2015.08.012>.
- Constancias, F., Dequiedt, S., Horrigue, W., Saby, N. P. A. y Ranjard, L. (2015). «Contrasting spatial patterns and ecological attributes of soil bacterial and archaeal taxa across a landscape». En: *Microbiology open* 4.3, págs. 518-531. DOI: [10.1002/mbo3.256](https://doi.org/10.1002/mbo3.256).
- Crits-Christoph, A., Robinson, C. K., Barnum, T., Fricke, W. F., Davila, A. F., Jedynak, B., McKay, C. P. y Diruggiero, J. (2013). «Colonization patterns of soil microbial



- communities in the Atacama Desert.» En: *Microbiome* 1.1, pág. 28. ISSN: 2049-2618. DOI: [10.1186/2049-2618-1-28](https://doi.org/10.1186/2049-2618-1-28). URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3971613&tool=pmcentrez&rendertype=abstract>.
- Dendooven, L., Alcántara-Hernández, R. J., Valenzuela-Encinas, C., Luna-Guido, M., Perez-Guevara, F. y Marsch, R. (2010). «Dynamics of carbon and nitrogen in an extreme alkaline saline soil: A review». En: *Soil Biology & Biochemistry* 42.6, págs. 865-877. ISSN: 00380717. DOI: [10.1016/j.soilbio.2010.02.014](https://doi.org/10.1016/j.soilbio.2010.02.014). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0038071710000672>.
- Edgar, R. C. (2010). «Search and clustering orders of magnitude faster than BLAST». En: *Bioinformatics* 26.19, págs. 2460-2461. DOI: [10.1093/bioinformatics/btq461](https://doi.org/10.1093/bioinformatics/btq461).
- Ettema, C. H. y Wardle, D. A. (2002). «Spatial soil ecology». En: *Trends in Ecology & Evolution* 17.4, págs. 177-183.
- Franklin, R. B. y Mills, A. L. (2003). «Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field». En: *FEMS Microbiology Ecology* 44.3, págs. 335-46. ISSN: 0168-6496. DOI: [10.1016/S0168-6496\(03\)00074-6](https://doi.org/10.1016/S0168-6496(03)00074-6). URL: <http://www.ncbi.nlm.nih.gov/pubmed/12830827>.
- Gallardo, A. y Maestre, F. T. (2008). «Métodos geoestadísticos para el análisis de datos ecológicos espacialmente explícitos». En: *Introducción al análisis espacial de datos en ecología y ciencias ambientales: Métodos y aplicaciones*. Ed. por Fernando T Maestre, Adrián Escudero-Alcántara y Andreu Bonet-Jornet. Sevilla, España: Universidad Rey Juan Carlos, Servicio de Publicaciones. Cap. 6, págs. 215-272. ISBN: 978-84-9849-308-5.
- Google Earth Pro. *Zona Federal del ex-lago de Texcoco*. Versión 7.3.1.4507. Digital Globe 2018, Google 2018, INEGI 2018. URL: <http://www.earth.google.com>.

- Goovaerts, P (1998). «Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties». En: *Biology and Fertility of Soils* 27, págs. 315-334.
- Gräler, B., Pebesma, E.J. y Heuvelink, G. (2016). «Spatio-Temporal Interpolation using gstat». En: *The R Journal* 8 (1), págs. 204-218. URL: <https://journal.r-project.org/archive/2016-1/na-pebesma-heuvelink.pdf>.
- Gu, Z., Eils, R. y Schlesner, M. (2016). «Complex heatmaps reveal patterns and correlations in multidimensional genomic data». En: *Bioinformatics*.
- Haas, B. J., Gevers, D., Earl, A. M., Feldgarden, M., Ward, D. V., Giannoukos, G., Ciulla, D., Tabbaa, D., Highlander, S. K., Desantis, T. Z., Sodergren, E., Methe, B., The Human Consortium and Microbiome, Petrosino, J. F., Knight, R. y Birren, B. W. (2011). «Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons». En: *Genome research* 21, págs. 494-504. DOI: [10.1101/gr.112730.110](https://doi.org/10.1101/gr.112730.110). URL: [www.genome.org](http://www.genome.org).
- Hahnke, R. L., Meier-kolthoff, J. P., García-lópez, M., Mukherjee, S., Huntemann, M., Ivanova, N. N., Woyke, T., Kyrpides, N. C., Klenk, H.-P. y Göker, M. (2016). «Genome-Based Taxonomic Classification of Bacteroidetes». En: *Frontiers in microbiology* 7.December, págs. 1-37. DOI: [10.3389/fmicb.2016.02003](https://doi.org/10.3389/fmicb.2016.02003).
- Hamamura, N., Itai, T., Liu, Y., Reysenbach, A. L., Damdinsuren, N. e Inskeep, W. P. (2014). «Identification of anaerobic arsenite-oxidizing and arsenate-reducing bacteria associated with an alkaline saline lake in Khovsgol , Mongolia». En: *Environmental microbiology reports* 6.5, págs. 476-482. DOI: [10.1111/1758-2229.12144](https://doi.org/10.1111/1758-2229.12144).
- Harrell Jr, F. E. y with contributions from Dupont, C. and many others (2016). *Hmisc: Harrell Miscellaneous*. R package version 4.0-2. URL: <https://CRAN.R-project.org/package=Hmisc>.

- Hatzenpichler, R. (2012). «Diversity, Physiology, and Niche Differentiation of Ammonia-Oxidizing Archaea». En: *AEM journal* 78.21, págs. 7501-7510. DOI: [10.1128/AEM.01960-12](https://doi.org/10.1128/AEM.01960-12).
- He, Y, Caporaso, J. G., Jiang, X., Sheng, H., Huse, S. M., Rideout, J. R., Edgar, R. C., Kopylova, E., Walters, W. A., Knight, R. y Zhou, H. (2015). «Stability of operational taxonomic units: an important but neglected property for analyzing microbial diversity». En: *Microbiome* 3.20, págs. 1-10. DOI: [10.1186/s40168-015-0081-x](https://doi.org/10.1186/s40168-015-0081-x).
- Hengl, T. (2009). *A Practical Guide to Geostatistical Mapping*. second. Amsterdam, Holland, pág. 293. ISBN: 978-90-9024981-0.
- Hollister, E. B., Engledow, A. S., Hammett, A. J. M., Provin, T. L., Wilkinson, H. H. y Gentry, T. J. (2010). «Shifts in microbial community structure along an ecological gradient of hypersaline soils and sediments». En: *The ISME Journal* 4.6, págs. 829-838. ISSN: 1751-7362. DOI: [10.1038/ismej.2010.3](https://doi.org/10.1038/ismej.2010.3). URL: <http://dx.doi.org/10.1038/ismej.2010.3>.
- Huang, P.M., Yuncong, L. y Sumner, M. E., eds. (2012). *Handbook of Soil Sciences: Resource management and Environmental impacts*. Segunda ed. Boca Ratón, FL: CRC Press, págs. 1-34. ISBN: 9781439803080. DOI: [10.13140/2.1.1923.1047](https://doi.org/10.13140/2.1.1923.1047).
- Inskip, W. P., Jay, Z. J., Tringe, S. G., Herrgård, M. J. y Rusch, D. B. (2013). «The YNP metagenome project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem». En: *Frontiers in microbiology* 4.67, págs. 1-15. DOI: [10.3389/fmicb.2013.00067](https://doi.org/10.3389/fmicb.2013.00067).
- Isaaks, E. H. y Srivastava, R. M. (1989). *Applied Geostatistics*. New York, USA: Oxford University press, pág. 589. ISBN: 0-19-505012-6.
- Islam, M.R., Sultana, T., Melvin-Joe, M., Cho, J. C. y Sa, T. (2012). «Comparisons of direct extraction methods of microbial DNA from different paddy soils». En: *Saudi Journal of Biological Sciences* 19.3, págs. 337-342. ISSN: 1319562X. DOI: [10.1016/j.sjbs.2012.03.001](https://doi.org/10.1016/j.sjbs.2012.03.001).

1016/j.sjbs.2012.04.001. URL: <http://dx.doi.org/10.1016/j.sjbs.2012.04.001>.

Jongman, R. H., Ter Braak, C. J. F. y Van Tongeren, O. F. R., eds. (2007). *Data analysis in community and landscape ecology*. New York, USA: Cambridge University Press, pág. 321. ISBN: 9780521475747. URL: [http://books.google.co.uk/books/about/Data{\\\\_}analysis{\\\\_}in{\\\\_}community{\\\\_}and{\\\\_}landscape.html?id=NBBqERya0ssC{\\&}redir{\\\\_}esc=y](http://books.google.co.uk/books/about/Data{\\_}analysis{\\_}in{\\_}community{\\_}and{\\_}landscape.html?id=NBBqERya0ssC{\\&}redir{\\_}esc=y).

Kahle, D. y Wickham, H. (2013). «ggmap: Spatial Visualization with ggplot2». En: *The R Journal* 5.1, págs. 144-161. URL: <http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf>.

Kalwasinska, A., Felföldi, T., Szabó, A., Deja-sikora, E., Kosobucki, P. y Walczak, M. (2017). «Microbial communities associated with the anthropogenic , highly alkaline environment of a saline soda lime, Poland». En: *Antonie van Leeuwenhoek* 110, págs. 945-962. DOI: [10.1007/s10482-017-0866-y](https://doi.org/10.1007/s10482-017-0866-y).

Kassambara, A. y Mundt, F. (2017). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. R package version 1.0.4. URL: <https://CRAN.R-project.org/package=factoextra>.

Konstantinidis, K. T. y Tiedje, J. M. (2005). «Towards a Genome-Based Taxonomy for Prokaryotes». En: *Journal of bacteriology* 187.18, págs. 6258-6264. DOI: [10.1128/JB.187.18.6258](https://doi.org/10.1128/JB.187.18.6258).

Lauber, C. L., Ramirez, K. S., Aanderud, Z., Lennon, J. y Fierer, N. (2013). «Temporal variability in soil microbial communities across land-use types.» En: *The ISME journal* 7.8, págs. 1641-50. ISSN: 1751-7370. DOI: [10.1038/ismej.2013.50](https://doi.org/10.1038/ismej.2013.50). URL: <http://www.ncbi.nlm.nih.gov/pubmed/23552625>.

Lavelle, P., Spain, A. V., (CSIRO Land y Water) (2003). *Soil Ecology*. New York, USA: Kluwer Academic Publishers, pág. 677. ISBN: EBOOK ISBN: 0-306-48162-6. DOI: [10.1038/455724a](https://doi.org/10.1038/455724a).

- Lê, S., Josse, J. y Husson, F. (2008). «FactoMineR: A Package for Multivariate Analysis». En: *Journal of Statistical Software* 25.1, págs. 1-18. DOI: [10.18637/jss.v025.i01](https://doi.org/10.18637/jss.v025.i01).
- Legendre, P. y Legendre, L. (2012). *Numerical Ecology*. Tercera ed. Elsevier B.V., pág. 992.
- Li, S., Hua, Z., Huang, L., Li, J., Shi, S., Chen, L. y Kuang, J. (2014). «Microbial communities evolve faster in extreme environments». En: *Scientific reports* 4.6205, págs. 1-9. DOI: [10.1038/srep06205](https://doi.org/10.1038/srep06205).
- Liu, J., Sui, Y., Yu, Z., Shi, Y., Chu, H., Jin, J., Liu, X. y Wang, G. (2014). «High throughput sequencing analysis of biogeographical distribution of bacterial communities in the black soils of northeast China». En: *Soil Biology and Biochemistry* 70, págs. 113-122. ISSN: 00380717. DOI: [10.1016/j.soilbio.2013.12.014](https://doi.org/10.1016/j.soilbio.2013.12.014). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0038071713004471>.
- Lozupone, C. A., Hamady, M., Kelley, S. T. y Knight, R. (2007). «Quantitative and qualitative alpha diversity measures lead to different insights into factors that structure microbial communities». En: *Applied and Environmental Microbiology* 73.5, págs. 1576-1585. ISSN: 00992240. DOI: [10.1128/AEM.01996-06](https://doi.org/10.1128/AEM.01996-06).
- Ma, B. y Gong, J. (2013). «A meta-analysis of the publicly available bacterial and archaeal sequence diversity in saline soils». En: *World Journal of Microbiology and Biotechnology* 29.12, págs. 2325-2334. DOI: [10.1007/s11274-013-1399-9](https://doi.org/10.1007/s11274-013-1399-9).
- Magurran, A.E. y McGill, B.J., eds. (2011). *Biological Diversity: Frontiers in Measurement and Assessment*. Primera Ed. Oxford y New York: Oxford University Press., pág. 182. ISBN: 978-0-19-958066-8 (hc); 978-0-19-958067-5 (pb). 2011.
- Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Adams-Krumins, J., Kuske, C. R., Morin, P. J., Naeem, S., Ovreås, L., Reysenbach, A-L., Smith, V. H. y Staley, J. T. (2006). «Microbial biogeography: putting microorganisms on the map.» En:

- Nature reviews. Microbiology* 4. February, págs. 102-112. ISSN: 1740-1526. DOI: [10.1038/nrmicro1341](https://doi.org/10.1038/nrmicro1341).
- Marusenko, Y., Bates, S. T., Anderson, I., Johnson, S. L., Soule, T. y García-Pichel, F. (2013). «Ammonia-oxidizing archaea and bacteria are structured by geography in biological soil crusts across North American arid lands». En: *Ecological processes* 2.9, págs. 1-10. DOI: [10.1186/2192-1709-2-9](https://doi.org/10.1186/2192-1709-2-9).
- Mcdonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., Desantis, T. Z., Probst, A., Andersen, G. L., Knight, R. y Hugenholtz, P. (2012). «An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea». En: *The ISME Journal* 6.3, págs. 610-618. ISSN: 1751-7362. DOI: [10.1038/ismej.2011.139](https://doi.org/10.1038/ismej.2011.139). URL: <http://dx.doi.org/10.1038/ismej.2011.139>.
- Moeys, J. (2016). *soiltexture: Functions for Soil Texture Plot, Classification and Transformation*. R package version 1.4.1. URL: <https://CRAN.R-project.org/package=soiltexture>.
- Moreno, C.E. (2001). *Métodos para medir la biodiversidad*. Ed. por CYTED, ORCYT-UNESCO y Sociedad Entomológica Aragonesa. Primera Ed. Vol. 1, pág. 84. ISBN: 1576 – 9526(colección); 84 – 922495 - 2 – 8 (volumen). DOI: [10.1371/journal.pone.0103709](https://doi.org/10.1371/journal.pone.0103709). eprint: [arXiv:1011.1669v3](https://arxiv.org/abs/1011.1669v3).
- Morris, E. K., Caruso, T., Buscot, F., Fischer, M., Hancock, C., Maier, T. S., Meiners, T., Müller, C., Obermaier, E., Prati, D., Socher, S. A., Sonnemann, I., Wäschke, N., Wubet, T., Wurst, S. y Rillig, M. C. (2014). «Choosing and using diversity indices: Insights for ecological applications from the German Biodiversity Exploratories». En: *Ecology and Evolution* 4.18, págs. 3514-3524. ISSN: 20457758. DOI: [10.1002/ece3.1155](https://doi.org/10.1002/ece3.1155). arXiv: [NIHMS150003](https://arxiv.org/abs/NIHMS150003).

- Muñoz, R., Rosselló Móra, R. y Amann, R. (2016). «Revised phylogeny of Bacteroidetes and proposal of sixteen new taxa and two new combinations including Rhodothermaeota phyl. nov.» En: *Systematic and Applied Microbiology* 39.5, págs. 281-296. ISSN: 0723-2020. DOI: <https://doi.org/10.1016/j.syapm.2016.04.004>. URL: <http://www.sciencedirect.com/science/article/pii/S0723202016300182>.
- Navarro-Noya, Y. E., Gómez-Acata, S., Montoya-Ciriaco, N., Rojas-Valdez, A., Suárez-Arriaga, M. C., Valenzuela-Encinas, C., Jiménez-Bueno, N., Verhulst, N., Govaerts, B. y Dendooven, L. (2013). «Relative impacts of tillage, residue management and crop-rotation on soil bacterial communities in a semi-arid agroecosystem». En: *Soil Biology & Biochemistry* 65.May, págs. 86-95. ISSN: 00380717. DOI: [10.1016/j.soilbio.2013.05.009](https://doi.org/10.1016/j.soilbio.2013.05.009). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0038071713001818>.
- Navarro-Noya, Y. E., Valenzuela-Encinas, C., Sandoval-Yuriar, A., Jiménez-Bueno, N. G., Marsch, R. y Dendooven, L. (2015). «Archaeal communities in a heterogeneous Hypersaline-Alkaline soil». En: *Archaea* 2015.May, págs. 1-13. ISSN: 14723646. DOI: [10.1155/2015/646820](https://doi.org/10.1155/2015/646820).
- Nguyen, N., Warnow, T., Pop, M. y White, B. (2016). «A perspective on 16S rRNA operational taxonomic unit clustering using sequence similarity». En: *Npj Biofilms and Microbiomes* 2.2, págs. 1-8. DOI: [10.1038/npjbio](https://doi.org/10.1038/npjbio).
- Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, H. M. H., Szoecs, E. y Wagner, H. (2017). *vegan: Community Ecology Package*. R package version 2.4-3. URL: <https://CRAN.R-project.org/package=vegan>.
- Oren, A. (2014). «Taxonomy of halophilic Archaea : current status and future challenges». En: *Extremophiles* 18, págs. 825-834. DOI: [10.1007/s00792-014-0654-9](https://doi.org/10.1007/s00792-014-0654-9).

- Oren, A., Arahal, D. R. y Ventosa, A. (2009). «Emended descriptions of genera of the family Halobacteriaceae». En: *International Journal of Systematic and Evolutionary Microbiology* 59.May, págs. 637-642. DOI: [10.1099/ijs.0.008904-0](https://doi.org/10.1099/ijs.0.008904-0).
- Pasternak, Z., Al-Ashhab, A., Gatica, J., Gafny, R., Avraham, S., Minz, D., Gillor, O. y Jurkevitch, E. (2013). «Spatial and temporal biogeography of soil microbial communities in arid and semiarid regions». En: *PLOS one* 8.7, págs. 1-9. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0069705](https://doi.org/10.1371/journal.pone.0069705). URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3724898&tool=pmcentrez&rendertype=abstract>.
- Pebesma, E. J. y Bivand, R. S. (2005). «Classes and methods for spatial data in R». En: *R News* 5.2, págs. 9-13. URL: <https://CRAN.R-project.org/doc/Rnews/>.
- Pebesma, E.J. (2004). «Multivariable geostatistics in S: the gstat package». En: *Computers and Geosciences* 30, págs. 683-691.
- Philippot, L., Bru, D., Saby, N. P. A., Cuhel, J., Arrouays, D., Simek, M. y Hallin, S. (2009). «Spatial patterns of bacterial taxa in nature reflect ecological traits of deep branches of the 16S rRNA bacterial tree.» En: *Environmental microbiology* 11.12, págs. 3096-3104. ISSN: 1462-2920. DOI: [10.1111/j.1462-2920.2009.02014.x](https://doi.org/10.1111/j.1462-2920.2009.02014.x). URL: <http://www.ncbi.nlm.nih.gov/pubmed/19638171>.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Regan, K. M., Nunan, N., Boeddinghaus, R. S., Baumgartner, V., Berner, D., Boch, S., Oelmann, Y., Overmann, J., Prati, D., Schloter, M., Schmitt, B., Sorkau, E., Stefens, M., Kandeler, E. y Marhan, S. (2014). «Seasonal controls on grassland microbial biogeography: Are they governed by plants, abiotic properties or both?» En: *Soil Biology and Biochemistry* 71, págs. 21-30. ISSN: 0038-0717. DOI: [10.1016/](https://doi.org/10.1016/)



- [j.soilbio.2013.12.024](http://dx.doi.org/10.1016/j.soilbio.2013.12.024). URL: <http://dx.doi.org/10.1016/j.soilbio.2013.12.024>.
- Ribeiro Jr, P. J. y Diggle, P.J. (2016). *geoR: Analysis of Geostatistical Data*. R package version 1.7-5.2. URL: <https://CRAN.R-project.org/package=geoR>.
- Rideout, J.R., He, Y., Navas-molina, J. A., Walters, W. A., Ursell, L. K., Gibbons, S. M., Chase, J., Mcdonald, D., Gonzalez, A., Robbins-pianka, A. y Clemente, J. C. (2014). «Subsampled open-reference clustering creates consistent , comprehensive OTU definitions and scales to billions of sequences». En: *PeerJ* 2.e545. DOI: [10.7717/peerj.545](https://doi.org/10.7717/peerj.545).
- RStudio Team (2012). *RStudio: Integrated Development Environment for R*. RStudio, Inc. Boston, MA. URL: <http://www.rstudio.com/>.
- Ruiz-Romero, E., Valenzuela-Encinas, C., López-Ramírez, M. P., Coutiño-Coutiño, M. A., Marsch, R. y Dendooven, L. (2013). «*Natronorubrum texcoconense* sp. nov., a haloalkaliphilic archaeon isolated from soil of the former lake Texcoco (Mexico)». En: *Archaea Microbiology* 195.2, págs. 145-151. ISSN: 14665026. DOI: [10.1099/ijs.0.053629-0](https://doi.org/10.1099/ijs.0.053629-0).
- Sambrook, J. y Russell, D.W. (2001). *Molecular Cloning: A Laboratory Manual*. Molecular Cloning: A Laboratory Manual v. 1. Cold Spring Harbor Laboratory Press. ISBN: 9780879695774.
- Sánchez, O., Garrido, L., Forn, I., Massana, R., Maldonado, M.I. y Mas, J. (2011). «Molecular characterization of activated sludge from a seawater-processing wastewater treatment plant». En: *Microbial biotechnology* 4.5, págs. 628-642. DOI: [10.1111/j.1751-7915.2011.00256.x](https://doi.org/10.1111/j.1751-7915.2011.00256.x).
- Santini, T. C., Warren, L. A. y Kendra, E. (2015). «Microbial diversity in engineered haloalkaline environments shaped by shared geochemical drivers observed in natural analogues». En: *Applied and Environmental Microbiology* 81.15, págs. 5026-5036. DOI: [10.1128/AEM.01238-15](https://doi.org/10.1128/AEM.01238-15).

- Simachew, A., Lanzén, A., Gessesse, A. y Øvreås, L. (2015). «Prokaryotic community diversity along an increasing salt gradient in a soda ash concentration pond». En: *Microbial Ecology* 71.2, págs. 326-338. DOI: [10.1007/s00248-015-0675-7](https://doi.org/10.1007/s00248-015-0675-7).
- Sullivan, T. S., McBride, M. B. y Thies, J. E. (2013). «Soil bacterial and archaeal community composition reflects high spatial heterogeneity of pH, bioavailable Zn, and Cu in a metalliferous peat soil». En: *Soil Biology and Biochemistry* 66, págs. 102-109. ISSN: 0038-0717. DOI: [10.1016/j.soilbio.2013.06.021](https://doi.org/10.1016/j.soilbio.2013.06.021). URL: <http://dx.doi.org/10.1016/j.soilbio.2013.06.021>.
- Taiyun, W. y Simko, V. (2016). *corrplot: Visualization of a Correlation Matrix*. R package version 0.77. URL: <https://CRAN.R-project.org/package=corrplot>.
- Tanner, M. A., Everett, C. L., Coleman, W. J., Yang, M. M. y Youvan, D. C. (2000). «Complex Microbial Communities Inhabiting Sulfide-rich Black Mud from Marine Coastal Environments». En: *Biotechnology* 8.408, págs. 1-16.
- Tejeda-Agredano, M. C., Gallego, S., Vila, J., Grifoll, M., Ortega-calvo, J. J. y Cantos, M. (2013). «Influence of the sun flower rhizosphere on the biodegradation of PAHs in soil». En: *Soil Biology and Biochemistry* 57, págs. 830-840. DOI: [10.1016/j.soilbio.2012.08.008](https://doi.org/10.1016/j.soilbio.2012.08.008).
- Thies, J.E. (2007). «4 - Molecular Methods for Studying Soil Ecology». En: *Soil Microbiology, Ecology and Biochemistry (Third Edition)*. Ed. por A. P. Eldor. Third Edition. San Diego: Academic Press, págs. 85 -118. ISBN: 978-0-12-546807-7. DOI: <https://doi.org/10.1016/B978-0-08-047514-1.50008-1>. URL: <https://www.sciencedirect.com/science/article/pii/B9780080475141500081>.
- Tinsley, J., Taylor, T.G. y Moore, J.H. (1951). «The determination of carbon dioxide derived from carbonates in agricultural and biological materials». En: *The Analyst* 76, págs. 300-310. ISSN: 0003-2654. DOI: [10.1039/an9517600300](https://doi.org/10.1039/an9517600300).

- U.S. Department of Agriculture. URL: [https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2\\_054167](https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2_054167).
- Valenzuela-Encinas, C., Neria-González, I., Alcántara-Hernández, R. J., Enríquez-Aragón, J. A., Estrada-Alvarado, I., Hernández-Rodríguez, C., Dendooven, L. y Marsch, R. (2008). «Phylogenetic analysis of the archaeal community in an alkaline-saline soil of the former lake Texcoco (Mexico)». En: *Extremophiles* 12, págs. 247-254. ISSN: 14310651. DOI: [10.1007/s00792-007-0121-y](https://doi.org/10.1007/s00792-007-0121-y).
- Valenzuela-Encinas, C., Alcántara-Hernández, R. J., Estrada-Alvarado, I., Zavala-Díaz de la Serna, F.J., Dendooven, L. y R. Marsch (2012). «The Archaeal diversity and population in a drained alkaline saline soil of the former lake Texcoco (Mexico)». En: *Geomicrobiology Journal* 29.1, págs. 18-22. ISSN: 0149-0451. DOI: [10.1080/01490451.2010.520075](https://doi.org/10.1080/01490451.2010.520075). URL: <http://www.tandfonline.com/doi/abs/10.1080/01490451.2010.520075>.
- Vogt, J. C., Abed, R.M.M., Albach, D. C. y Palinska, K. A. (2017). «Bacterial and Archaeal Diversity in Hypersaline Cyanobacterial Mats Along a Transect in the Intertidal Flats of the Sultanate of Oman». En: *Microbial Ecology* 75.2, págs. 331-347. DOI: [10.1007/s00248-017-1040-9](https://doi.org/10.1007/s00248-017-1040-9).
- Wallenius, K. (2011). «Microbiological characterisation of soils: Evaluation of some critical steps in data collection and experimental design». Tesis doct. University of Helsinki, pág. 51. ISBN: 9789521072550. URL: <http://ethesis.helsinki.fi/>.
- Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J. R. y Al, W.E.T. (2007). «Naive bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy». En: *Applied and Environmental Microbiology* 73.16, págs. 5261-5267. DOI: [10.1128/AEM.00062-07](https://doi.org/10.1128/AEM.00062-07).
- Wasserstrom, H., Kublik, S., Wasserstrom, R., Schulz, S., Schloter, M. y Steinberger, Y. (2017). *Bacterial community composition in costal dunes of the Mediterranean along*

- a gradient from the sea shore to the inland*. Inf. téc. 40266, págs. 1-7. DOI: [10.1038/srep40266](https://doi.org/10.1038/srep40266). URL: <http://dx.doi.org/10.1038/srep40266>.
- Webster, G., Sullivan, L. A. O., Meng, Y., Williams, S., Sass, A. M., Watkins, A. J., Parkes, R. J. y Weightman, A. J. (2015). «Archaeal community diversity and abundance changes along a natural salinity gradient in estuarine sediments». En: *FEMS Microbiology Ecology* 91.2, págs. 1-18. DOI: [10.1093/femsec/fiu02](https://doi.org/10.1093/femsec/fiu02).
- Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J. R., Vázquez-Baeza, Y., Birmingham, A. e Hyde, E. R. (2017). «Normalization and microbial differential abundance strategies depend upon data characteristics». En: *Microbiome* 5.27, págs. 1-18. DOI: [10.1186/s40168-017-0237-y](https://doi.org/10.1186/s40168-017-0237-y).
- Wessén, E., Söderström, M., Stenberg, M., Bru, D., Hellman, M., Welsh, A., Thomssen, F., Klemmedtson, L., Philippot, L. y Hallin, S. (2011). «Spatial distribution of ammonia-oxidizing Bacteria and Archaea across a 44-hectare farm related to ecosystem functioning.» En: *The ISME journal* 5.7, págs. 1213-25. ISSN: 1751-7370. DOI: [10.1038/ismej.2010.206](https://doi.org/10.1038/ismej.2010.206). URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3146283&tool=pmcentrez&rendertype=abstract>.
- Winston, F. y Hoffman, C. S. (1987). «A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of Escherichia coli.» En: *Gene* 57, págs. 267-272. ISSN: 03781119. DOI: [10.1016/0378-1119\(87\)90131-4](https://doi.org/10.1016/0378-1119(87)90131-4).
- Youssef, N. H., Couger, M. B., McCully, A. L., Guerrero-Criado, A. E. y Elshahed, M. S. (2015). «Assessing the global phylum level diversity within the bacterial domain: A review». En: *Journal of Advanced Research* 1, págs. 269-282. DOI: [10.1016/j.jare.2014.10.005](https://doi.org/10.1016/j.jare.2014.10.005).

---

Zhang, Y., Wang, H., Yang, L., Zhou, X., Zhi, X., Duan, Y., Xiao, M, Zhang, Y. y Li, W. (2016). «Egibacter rhizosphaerae gen. nov., sp. nov., an obligately halophilic, facultatively alkaliphilic actinobacterium and proposal of Egibaceraceae fam. nov. and Egibacterales ord. nov.» En: *International Journal of Systematic and Evolutionary Microbiology* 66.May, págs. 283-289. DOI: [10.1099/ijsem.0.000713](https://doi.org/10.1099/ijsem.0.000713).